USGS

science for a changing world

Techniques of Water-Resources Investigations of the United States Geological Survey

Chapter B4

REGRESSION MODELING OF GROUND-WATER FLOW

By Richard L. Cooley and Richard L. Naff

Book 3
APPLICATIONS OF HYDRAULICS

Preface

Scientists and engineers have been using ground-water flow models to study ground-water flow systems for more than 20 years. The basic modeling process seems to be relatively straightforward. Initially, a sound conceptual model is formed and is translated into a tractable, mathematical model. Contributing to (and following) this conceptualization process is the collection of field information, such as (1) location and extent of hydrostratigraphic units, recharge areas, discharge areas, and system boundaries; (2) hydraulic head measurements; and (3) pumping discharges. These data form the basis for input to the flow model. Finally, the model is run, and the desired information such as head distribution or flux rates is extracted. However, people engaged in modeling usually observe that two pervasive problems considerably complicate the situation. One problem is that good, general methods of measuring (or computing) some of the variables that characterize the flow system and its geologic framework do not exist. One example is measurement of ground-water recharge. No direct ways of measuring recharge exist, and the accuracy of indirect methods is often unknown. Furthermore, many indirect methods are applicable only to unique situations. The second problem relates to errors in the measurements and their propagation into model results. No error-free measurement (or computation) methods for obtaining data on the flow system exist. Thus, even the variables that can be estimated will contribute to error, so that model results will always be unreliable to some extent. As a consequence of these two problems, measurement (or computation) of the necessary input variables, application of them to an adequate model, and calculation of the desired results to an acceptable accuracy generally are not possible. Other methods that recognize and deal with the problems of incompleteness and (or) inaccuracy of data must also be applied. The present text has been designed to teach these methods to scientists and engineers engaged in ground-water modeling.

The basic methodology is multiple, nonlinear regression, in which the regression model is some type of ground-water flow model. As seen subsequently, this methodology is consistent with known aspects of the physical systems to be analyzed and requires relatively few assumptions. Even though the present text is directed specifically toward ground-water modeling, the procedures to be discussed are applicable to a number of different types of modeling problems. Thus, the methods are usually discussed in a general context; in other words, without reference to any specific model.

Material in the present text evolved from notes developed for training courses in parameter estimation for ground-water flow models taught by the authors and others at the U.S. Geological Survey National Training Center, Denver Federal Center, Lakewood, Colo. The philosophy of these courses, and of this text, is to teach general methods that are applicable to a wide range of problems and to teach these methods in sufficient depth so that students can apply them to many problem situations not considered in the courses or text.

The main body of the text is organized into six major sections. The first section is an introduction that discusses the general topic of modeling ground-water flow. This section shows that ground-water modeling problems are an incomplete combination of direct-type problems (solution for hydraulic head given values of flow system and framework variables) and inverse-type problems (solution for flow system and framework variables given values of hydraulic head) that commonly require solution by optimization procedures which give the best fit between observed and calculated results. Because the specific optimization approach employed here is regression and regression procedures are based on statistical concepts, the second section is included to provide the student with the necessary statistical background material. It is not designed to be an exhaustive review of basic statistics; rather, it presents material essential to understanding the following sections. The third section presents detailed material on linear and nonlinear regression. Although most

of the material on linear regression is fairly standard, some of the material on nonlinear regression is not. In particular, specific modifications presented to induce convergence of the iterative solution procedure for nonlinear regression have not, to the writers' knowledge, been presented elsewhere in the form given here. The fourth section applies the nonlinear regression method to the specific problem of developing a general finite-difference model of steady-state ground-water flow. In the fifth section, statistical procedures are given to analyze and use general linear and nonlinear regression models. The tests and analytical procedures presented are not exhaustive; they are the ones that the writers have found to be most useful for analyzing the real systems examined to date. The sixth section is designed to be supplemental to the preceding sections. Specialized procedures presented include nonlinear regression for models that cannot be solved directly for the dependent variable, a measure of model nonlinearity called Beale's measure, and a statistical test for compatibility of prior information on parameters and parameter estimates derived from sample (observed head) information.

A number of exercises have been included, and a complete discussion of the answers can be found in the seventh major section at the end of the text. These problems exercise the student on nearly all methods presented. In addition, three computer programs are documented and listed: the program for nonlinear-regression solution of ground-water flow problems of section four, a program to calculate Beale's measure, and a program to calculate simulated errors in computed dependent variables such as hydraulic head.

The mathematical background necessary to use this text includes basic mathematics through differential and integral calculus, including partial derivatives, and matrix algebra. A background in elementary statistics would be useful but is not essential. In addition, a sound knowledge of ground-water hydrology and ground-water flow modeling are needed to effectively apply the methods presented.

References for cited material are given at the end of each major section. Good supplemental sources for the unreferenced material not peculiar to this text are presented as "Additional Reading" at the end of each reference list. It is expected that students who have difficulty with the material in this text will consult the more expanded developments in these supplemental sources.

TECHNIQUES OF WATER-RESOURCES INVESTIGATIONS OF THE UNITED STATES GEOLOGICAL SURVEY

The U.S. Geological Survey publishes a series of manuals describing procedures for planning and conducting specialized work in water-resources investigations. The manuals published to date are listed below and may be ordered by mail from the U.S. Geological Survey, Books and Open-File Reports, Federal Center, Box 25425, Denver, Colorado 80225 (an authorized agent of the Superintendent of Documents, Government Printing Office).

Prepayment is required. Remittance should be sent by check or money order payable to U.S. Geological Survey. Prices are not included in the listing below as they are subject to change. Current prices can be obtained by writing to the USGS, Books and Open File Reports. Prices include cost of domestic surface transportation. For transmittal outside the U.S.A. (except to Canada and Mexico) a surcharge of 25 percent of the net bill should be included to cover surface transportation. When ordering any of these publications, please give the title, book number, chapter number, and "U.S. Geological Survey Techniques of Water-Resources Investigations."

CONTENTS

COMPUTER DISKETTE INFORMATION

This report contains a computer diskette (in the pocket at the back of the report) that has the source codes and data sets to be used with the report. The source codes were developed using the Microsoft Fortran Compiler, Version 3.3, with the DOS 2.0 operating system on an IBM PC/XT computer with the IBM 8088 Math Coprocessor and 256KB memory. Except for the OPEN statements near the beginning of the codes, Fortran 66 was used throughout to make the codes as machine independent as possible. For more information concerning the contents of the diskette, insert the diskette in drive A of the computer, type README, and push the enter key.

METRIC CONVERSION FACTORS

For those readers who prefer to use metric units, conversion factors for terms used in this report are listed below:

Multiply	By	To obtain
foot (ft)	0.3048	meter (m)
square foot (ft^2)	0.09290	square meter (m^2)
foot per day (ft/d)	0.3048	meter per day (m/d)
square foot per second (ft^2/s)	0.09290	square meter per second (m^2/s)
square foot per day (ft^2/d)	0.09290	square meter per day (m^2/d)
gallon per day per foot (gal/d/ft)	0.01242	square meter per day (m^2/d)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
cubic foot per day (ft^3/d)	0.02832	cubic meter per day (m^2/d)
gallon per minute (gal/min)	0.00006309	cubic meter per second (m^3/s)
ounces per ton	31.25	grams per metric ton[1]

[1] 1 metric ton = 1 megagram

REGRESSION MODELING OF GROUND-WATER FLOW

By Richard L. Cooley and Richard L. Naff

1 Introduction

1.1 Flow Equation and Boundary Conditions

The most general form of the ground-water flow equation that we consider here is given as

$$\frac{\partial}{\partial x}(T_{xx}\frac{\partial h}{\partial x})+\frac{\partial}{\partial y}(T_{yy}\frac{\partial h}{\partial y})+R(H-h)+W$$

$$+\sum_{\ell=1}^{N}\delta(x-a_\ell)\delta(y-b_\ell)Q_\ell=S\frac{\partial h}{\partial t} \qquad (1.1\text{-}1)$$

where

$T_{\xi\xi}$ $(x,y)=$ transmissivity $(K_{\xi\xi}b)$ in the $\xi=x$ or y direction;

$K_{\xi\xi}$ $(x,y)=$ hydraulic conductivity of the aquifer in the ξ direction;

$b(x,y)=$ thickness of the aquifer;

$R(x,y)=$ hydraulic conductance (hydraulic conductivity divided by thickness) of sediments underlying a stream or of an aquitard underlying or overlying the aquifer;

$W(x,y,t)=$ source-sink term (positive for a source), distributed areally;

$\sum_{\ell=1}^{N}\delta(x-a_\ell)\delta(y-b_\ell)Q_\ell=$Dirac delta designation for N wells, each one pumping at rate Q_ℓ (t) (positive for injection) and located at (a_ℓ,b_ℓ);

$S(x,y)=$ storage coefficient;

$h(x,y,t)=$ hydraulic head in the aquifer;

$H(x,y,t)=$ head at the stream bottom or at the distal side of the aquitard;

$x,y=$ Cartesian coordinates;

$t=$ time;

and T_{xx} and T_{yy} are continuous functions of x and y.

With suitable internal boundary conditions, the region can be zoned with respect to $T_{\xi\xi}$. Such boundary conditions involve head and specific discharge multiplied by thickness normal to the boundary (q_n) and can be stated for a boundary between $T_{\xi\xi}$ zones k and ℓ as

$$(h)_k=(h)_\ell \qquad (1.1\text{-}2)$$

$$(q_n)_k=(q_n)_\ell \qquad (1.1\text{-}3)$$

where $(\cdot)_k$ indicates that the quantity in parentheses is evaluated just within the k side of the boundary and similarly for ℓ. Zonation with respect to R, W, or S requires no internal boundary conditions.

External boundary conditions applying at the periphery of the domain being modeled are given as

$$T_{xx}\frac{\partial h}{\partial x}n_x+T_{yy}\frac{\partial h}{\partial y}n_y+\alpha h=\beta \qquad (1.1\text{-}4)$$

where $\alpha(x,y,t)$ and $\beta(x,y,t)$ are given functions, and $\{n_x(x,y),n_y(x,y)\}$ is the outward-pointing unit normal at the boundary. The sum of the first two terms is the flux q_B normal to the boundary (positive for outflow). Equation 1.1–4

1

incorporates the standard boundary conditions of specified flux (q_B) and specified head (h_B) but also allows for linear combinations to be given.

1.2 Types of Solutions

1.2.1 Direct Solution for Head

The classical problem of mathematical physics (and, by assumption, of ground-water hydrology) is to directly solve equations 1.1-1 through 1.1-4 for $h=h(x,y,t)$. Given that any specific problem is properly posed, such a solution will always exist. The conditions for properly posing a problem are the following.

1. The positions of all internal boundaries are known exactly. Examples of internal boundaries are abrupt changes in $T_{\xi\xi}$, R, S, or W; internal known flux (q_n) boundaries; and internal known head boundaries. Note that a river is often treated as either an internal known head boundary where the river is assumed to have no width, or a zone of differing R where each bank is a zone boundary.
2. The positions and types of all external boundaries are known exactly. External boundaries frequently are known flux (q_B) types or known head (h_B) types. Sometimes some linear combination is known.
3. Hydrogeologic variables $T_{\xi\xi}$, R, and S and hydrologic variables W and Q_ℓ are known at all points in the region.
4. All boundary-condition variables H, α, and β are known. The initial head (at $t=0$) is a boundary condition and must also be known.

Obviously, ground-water flow problems are not actually of the classical type because none of the conditions cited above ever are met exactly. Conditions 1 and 2 are often most closely fulfilled, but estimates (often crude) usually must suffice for the variables in conditions 3 and 4. Any errors in these input variables are propagated directly into the solution. However, reasonable (but incorrect) estimates of the variables can be shown to yield errors in predicted $h(x,y,t)$ that have the characteristic of being bounded (that is, they do not tend to plus or minus infinity). Also, as the errors in the input variables tend to zero, the errors in computed head do also.

1.2.2 Inverse Solution for Parameters

An inverse solution involves solving equations 1.1-1 through 1.1-4 for one or more of the variables $T_{\xi\xi}$, R, S, W, Q_ℓ, α, or β, over the region; these variables are termed <u>parameters</u> here. Because R, S, W, Q_ℓ, α, and β are not involved in derivatives, theoretically they may be solved for algebraically. Unless $T_{\xi\xi}$ is constant, it is involved in derivatives and, thus, must be obtained by solving a differential equation. To understand this, note that equation 1.1-1 may be rearranged to give

$$a\frac{\partial T_{xx}}{\partial x}+b\frac{\partial T_{yy}}{\partial y}+cT_{xx}+dT_{yy}-F=0 \quad (1.2-1)$$

where

$$a(x,y,t)=\frac{\partial h}{\partial x} \quad , \quad b(x,y,t)=\frac{\partial h}{\partial y} \quad ,$$

$$c(x,y,t)=\frac{\partial^2 h}{\partial x^2} \quad , \quad d(x,y,t)=\frac{\partial^2 h}{\partial y^2} \quad , \text{ and}$$

$$F(x,y,t)=S\frac{\partial h}{\partial t}-R(H-h)-W-\sum_{\ell=1}^{N}\delta(x-a_\ell)\delta(y-b_\ell)Q_\ell.$$

In general, if T_{xx} and T_{yy} are known, conditions for finding R, S, W, Q_ℓ, α, or β are:

1. Conditions 1 and 2 for the classical direct solution are met.
2. Head distribution $h(x,y,t)$ is known exactly.
3. The solution for the desired combination of parameters to be obtained is unique.

The latter condition is completely problem dependent. Because solution involves only algebraic manipulations, the condition reduces to the requirement that the system of algebraic equations involving the desired parameters has a unique solution. Generally, solution involves picking the required number of points spatially and through time to yield the necessary number of equations.

To find T_{xx} and T_{yy}, more conditions are

required than for finding R, S, W, Q_p, α, or β. These conditions are:

1. Conditions 1 and 2 for finding R, S, W, Q_p, α, or β must be met.
2. The direction of the velocity vector must be known everywhere, or T_{xx}/T_{yy} must be known everywhere, or quantities a, b, c, d, and F in equation 1.2-1 must be known at two (or more) points in time to give a unique solution to equation 1.2-1 written in the form of a pair of simultaneous differential equations. These requirements result because 1.2-1 is one equation in two unknowns. Hence, an additional relationship is required. If the velocity direction is known everywhere, then by employing Darcy's law the additional relationship is derived as

$$\frac{T_{xx}}{T_{yy}} = \frac{q_x b}{q_y a} \qquad (1.2\text{-}2)$$

where a, b, and q_x/q_y (the ratio of the x and y direction fluxes) are known.

3. If either the direction of the velocity vector or T_{xx}/T_{yy} is known, then either T_{xx} or T_{yy} must be known on a possibly discontinuous curve crossing all flowlines. If solution is to be obtained by solving a simultaneous pair of differential equations, then T_{xx} must be known on a possibly discontinuous curve that spans the range of y, and T_{yy} must be known on a possibly discontinuous curve that spans the range of x. These are extensions of the Cauchy boundary condition for a first-order differential equation involving a single dependent variable and are required for solution of the problem.
4. The function F in equation 1.2-1 must be known everywhere. This means that all quantities in F must be known or that a mathematical form for F can be assumed.

Ground-water flow modeling does not fit into the category of inverse solutions, although a significant part of most model studies is to find values of the parameters that allow values of calculated head to match those observed in the field. The difficulty is that the required conditions are almost never met. Head distribution is never known exactly because measurements do not exist at all points and, where these measurements do exist, they are not exact. Furthermore, some measure of $T_{\xi\xi}$ is virtually never available on the required curves, and information on directions of flow vectors for even scattered locations usually is nonexistent. Assumptions concerning zonations in which T_{xx}/T_{yy} and (or) T_{xx} and T_{yy} may be considered constant simplify the problem, but the fact that h must be known still remains.

Because the head distribution is not known exactly, coefficients a, b, c, d, and F in equation 1.2-1 are in error. Furthermore, head appears as a derivative in all of these quantities. Hence, any error in h is propagated into the inverse solution as a derivative of error. The effects of this propagation are often disastrous because, if ϵ_h is defined as error in head, $\epsilon_h \to 0$ does not imply that $\partial \epsilon_h / \partial \xi \to 0$. Also, $|\partial \epsilon_h / \partial \xi| \gg |\epsilon_h|$ is common, and it can happen that $|\partial \epsilon_h / \partial \xi| \to \infty$ even if ϵ_h is bounded. Therefore, the error in computed $T_{\xi\xi}$ (or other parameter) may not approach zero as $\epsilon_h \to 0$, and may, in fact, be quite large (Neuman, 1980, p. 342-344).

1.2.3 Solution Using Real Data

In the previous section, we argued that problems involving ground-water flow modeling of real field systems are neither of the classical nor inverse type, because the data necessary for the problems to be classified as either type are usually lacking. An estimate of the hydraulic head distribution based on measurements (that are in error with respect to the model) taken at selected points usually exist. Estimates of the parameters are usually either completely unknown or have been obtained by spot measurements, few of which are directly useful for construction of appropriate effective values for use in equation 1.1-1. That modeling problems in ground-water hydrology involve an incomplete combination of several types of data in which error and error propagation are important considerations is evident.

1.3 Sources of Error in Ground-Water Data

Uncertainty (or errors) in ground-water data may have many sources, and enumeration of all

possible sources would be a nearly impossible task. However, a consideration of some of the more important sources of error serves to illustrate the importance of the error component.

1.3.1 Sources of Error in Head Data

Some major potential sources of random-appearing error in head data with respect to the model (equations 1.1-1 through 1.1-4) are:

1. Areal ground-water models assume that the head used is the average over the vertical. However, wells may not be open over the entire interval modeled, and if they are, they may not measure the average. Flow into and (or) out of a well distorts the hydraulic head field in the vicinity of the well so that the recorded water level does not represent the average head.
2. Permeability varies from point to point, which causes water levels to vary from values they would have if permeability were uniform. However, models usually do not take this detailed variation into account. This phenomenon has been extensively studied during the last 10 years, and literature reviews are contained in the works of Dagan (1986) and Gelhar (1984, 1986).
3. Water levels measured in wells in use may contain unknown amounts of residual drawdown. In addition, unused wells may be near wells that are in use, with resulting unknown drawdown in the unused well.
4. Measurement of well-head elevation may be in error.

Actual total error from the above sources is highly problem dependent, but it is easy to imagine errors of several feet. It should be noted that measurement error in water levels was not mentioned as a major source of error because it commonly amounts to one- or two-tenths of one foot or less. Finally, major model error in equations 1.1-1 through 1.1-4 (for example, head dependence in one or more parameters or three-dimensional flow) was also not mentioned because error resulting from this source is bias and should be detected and eliminated by analysis of model results.

1.3.2 Sources of Error in Parameter Data

Because there are several different parameters to be considered, and each can be estimated or measured in several different ways, a large number of sources of error exist in parameter data. Model error is not considered here, but other types of bias are potentially important and are often difficult to detect. Some examples of errors in parameter data illustrating the nature of the problem are:

1. Too few estimates of parameters are available to compute stable estimates of statistics, such as mean and variance.
2. Results of point sampling are often biased because a large amount of data does not necessarily allow computation of nearly true or effective values of a parameter and its variance. For example, permeability values from core analyses often are not representative of regional values, because flow through large fractures is not reproduced by core analyses. Also, effective values of a parameter and its variability are usually not directly given by standard mean and variance formulas.
3. Transmissivities estimated from specific-capacity data collected by drillers are subject to numerous sources of error. Common sources include (1) mismeasuring water levels or pumping rates, (2) allowing the water level to recover after bailing, (3) clogging of the slots or screen, and (4) inaccurate reporting. There are so many sources of error that the errors may often appear to be random. A persistent source of bias results because drillers drill wells in favorable locations and only screen (or slot) the most productive zones.
4. Transmissivities and storage coefficients estimated from pumping-test analysis are subject to many of the same errors as above, but the more carefully controlled tests should reduce their frequency and magnitude. In addition, a single test may not be representative of an entire hydrostratigraphic unit.
5. Transmissivities estimated from lithological data are usually biased to an unknown degree.

1.4 Model Construction

Ground-water models are constructed by using the types of data alluded to in the previous section. Hence, measured or estimated parameter data, either reliable or complete enough to employ directly in a model to reproduce measured head data with an acceptable model fit, are rare. As a result, adjustment of parameter values, and sometimes basic model structure, is used to improve model fit. Two basic groups of methods currently in use to accomplish this are: (1) trial and error procedures and (2) optimization methods that minimize a formal objective function.

1.4.1 Trial and Error Methods

Trial and error is the method of repeated simulation until the calculated head distribution obtained with a reasonable set of parameters fits closely enough to satisfy the analyst. Sometimes an objective measure of goodness of fit, such as $\Sigma(h^{calc}-h^{obs})^2$, is used to aid the analyst in deciding whether or not a change in parameters (or model structure) has improved the overall model fit. However, no matter how the method is applied, it has several inherent critical deficiencies:

1. No methodology exists to guarantee that the simulations will proceed in a direction that could lead to the best set of parameters.
2. Determining when that best set has been reached is difficult.
3. No practical way of determining how many other sets of parameters could yield similar correspondence between h^{calc} and h^{obs} exists.
4. Deciding whether or not additional parameters or a more refined model would significantly improve model fit is difficult.
5. No way of quantitatively assessing the predictive reliability of the model exists.

A method of model construction that addresses these deficiencies would allow construction and use of a model with a much greater degree of confidence than that provided by trial and error methods. Hence, attention is turned to formal optimization procedures.

1.4.2 Formal Optimization Procedures

Optimization procedures utilize a formal criterion of goodness of fit, often called an objective function. This function is minimized (or sometimes maximized, depending on the form of the function) with respect to the parameters to yield an optimum or best-fit solution. Minimization (or maximization) sometimes is subject to certain other criteria regarding values that the parameters, or pertinent functions of the parameters related to the model, may take on. These criteria are called constraints.

Examples of objective functions are:

$$\sum_{\ell=1}^{n_s}(h_\ell-\hat{h}_\ell)^2 \ ,$$

$$\sum_{\ell=1}^{n_s}|h_\ell-\hat{h}_\ell| \ ,$$

$$\max_\ell |h_\ell-\hat{h}_\ell| \ , \text{ and}$$

$$\sum_{\ell=1}^{n_s} w_\ell(h_\ell-\hat{h}_\ell)^2 + \sum_{m=1}^{n_p} k_m(p_m-\hat{p}_m)^2$$

where

h_ℓ = observed head,
$\hat{h}_\ell$ = calculated head,
p_m = observed or prior estimate of a parameter,
$\hat{p}_m$ = calculated parameter value that, when used in the model, produces $\hat{h}_\ell$,
w_ℓ = weight related to the reliability of the observation h_ℓ,
k_m = similar weight applied to p_m,
n_s = number of observations of head, and
n_p = number of observations of parameters.

The last example is called a compound objective function because it contains both head and parameters explicitly. Note that minimization of each of the functions with respect to the parameters of the model produces a solution that is overall a best fit to the data, according to the objective function. If the signs of the

functions were changed, maximization would produce the same result.

Examples of constraints are:

$$p_m^L < \hat{p}_m < p_m^U ,$$

$$a\hat{p}_k + b\hat{p}_m + c\hat{p}_n = f ,$$

$$a\hat{p}_k + b\hat{p}_m + c\hat{p}_n < f ,$$

where a, b, c, and f are constants or known functions; superscript L refers to a lower limit; and superscript U refers to an upper limit. The best-fit solution obtained by minimizing (or maximizing) the appropriate objective function must simultaneously satisfy the appropriate constraints.

Because the solution obtained by an optimization procedure has known properties, it may be analyzed. The exact procedures used and the extent to which the model may be analyzed depend on the type of optimization method selected for use. Statistical regression procedures handle, on a probabilistic basis, the propagation of data errors (with respect to the model) into the estimates of parameters and predictive capability of the model. Methods have been developed for estimating parameters, testing assumptions made during development

of techniques, testing model fit, determining the reliability and significance of the model and the parameters contained in it, effecting corrective measures for violation of some assumptions, and estimating the reliability of predictions to be made with the model. These procedures and the statistical background necessary to apply them are detailed in the remainder of the text.

References Cited

Dagan, Gedeon, 1986, Statistical theory of groundwater flow and transport—pore to laboratory, laboratory to formation, and formation to regional scale: Water Resources Research, v. 22, no. 9, p. 120S-134S.

Gelhar, L.W., 1984, Stochastic analysis of flow in heterogeneous porous media, in Bear, Jacob, and Corapcioglu, M.Y., eds., Fundamentals of transport phenomena in porous media: Dordrecht, The Netherlands, Martinus Nijhoff, p. 673-720.

_____1986, Stochastic subsurface hydrology from theory to applications: Water Resources Research, v. 22, no. 9, p. 135S-145S.

Neuman, S.P., 1980, A statistical approach to the inverse problem of aquifer hydrology, 3—Improved solution method and added perspective: Water Resources Research, v. 16, no. 2, p. 331-346.

Additional Reading

Freeze, R.A., and Cherry, J.A., 1979, Groundwater: Englewood Cliffs, N.J., Prentice-Hall, p. 15-77, 356-359.

2 Review of Probability and Statistics

Casual observation of our environment indicates that many phenomena are not strictly predictable. We cannot, for instance, exactly say what the maximum air temperature at any particular location will be tomorrow, although we might be able to give a probable range. This probable range might be based on our past experience, which would enable us to say that tomorrow's high, considering the location and season, will probably fall within a specified interval. A more sophisticated forecasting model may enable us to reduce the range within which we think tomorrow's high will fall, but random elements in the forecasting procedure would preclude giving an exact answer. As another example of randomness, consider the toss of a coin. Prior to the toss, we can only give the possible outcomes, either a head or a tail, and, if the coin is fair, say that either have equal likelihood of occurring. However, this ability to state precisely that any future outcome of this experiment can, with equal probability, result in either a head or a tail is an important advantage over that offered for predicting tomorrow's maximum temperature. In this latter case, because of the complex nature of the processes resulting in tomorrow's maximum, the likelihood that we could give a precise statement concerning the probability that it will fall in our predicted interval is remote. Instead of attempting to untangle these complexities, we might opt to study the history of maximum temperatures at the location and annual date in question. By assuming that this history will extrapolate into the future (that is, that weather dynamics in future years will remain essentially unchanged from those in previous years), we could give an estimate of the likelihood that tomorrow's maximum will fall in a particular interval. However, tools need to be developed to carry out this investigation.

2.1 Basic Concepts

Randomness itself can be considered to be centered around an experiment; the outcome of the experiment will have a random quality attached to it. For example, in a coin-toss experiment, the outcome is dominated by the random element (either a head or a tail). On the other hand, many experiments have a large deterministic factor. For example, in a chemical titration experiment we measure the unknown and the amount of titrant used, then calculate the amount of a specific substance in the unknown. However, measurement error creeps into our technique, and results vary from realization to realization of the experiment. Some experiments, such as annual, peak river flows, are not ours to perform but only to observe. This experiment is an example of an event in nature that has a large random component which nature provides. As we attempt to measure these flows, we introduce additional randomness, which we generally ignore. Hydraulic conductivities measured from core samples are similar to peak flows; nature has already provided for randomness, which is constrained by certain deterministic factors, such as type of source material, distance of transport, climate, and diagenesis. Again, for every realization of this experiment, measurement error is introduced, which may not be small.

All possible outcomes of an experiment are known as its sample space. The sample space of a coin-toss experiment consists of either a head (H) or tail (T):

$$S = \{H, T\}.$$

If the experiment consists of the toss of two coins, then the sample space consists of

$$S_1 = \{(H,H),(H,T),(T,H),(T,T)\}.$$

On the other hand, if we are only interested in the total number of heads which might result from a single toss of two coins, we could define the experiment as this sum, which would result in the sample space

$$S_2 = \{0,1,2\}.$$

In the case of S_1, every member of the sample space is equally likely to occur, whereas for S_2, a one is twice as likely to occur as either zero or two, provided that the coin is fair.

The sample space for a hydraulic-conductivity experiment could be defined as all positive real

numbers; that is, measurements from cores might result in values (outcomes) which could be *as small as zero or, if we stretch our imagina-tions, infinitely large.* This space could be con-sidered to be a continuous equivalent of the S_2 space for the two-coin experiment. That is, a porous medium is an extremely complex random process itself. By conducting hydraulic-conductivity measurements on cores, we quan-tify this randomness in much the same way that counting heads quantifies an outcome of the two-coin experiment. However, by quantifying the randomness of the porous medium in this manner, we have never investigated the possi-ble existence of more basic, perhaps nonnumeric sample spaces similar to S_1 of the two-coin ex-periment for a porous medium. Even if we were to discover the existence of such a space, we would then need to find a rule, or algorithm, which would allow us to connect the two spaces. We shall not worry about the possibilities of an S_1-like space for many processes; however, when they are available, they provide an ex-cellent mechanism for investigating the charac-teristics of S_2-like spaces.

An <u>event</u> is defined as any subset of the sample space. The investigator is usually inter-ested in the relative frequency of occurrence of an event. In the case of the S_2 space and the two-coin experiment, it is apparent that half the time a realization experiment should result in a one. This event is equivalent to the event in the S_1 space corresponding to the union of (H,T) and (T,H), which occurs with a relative frequency of one-half. Thus, the relative fre-quency of a head occurrence for the two-coin ex-periment is not dependent upon the definition of the sample space, but on the basic ran-domness controlling the experiment.

The investigator is frequently confronted with the problem of needing a numerical result for the outcome of a random, but not necessarily numerical, experiment. In the case of coin-toss experiments, the basic outcome is seen to be a particular arrangement of heads and (or) tails. By assigning a head a value of one and a tail a value of zero and then summing, it is possible to translate these basic results into something measurable. This process of assigning a numerical value to a nonnumerical outcome leads to the definition of a <u>random variable.</u>

Definition: A random variable is a function whose value is a real number determined by each element in a sample space.

When the outcome of the experiment is numerical, then this result can be considered to be the random variable (this statement is merely a special case of the above definition). From the above definition, we see that a mathematical transformation of a random variable is also a random variable. (Throughout this review, a random variable is indicated by an upper case English or Greek letter, whereas a value that it may take on is indicated by another letter, usually lower case of the same type as used for the random variable.)

The concepts of a random experiment, sample space, and random variable are flexible. For in-stance, if in the case of the toss of two coins, the experiment is defined as the total number of heads appearing, then the S_2 sample space is an automatic result, and the random variable can also be considered to be this result. How-ever, if the experiment is defined to be the ar-rangement of heads and (or) tails resulting from a toss (that is, the S_1 space), then the same ef-fect can be obtained by letting the random variable over the S_1 space be a function that assigns a one to a head and a zero to a tail and then sums the result. The investigator usually defines the sample space, or experiment, to suit a particular objective. As a matter of conveni-ence, the space is usually selected such that the relative frequencies of occurrence of events within the space are definable. Access to such basic sample spaces as S_1 for the two-coin ex-periment allow for the calculation of relative fre-quencies for events in both S_1 and S_2. Without the existence of a space like S_1, determining the true relative frequency of occurrence for an event in S_2 is difficult, if not impossible. This situation is also evident from the hydraulic-conductivity experiment, where only an S_2-like sample space is available to the investigator.

A random variable can also be described as either being <u>discrete</u>, as in the coin-toss experi-ment, or <u>continuous</u>, as represented by the hydraulic-conductivity experiment. A discrete random variable is defined over a sample space whose elements are discrete, although there may be as many as there are whole numbers

(mathematicians refer to this phenomenon as being <u>countably infinite</u>). A continuous random variable is defined over a continuous sample space whose elements are infinite in number (therefore these elements are <u>uncountably infinite</u>).

2.2 Frequencies and Distributions

2.2.1 Discrete Random Variables

Although frequencies of occurrence are usually associated with events in a sample space, they are also associated with values of random variables, since random variables are functions of the elements in a sample space. That is, particular values of a random variable correspond to particular events in the sample space and, therefore, have frequencies of occurrence. Even though we will speak of the relative frequency of occurrence for particular values of a random variable, we are, in reality, speaking of a corresponding event in the sample space. In fact, we frequently use a range of values of a random variable to define an event in a sample space, thus avoiding the task of describing which elements of the sample space compose the event.

Frequencies of occurrence for events in many discrete sample spaces can be deduced from the following axiomatic premise: If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the relative frequency of occurrence of A is n/N. As a simple example of employment of this premise, consider an experiment consisting of a toss of a die. The sample space consists of the integers 1 through 6 and, for any realization of the experiment, each element of the sample space has equal likelihood of occurrence. By considering each element of the sample space to be an event, one can calculate the frequency of occurrence, $f(x_i)$, with which a random variable takes on the value x_i. For this experiment, only the integer values 1 through 6 of x_i have frequencies of occurrence other than 0; $f(x_i)$ can be graphically represented as shown in figure 2.2-1. In this case, $f(x_i)$ is referred to as the <u>discrete density function</u> of the discrete random variable
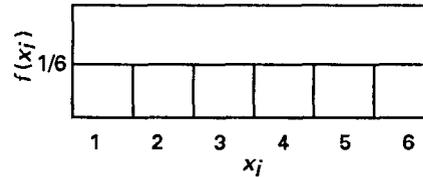
Figure 2.2-1

consisting of the outcome of a toss of a single die.

When two dice are cast, the experiment can be defined either as the sum that results from the toss or simply as all possible arrangements that could appear on the dice. If the sum is chosen, then the sample space consists of the integers 2-12, which would also be the range of values that the random variable could take on. The elements of this space, however, are not equally likely to occur. The sample space consisting of all arrangements of the numbers appearing on the two dice, presented graphically in table 2.2-1, has elements which are equally likely to occur.

Table 2.2-1

Second die	First die					
	1	2	3	4	5	6
1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

The relative frequency of occurrence of an event corresponding to any subset of elements in this space can be calculated by using the premise concerning equally likely outcomes.

A random variable, consisting of the sum that results from any outcome of the two dice experiment, takes on the integer values 2-12 over the sample space represented by table 2.2-1. The discrete density function for this random variable can now be derived from the basic premise concerning outcomes that are equally likely, since each value for this discrete random variable corresponds to a particular event consisting of a particular subset of elements in the

sample space indicated by table 2.2–1. Thus, the value of $x_i=3$ corresponds to the event containing the elements (2,1) and (1,2) and has a relative frequence of occurrence of 2/36. Letting x_i represent the integer values that this random variable can obtain, its density function, $f(x_i)$, can be represented as shown in figure 2.2–2.

Note that had the first definition of the experiment been used, then every element of the sample space consisting of the integers 2–12 would have frequencies of occurrence, when considering each element as an event, equivalent to those shown in figure 2.2–2.

Frequencies of occurrence, or deduced frequencies of occurrence as indicated in figures 2.2–1 and 2.2–2, are indications of the future. We can make probability statements concerning the possibility of a random variable taking on future values from such knowledge. In a craps (two-dice) game, we know that the probability of rolling a natural, an outcome of 7 or 11 on the first cast, is 2/9 simply because these values of the random variable for the two-dice experiment correspond to elements in the sample space which occur with a relative frequency of 2/9. Formally, the statement that this discrete random variable X take on the values of 7 or 11 with a probability of 2/9 is written

$$P(X=7 \text{ or } X=11)=2/9.$$

The probability that this random variable takes on any integer value between 2 and 12 is obtainable directly from its frequency density, figure 2.2–2.

A probability statement that is frequently encountered concerns the probability that a random variable is less than or equal to a specific value. For the random variable corresponding to the sum of outcomes of the cast of two dice, we may ask, what is the probability that the random variable X is less than or equal to 5? The probability of this event is equal to the probability that X take on any integer value 2 through 5:

$$P(X \le 5)=P(X=2 \text{ or } X=3 \text{ or } X=4 \text{ or } X=5).$$

This probability is the sum of the probabilities of the individual events that X take on the integer values 2 through 5:

$$P(X \le 5)=1/36+2/36+3/36+4/36=5/18.$$

(If the student is not convinced of this relationship, he or she should examine the elements of the sample space represented by table 2.2–1 to ascertain that it holds.) Note that $P(X \le 12)$ is unity; that is, an event which occurs with a probability of one will, undoubtedly, take place. A probability of zero indicates, on the other hand, that the event of concern cannot possibly occur.

The probability statement $P(X \le a)$, where a is any real number, is given a special definition for both discrete and continuous random variables. That is, $F(a)=P(X \le a)$ is known as the
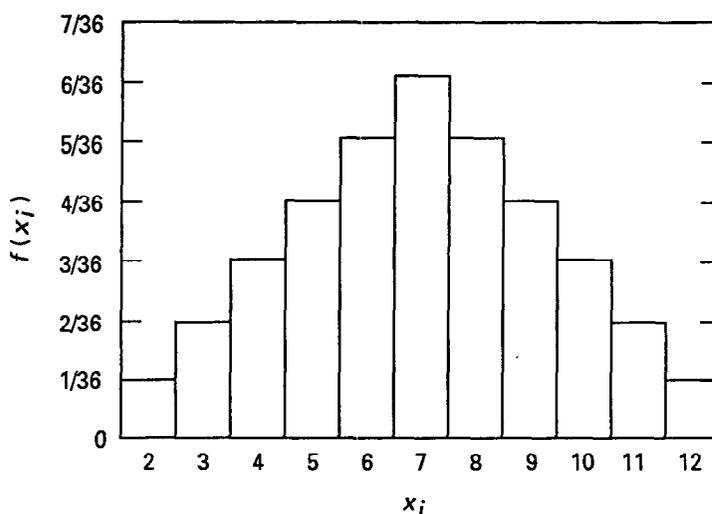
Figure 2.2–2

cumulative distribution function of the random variable X. For the case of the sum of outcomes for two dice, $F(a)$ appears as illustrated in figure 2.2-3. Because a random variable represents a functional mapping from the sample space to the real number space, we can be assured that the probability of the event $X \leq a$ exists and is equal to the sum of the probabilities of all events corresponding to values of the random variable which are less than or equal to a. In general, for discrete random variables, the cumulative distribution function can be evaluated by summing the appropriate relative frequencies of occurrence:

$$F(a) = \sum_{x_i \leq a} f(x_i). \qquad (2.2-1)$$

The cumulative distribution function for all random variables, discrete or continuous, has the following properties:

1. $F(a)$ is a nondecreasing function of a,

2. $\lim_{a \to \infty} F(a) = 1$,

3. $\lim_{a \to -\infty} F(a) = 0$.

These properties will be demonstrated in detail for continuous random variables in a later section. For a discrete random variable, these properties reflect the fact that, by definition, the discrete density function can never have a negative frequency of occurrence and that the sum of frequencies must equal one.

In the next section, an estimator for the density function of continuous random variables is developed, which will eventually allow us to explore the nature of density and cumulative distribution functions of continuous random variables.

Problem 2.2-1

An urn contains one red, one white, and two blue balls, all of equal dimensions. A ball is
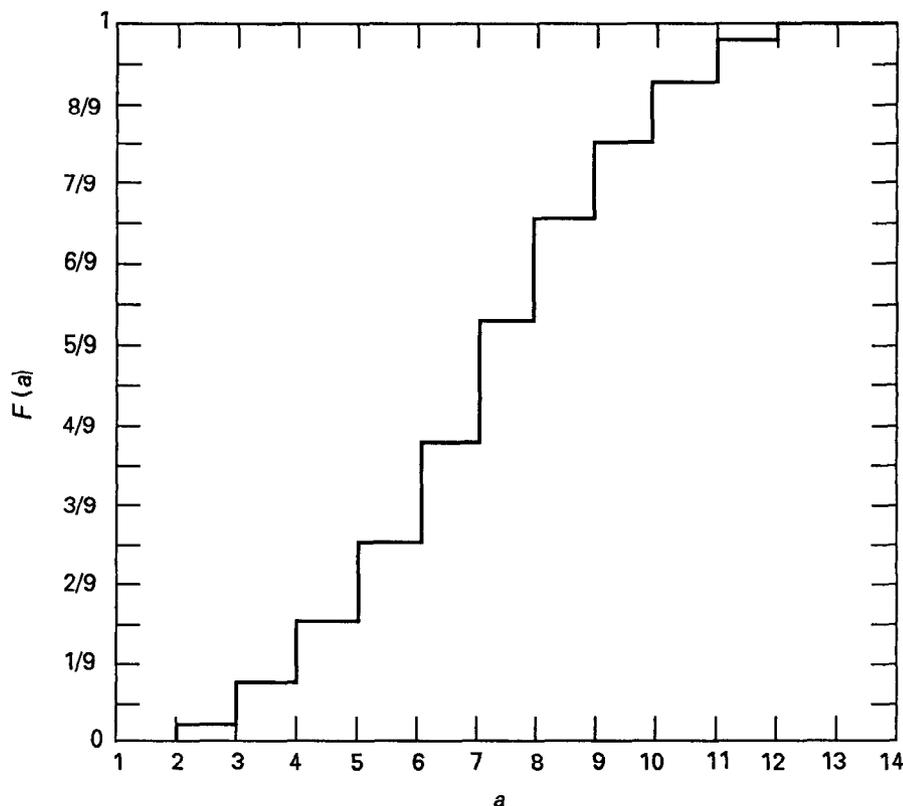
Figure 2.2-3

drawn from the urn, replaced, and then another draw is made.

    a. What possible arrangements (red, white, and (or) blue) of the two balls, considering order of selection, could occur (see, for example, table 2.2–1 for two dice)?

    b. What is the frequency of occurrence of any of the above events? (Hint: let the balls be represented by the symbols R, W, $B1$, and $B2$.)

    c. A value of one is assigned to a blue ball, two to a red ball, and three to a white. A random variable consists of the sum of any outcome consisting of two draws with replacement. Develop a discrete density function for this random variable.

    d. What is the probability that this random variable takes on a value of 4? What arrangements of balls correspond to this value of the random variable?

2.2.2 Histograms

In many cases, we do not have access to all values of random variables in a sample space (in particular, for many continuous random variables). We sample the population consisting of all possible values of the random variable and hope to draw inferences from this sample. The inferences we draw are usually in the form of statistics, which we refer to as sample statistics. We like to think that sample statistics estimate values of population parameters, which are constants reflecting the true frequency distribution of the random variable. This is frequently the case if the observations composing samples are made randomly and without bias. Samples composed of such observations are referred to as random samples and are expected to be representative of the population.

Estimates of density functions for random variables are frequently made from random samples. Although certain experiments, such as a coin toss, allow for the deduction of frequencies of occurrence of events, other experiments defy a theoretical calculation, forcing us to estimate from a random sample. These estimates, known as histograms, are generally constructed by repeating the experiment a large number of times (thus, sampling the population

of all possible outcomes), dividing the range of these outcomes into class intervals, and calculating the relative number of points that fall in each interval. We might imagine, for example, that we could watch a craps game and note the outcome of each roll of dice. After a thousand rolls, we would calculate the relative percentage of each integer, 2–12, which occurred. If these sample frequencies of occurrence were not close to that shown previously for the theoretical result, we would suspect that the dice had been tampered with.

As an example of a histogram constructed from observed values of a continuous random variable, consider the transmissivity data shown in table 2.2–2. Figure 2.2–4 represents a histogram constructed directly from these data, which constitute a random sample from the population of transmissivities as determined from specific capacities of wells in carbonate rocks of central Pennsylvania. A second histogram, figure 2.2–5, was constructed from a logarithmic transformation of these data as shown in table 2.2–3. The first histogram was constructed by using a class interval of 50,000 gal/d/ft, and the second is based upon an interval of one-half a $\log_{10}$ cycle. The first histogram is not very illustrative because most of the wells have transmissivities less than 50,000 gal/d/ft (the underlying population frequency is probably heavily skewed to the right). By logarithmically transforming of the random variable, we scale the abscissa so as to remove the skewness in the histogram, causing it to be more bell shaped. This type of transformation is frequently used on random variables that have a zero lower bound, causing the transformed variable to have tails that tend to infinity in both directions. The transformation also tends to remove any right skewness in the frequency distribution of these random variables. With regard to the transformed variate, the histogram in figure 2.2–5 suggests a bell-shaped population frequency distribution. More data and smaller class intervals, as suggested in the following paragraphs, should cause the histogram shown in figure 2.2–5 to approach its population shape, which we may suspect to be a normal distribution; the untransformed random variable would then result from a log-normal distribution.

Table 2.2-2

[From Siddiqui (1969, p. 433–436)]

Transmissivity gal/d/ft	$Log_{10}T$	Transmissivity gal/d/ft	$Log_{10}T$
15.0	1.176	2,370.0	3.375
18.0	1.255	2,440.0	3.387
21.0	1.322	2,540.0	3.405
29.0	1.462	2,800.0	3.447
32.0	1.505	2,820.0	3.450
35.0	1.544	3,380.0	3.529
50.0	1.699	4,410.0	3.644
52.0	1.716	4,520.0	3.655
56.0	1.748	5,500.0	3.740
62.0	1.792	5,650.0	3.752
84.0	1.924	6,030.0	3.780
92.0	1.964	6,240.0	3.795
106.0	2.025	6,340.0	3.802
118.0	2.072	7,290.0	3.863
142.0	2.152	8,130.0	3.910
160.0	2.204	11,000.0	4.041
175.0	2.243	13,100.0	4.117
184.0	2.265	13,700.0	4.137
202.0	2.305	14,500.0	4.161
264.0	2.422	17,200.0	4.236
354.0	2.549	17,700.0	4.248
370.0	2.568	19,700.0	4.294
374.0	2.573	23,100.0	4.364
455.0	2.658	24,200.0	4.384
463.0	2.666	26,400.0	4.422
515.0	2.712	33,400.0	4.524
528.0	2.723	34,700.0	4.540
615.0	2.789	42,400.0	4.627
705.0	2.848	46,300.0	4.666
753.0	2.877	52,000.0	4.716
800.0	2.903	66,500.0	4.823
984.0	2.993	68,400.0	4.835
1,150.0	3.059	132,000.0	5.121
1,290.0	3.111	152,000.0	5.182
1,500.0	3.176	423,000.0	5.626
1,580.0	3.199	423,000.0	5.626
1,670.0	3.223	528,000.0	5.723
1,850.0	3.267	528,000.0	5.723
2,310.0	3.364	528,000.0	5.723

Table 2.2-3

Class interval[1]	Number of occurrences	Relative frequency	Cumulative frequency
1.0–1.5	4	0.051	0.051
1.5–2.0	8	.103	.154
2.0–2.5	8	.103	.257
2.5–3.0	12	.154	.411
3.0–3.5	12	.154	.565
3.5–4.0	10	.128	.693
4.0–4.5	10	.128	.821
4.5–5.0	7	.090	.911
5.0–5.5	2	.026	.937
5.5–6.0	5	.064	1.001
Total	78		

[1]Based on $log_{10}T$, table 2.2-2
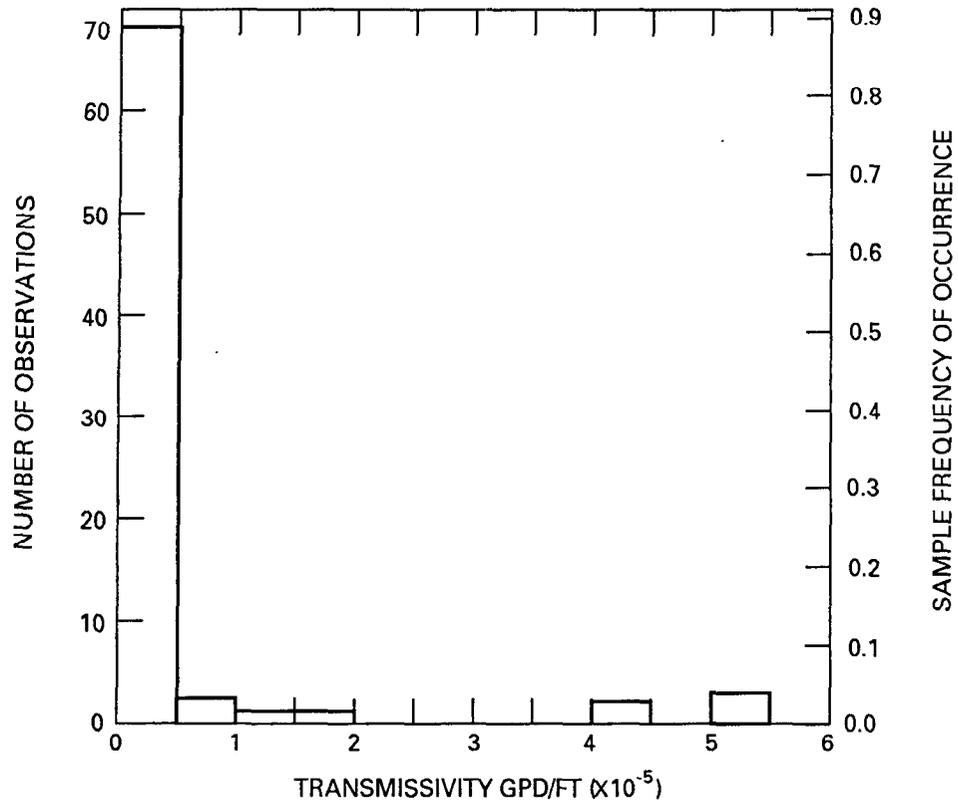
Figure 2.2-4
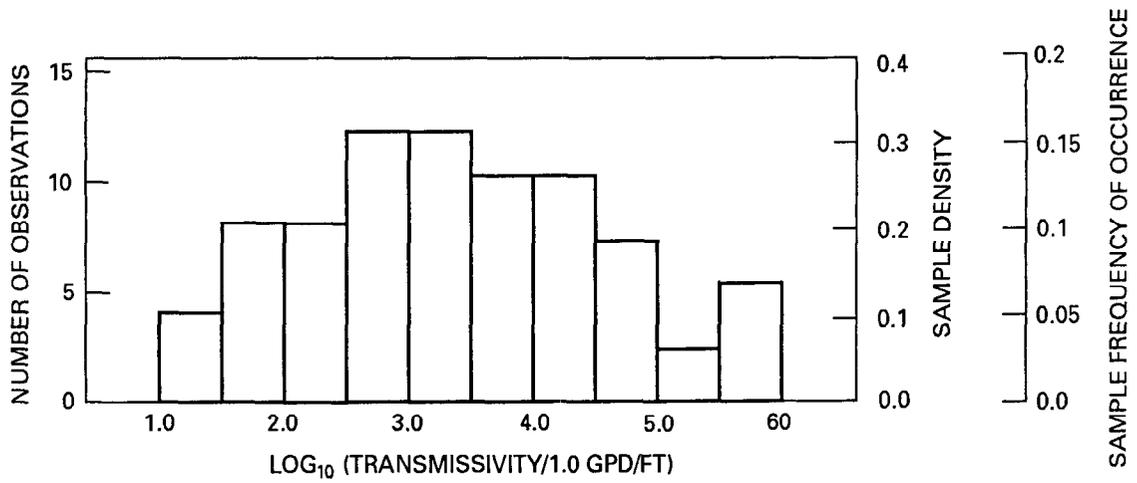
Figure 2.2-5

We are now in position to estimate the probability of occurrence of an event associated with the log-transformed random variable. Let $X = \log_{10} T$ represent the transformed variate plotted in figure 2.2-5, and assume that the histogram for X is representative of the population frequency. The estimated probability that X is less than or equal to 5.0 but greater than 4.5, $P(4.5 < X \le 5.0)$, then, is the sample frequency of occurrence of this event (equal to 0.09). The probability that X is less than or equal to 5.0 can be estimated by summing the frequencies of occurrence of all events smaller than 5.0; thus $P(X \le 5.0) \approx 0.911$. Thus, the chances are about 91 in 100 that the transmissivity of the carbonate rocks in central Pennsylvania, as determined by any random well, will be less than or equal to 1×10^5 gal/d/ft. The reader should realize that these results are only approximate, as the histogram is an approximation of the true population frequency distribution.

An estimate of the cumulative distribution function can also be constructed from a random sample. Let $F_n(a)$ represent this estimate, known as the <u>sample distribution function</u>; an appropriate estimator for $F_n(a)$ is the sum of all estimated relative frequencies for values of the random variable X less than a:

$$P(X \le a) \approx F_n(a) = \sum_{i \le a/\Delta x} f_i^* \qquad (2.2\text{-}2)$$

where

$f_i^* = n_i/n = $ sample frequency of occurrence of an event represented by the ith class interval,

$\Delta x = $ size of class interval,

$n_i = $ number of outcomes having values in interval i, and

$n = $ size of random sample.

An application of this procedure for the logarithmic transformation of transmissivity is shown in figure 2.2-6.

2.2.3 Continuous Random Variables

The definition of frequency f_i^* used in equation 2.2-2 suffers from the deficit that it is dependent upon the size of the class interval; that is, if Δx decreases in size while n remains constant, then f_i^* must also decrease, as we are also effectively decreasing the value of n_i within this interval. Indeed, even if n were allowed to become large as Δx decreases, thus causing n_i for any arbitrary interval to be large, f_i^* could still be made arbitrarily small by decreasing the interval size sufficiently. However, this phenomenon would prevent us from defining a frequency for a single point in a continuous
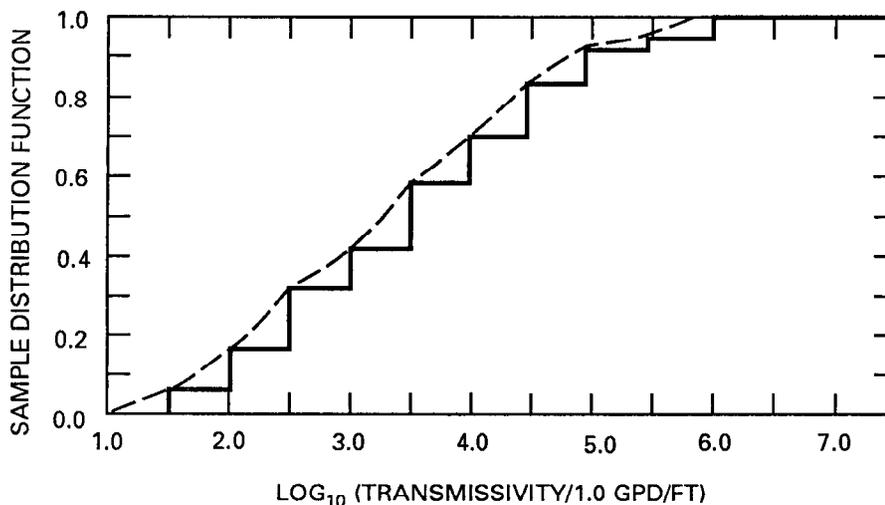
Figure 2.2-6

random variable, unless we are content to associate it with some arbitrary class interval. To overcome this problem, probabilists have defined a different measure of frequency for continuous random variables that consists of the frequency of occurrence f_i^* scaled by its class interval:

$$f_i = f_i^*/\Delta x \qquad (2.2\text{-}3)$$

This normalized frequency, referred to as the sample density, should be relatively stable for reasonable choices of Δx and n, and in the limiting case of n approaching infinity and Δx approaching zero, f_i should be constant. An additional ordinate has been added to figure 2.2-5 to show the sample-density distribution of the $\log_{10} T$ data.

The sample distribution function of equation 2.2-2 can now be redefined in terms of equation 2.2-3 as follows:

$$F_n(a) = \sum_{i \le a/\Delta x} f_i \Delta x \qquad (2.2\text{-}4)$$

This definition lends itself to an exploration of the population equivalents of $F_n(a)$ and f_i. If the random sample is of sufficient size to sample every member of the sample space and Δx is taken infinitely small, then the population equivalents of $F_n(a)$ and f_i should be approached. By letting n become large and Δx small, we see that

$$\sum_{i \le a/\Delta x} f_i \Delta x \approx \int_{-\infty}^{a} f(x)dx \qquad (2.2\text{-}5)$$

where $f(x)$, the population equivalent of f_i, is known as the probability density function. Because $f(x)$ is the population equivalent of f_i, then the integral representation in equation 2.2-5 of summing these scaled frequencies must be the population equivalent of $F_n(a)$, which of course is the same cumulative distribution function defined earlier in section 2.2.1:

$$F(a) = \int_{-\infty}^{a} f(x)dx = P(X \le a) \quad . \qquad (2.2\text{-}6)$$

However, because a random sample, whether it be finite or infinite, is countable, equation 2.2-5 must be given a special interpretation. Note that, because $f(x)$ is the continuous analog of f_i, it is always a non-negative quantity.

A stronger statement than equation 2.2-5 can be made concerning the equivalence of $F(a)$ and $F_n(a)$ for large sample sizes by noting that $F_n(a)$, prior to sampling, is a random variable. That is, if we were to collect different samples of the same size n from the same population, we would not expect that $F_n(a)$, computed from each random sampling, would have the same value. We would only hope that, as n becomes large, these different values would approach some constant. Indeed, probabilists have shown that, with a probability of one, $F_n(a)$ becomes the constant $F(a)$ as n goes to infinity. This result is particularly remarkable if we first consider that $F_n(a)$ can only take on a countable number of values k/n, $0 \le k \le n$, where k is an integer (see equation 2.2-2). Thus, although the values of $F(a)$ are uncountably infinite (continuous), $F_n(a)$ can only be, in the case that the random sample is infinitely large, at most, countably infinite. We will use this result loosely by allowing equation 2.2-5 to take on the indicated limits,

$$\lim_{\substack{\Delta x \to 0 \\ n \to \infty}} F_n(a) = F(a) \qquad (2.2\text{-}7)$$

and noting that this result only can occur with a probability of one.

Both $f(x)$ and $F(x)$ are continuous functions of values of the random variable X. For the previously illustrated case of $X = \log_{10} T$, the density function might appear as in figure 2.2-7. Figure 2.2-7 represents the population equivalent of figure 2.2-5, as if all possible outcomes of the random variable were available to us. Similarly, the cumulative frequency distribution, the population equivalent of figure 2.2-6, for this random variable might appear as in figure 2.2-8.

Because of equation 2.2-6, the density function $f(x)$ can be defined in terms of the cumulative distribution function $F(x)$ by differentiation:
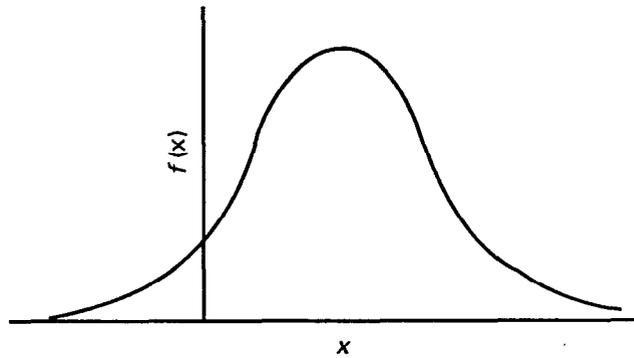
Figure 2.2-7
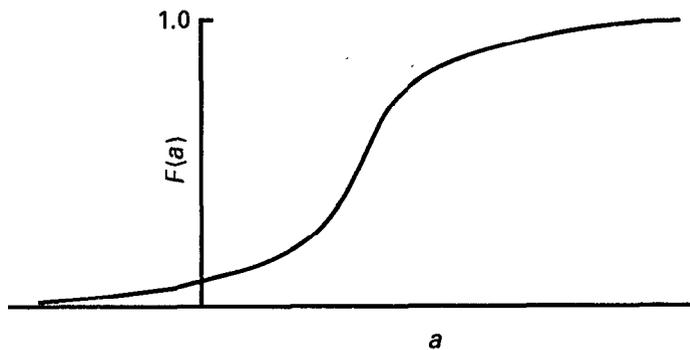
Figure 2.2-8

$$f(a) = \frac{dF(a)}{da} = \frac{d}{da} \int_{-\infty}^{a} f(x)dx. \qquad (2.2\text{-}8)$$

This result follows directly from the fundamental theorem of integral calculus, and is applicable only to density functions of continuous random variables. Equation 2.2-8 is one of three concepts which defines density functions of continuous random variables. The other two state that $f(x)$ must be greater than or equal to zero for any possible value of the random variable and, as will be demonstrated in the next section, that the total mass under the frequency curve must be unity. All density functions of continuous random variables have these concepts in common.

Problem 2.2-2

a. Construct histograms for the following specific-conductance data using class intervals of 100 and 200 μmho/cm, such that the abscissa and ordinate of both histograms are scaled equally. What is the effect of changing the class interval?

b. Construct a cumulative frequency distribution from your 100 μmho/cm class-interval results. Let X represent the specific-conductance random variable; what is

$$P(X \leq 600)?$$
$$P(X > 400)?$$
$$P(400 < X \leq 600)?$$
$$P(X \leq 1300)?$$

Ordered specific-conductance data

[Data in μmho/cm for wells in carbonate rocks of Maryland. From Nutter, 1973, p. 63–68]

63	423	501	582	685	836
76	433	504	596	697	839
168	439	509	598	700	876
278	440	512	600	704	882
301	440	518	604	710	895
304	440	518	617	721	897
310	444	527	620	723	904
315	452	529	627	724	906
319	452	533	629	726	915
323	452	537	632	728	948
332	456	538	636	740	968
347	462	542	641	750	969
357	469	552	647	750	982
359	471	562	659	764	997
363	473	564	659	765	1,030
389	477	564	661	779	1,080
407	487	565	664	783	1,106
408	490	566	665	789	1,120
411	492	570	670	808	1,170
413	493	575	673	808	1,230
417	493	578	675	815	
418	499	582	677	820	

2.2.4 Properties of Cumulative Distribution Functions

In the previous section, the cumulative distribution function $F(a)$, defined by the probability statement $P(X \leq a)$, was noted to have the integral form of equation 2.2-6 for continuous random variables. We state all manner of probability statements in terms of the cumulative distribution function, as this is a standard form. For this purpose, properties of cumulative distribution functions, with applications to other probability statements, are developed in this section.

The probability that a random variable X takes on a value in the interval $(a, b]$ can be expressed in terms of cumulative distribution functions as

$$P(a < X \leq b) = \int_a^b f(x)dx = F(b) - F(a) \quad . \quad (2.2-9)$$

This statement is a direct result of integral calculus, whereby integration is used to sum all the frequencies of occurrences of values of the random variable between a and b. From equation 2.2-9 one sees that the cumulative distribution function is a nondecreasing function of x, because

$$0 \leq P(a < X \leq b) \leq 1.$$

That the total mass under the sample density curve f_i is unity is evident from equation 2.2-4; that is,

$$\lim_{a \to \infty} F_n(a) = \lim_{a \to \infty} \sum_{i \leq a/\Delta x} f_i \Delta x = 1 \quad . \quad (2.2-10)$$

Because the probability density function $f(x)$ of a continuous random variable X is a limiting form of the sample density f_i, the mass under its curve is also unity:

$$\lim_{a \to \infty} F(a) = \lim_{a \to \infty} \int_{-\infty}^a f(x)dx = 1 \quad . \quad (2.2-11)$$

Equation 2.2-11 is a property of all cumulative distribution functions. Similarly,

$$\lim_{a \to -\infty} F(a) = \lim_{a \to -\infty} \int_{-\infty}^a f(x)dx = 0 \quad , \quad (2.2-12)$$

which follows from integral calculus, is also a property of cumulative distribution functions.

Equation 2.2-11 allows one to express $P(X>a)$ as

$$P(X>a)= \int_a^\infty f(x)dx=1-\int_{-\infty}^a f(x)dx=1-F(a), \quad (2.2\text{-}13)$$

which is also a result of Riemannian integration. An alternate statement of equation 2.2-13 is that $P(X>a)=1-P(X\leq a)$.

By considering equation 2.2-9 in a limit form, we can also find the probability that $X=a$:

$$P(X=a)=\lim_{\Delta x\to 0} P(a<X\leq a+\Delta x) \quad ,$$

$$=\lim_{\Delta x\to 0} [F(a+\Delta x)-F(a)]=0. \quad (2.2\text{-}14)$$

This result is unique to continuous random variables, in contradistinction to discrete random variables. From equation 2.2-14 one sees that $P(X\leq a)$ is equivalent to $P(X<a)$ for continuous random variables, as the endpoint, a, of the semi-infinite interval does not contribute mass to the probability statement.

Equations 2.2-9, 2.2-11, and 2.2-13 can be demonstrated for discrete random variables by using the summation form of the cumulative distribution function (equation 2.2-1). In contradistinction to continuous random variables, the endpoint in $P(X\leq a)$ for a discrete random variable can contribute significant mass to the statement.

A number of frequency densities that result from randomness in nature, or probabilistic models of random events, have been investigated and published. Cumulative distributions of these densities are frequently tabulated and are found in many reference books on probability and statistics. Equations 2.2-9 and 2.2-13 are especially useful in evaluating probability statements of tabulated random variables.

2.2.5 An Example: The Normal Distribution

Let the random variable Y represent the amount of titrant used in a titration experiment to neutralize measured amounts of the unknown x. A scatter diagram of titrant versus unknown might appear as in figure 2.2-9. The solid line represents the true stoichiometric balance between titrant and unknown. The dots, representing repetitions of the experiment, deviate from this line by an amount ϵ, which represents a value of the measurement error $\mathcal{E}$. These errors represent a continuous random variable that could theoretically vary from $-\infty$ to $+\infty$ (the graphed points only represent a random sample from the population). If the experimental apparatus is functioning properly, however, we would expect these dots to be concentrated in the general vicinity of the solid line.

A distribution that is frequently used to model errors that are symmetrically distributed about some common point is the normal
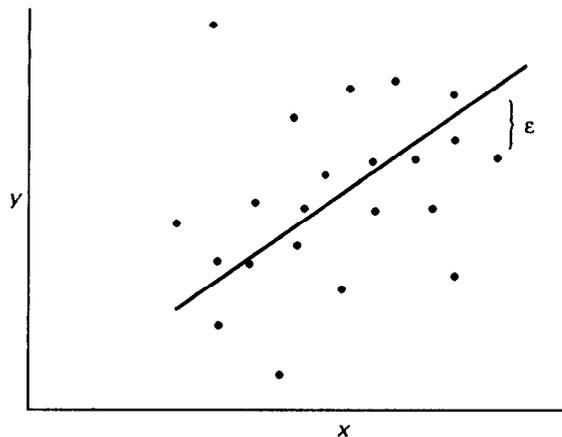
Figure 2.2–9

distribution. The density of the normal distribution is a bell-shaped curve, symmetric about its mean μ_ε, and with most of the mass concentrated within one standard deviation σ_ε of the mean (see figure 2.2-10). In the case of the titration experiment, we would hope that the most frequently found value of the error would be near-zero and expect that μ_ε would equal zero. The standard deviation σ_ε is a measure of the dispersion, or spread, of the errors about the mean and is equal to the distance from the mean to an inflection point on the curve $f(\epsilon)$. The mean and standard deviation will be formally defined in a later section.

A normal random variable is frequently standardized with its mean and standard deviation by the following transformation:

$$Z = (\varepsilon - \mu_\varepsilon)/\sigma_\varepsilon \quad . \qquad (2.2\text{-}15)$$

The cumulative distribution for this standard normal random variable is tabulated (table 2.10-1) for use by the investigator, since its probability density function, $f_Z(z)$, is parameter free:

$$f_Z(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \quad . \qquad (2.2\text{-}16)$$

Given the density function for the standard normal random variable, it is natural to inquire about the form of density, $f_\varepsilon(\epsilon)$, of the unnormalized random variable ε. Consider the cumulative frequency distribution for Z. By making the change of variables $z = (s - \mu_\varepsilon)/\sigma_\varepsilon$,

$$F_Z(a) = \int_{-\infty}^{a} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz$$

$$= \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \int_{-\infty}^{\epsilon} \exp\left[-\left(\frac{s - \mu_\varepsilon}{\sigma_\varepsilon}\right)^2/2\right] ds \quad (2.2\text{-}17)$$

results where $\epsilon = a\sigma_\varepsilon + \mu_\varepsilon$ is a value of the unnormalized random variable. Since differentiation is the inverse operator of integration, equation 2.2-17 is differentiated with respect to ϵ to find $f_\varepsilon(\epsilon)$ (see also equation 2.2-8):

$$f_\varepsilon(\epsilon) = \frac{d}{d\epsilon} F_Z(a(\epsilon))$$

$$= \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left[-\left(\frac{\epsilon - \mu_\varepsilon}{\sigma_\varepsilon}\right)^2/2\right] . \qquad (2.2\text{-}18)$$

$\mu_\varepsilon = 0$
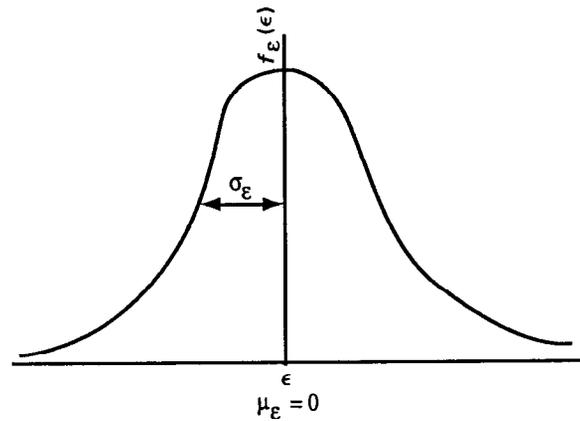
Figure 2.2-10

Note that equation 2.2-18 is not parameter free, as this density is a function of the parameters μ_ε and σ_ε.

2.3 Expectation and the Continuous Random Variable

The discussion in this section is largely presented with continuous random variables in mind. All the results, however, are applicable to discrete random variables; whenever a quantity is defined by an integration over a probability density function for the continuous case, this same quantity can almost invariably be defined by a summation over the discrete density function for the discrete case. The reader should demonstrate the veracity of this statement.

2.3.1 The Mean

The mean is a measure of central tendency of a population. As an estimator of this central tendency, consider a finite random sample consisting of n values x_i of the random variable X. If the sample frequency of occurrence f_i^* is estimated from this random sample, then a logical estimator of the central tendency is to sum the product of the central value $\bar{x}_i$ of each class interval and the frequency of occurrence for that interval:

$$\bar{x} = \sum_{i \leq x_m/\Delta x} f_i^* \bar{x}_i = \sum_{i \leq x_m/\Delta x} f_i \bar{x}_i \Delta x \qquad (2.3\text{-}1)$$

where x_m is the upper limit of the largest class interval necessary to construct f_i^*. The frequencies of occurrence f_i^* in equation 2.3-1 can be looked upon as weights that sum to one, and the quantities $\bar{x}_i$ as equally spaced values of the random variable. The values of the random variable that occur more frequently, as indicated by the random sample, receive larger weights through equation 2.3-1 and will have a greater influence on $\bar{x}$.

Equation 2.3-1 should be recognized by the reader as also being the definition for the center of mass of physical weights distributed along a line. That is, if mf_i^* represents the mass of a weight located at $\bar{x}_i$, where m is the total mass of all the weights, then equation 2.3-1 would give us the center of mass of the line with respect to the origin. In the case of a histogram, the role of the weights is played by the sample frequency of occurrence for an interval, which gives us the approximate relative likelihood that any future value of the random variable will occur in that interval. For calculation purposes, this distributed weight over any interval i is replaced by a point weight having the same mass as the distributed weight, but located at the center $\bar{x}_i$ of the interval. The sum of the products of the relative masses of these point weights, f_i^*, with their relative distances from the origin, $\bar{x}_i$, gives us the center of mass, which is also a measure of the central tendency. Of course, if the sample size n were to become very large, then Δx could be made very small, refining equation 2.3-1 as an estimator of the central tendency of a random variable.

Reasoning similar to that leading to $\bar{x}$ as an estimator of the population mean can be applied directly to defining this parameter. First, given that the density function $f(x)$ is known, then the approximate frequency of occurrence of an event corresponding to an interval of size Δx that has as its central value $\bar{x}_i$ is $f(\bar{x}_i)\Delta x$. Thus, assuming that these relative frequencies are centered at each $\bar{x}_i$, an approximate measure of the population central tendency, μ_X, is

$$\mu_X \approx \sum_{(\text{all } i)} \bar{x}_i f(\bar{x}_i) \Delta x \qquad (2.3\text{-}2)$$

where the values $\bar{x}_i$ are equally spaced by Δx from each other. Of course, by letting Δx

become smaller, a more accurate measure of μ_X is developed, until μ_X, also known as the expected value, $E[X]$, of the random variable X, is defined by the following integral expression:

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f(x)dx \quad . \qquad (2.3\text{-}3)$$

This equation is the standard form for the expected value of a univariate random variable.

Equation 2.3-3 can also be developed directly from equation 2.3-1 by letting $n \rightarrow \infty$ and $\Delta x \rightarrow 0$:

$$\mu_X = \lim_{\substack{n \to \infty \\ \Delta x \to 0}} \bar{x} = \lim_{\substack{n \to \infty \\ \Delta x \to 0}} \sum_{i \leq x_m/\Delta x} \bar{x}_i f_i \Delta x = \int_{-\infty}^{\infty} x f(x)dx. \ (2.3\text{-}4)$$

That is, as Δx becomes smaller and as the number of observations becomes very large, $\bar{x}_i$ becomes a unique continuous value of X, f_i becomes the continuous function $f(x)$, and the summation can be replaced by an integration. As in the case of equation 2.2-7, we can only say that the limit indicated in equation 2.3-4 is reached with a very high probability as n becomes large; however, this probability should be unity as n becomes infinite.

Note that μ_X is a population parameter that is characteristic of the random variable X, while $\bar{x}$, being derived from values of a finite random sample from the population of X, is only an estimate for μ_X. Estimators such as $\bar{x}$ will be developed in greater detail in a later section.

Problem 2.3-1

a. Find $\bar{x}$ from 100 μmho/cm histogram of problem 2.2-2.
b. Find μ_X for the random variable of problem 2.2-1.

2.3.2 Generalization and Application of the Expectation Operator

The operation of finding an expected value can be generalized by considering a function $g(X)$ of continuous random variable X. If we

wish to find the average effect of the function $g(x)$ over the outcomes of a random sampling, we again resort to the approximation

$$\overline{g(x)} = \sum_{i \le x_m/\Delta x} f_i^* g(\bar{x}_i) \ . \qquad (2.3\text{-}5)$$

That is, we weight $g(x)$, where $g(x)$ is evaluated at the center of every class interval, by the frequency of occurrence of that interval and sum all the weighted values of $g(\bar{x}_i)$. To obtain the population equivalent of $\overline{g(x)}$, n is taken to be large, while Δx is taken to be small; this equivalent is denoted by the expectation symbol $E[g(X)]$:

$$E[g(X)] = \lim_{\substack{n \to \infty \\ \Delta x \to 0}} \overline{g(x)} = \int_{-\infty}^{\infty} g(x) f(x) dx \ . \qquad (2.3\text{-}6)$$

Equation 2.3-6 represents the general form of the expectation operator for a univariate distribution when the random variable is continuous. A similar form exists for discrete random variables, in which the integration has been replaced by summation.

A trivial but useful property of the expectation operator is that the expected value of any constant c is that constant; for the continuous case, this is easily demonstrated as

$$E[c] = \int_{-\infty}^{\infty} c f(x) dx = c \int_{-\infty}^{\infty} f(x) dx = c \qquad (2.3\text{-}7)$$

where equation 2.2-11 has been invoked. A more important property of E is that it is a linear operator; that is,

$$E[a g_1(X) + b g_2(X)] = a E[g_1(X)] + b E[g_2(X)]. \qquad (2.3\text{-}8)$$

This property results because in the continuous case, integration itself is a linear operator:

$$\int_{-\infty}^{\infty} (a g_1(x) + b g_2(x)) f(x) dx = a \int_{-\infty}^{\infty} g_1(x) f(x) dx$$

$$+ b \int_{-\infty}^{\infty} g_2(x) f(x) dx \qquad (2.3\text{-}9)$$

As a practical example of finding the expected value of a random variable, consider the problem of finding the mean of X where X is a normal random variable with mean μ_X and standard deviation σ_X:

$$E[X] = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{\infty} x \exp\left[-\left(\frac{x-\mu_X}{\sigma_X}\right)^2 / 2\right] dx. \qquad (2.3\text{-}10)$$

By a change of variable $z = (x - \mu_X)/\sigma_X$, we see that equation 2.3-10 becomes

$$E[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu_X + \sigma_X z) e^{-z^2/2} dz = \mu_X, \qquad (2.3\text{-}11)$$

because $z \cdot \exp(-z^2/2)$ is an odd function of z in the interval $(-\infty, \infty)$, and equation 2.3-7 holds (μ_X, being a population parameter, is constant). Equation 2.3-11 is the reason why μ_X is defined to be $E[X]$.

2.3.3 The Variance, Standard Deviation, and Coefficient of Variation

Although the mean μ_X is a measure of the central tendency of a random variable, it gives no information as to how frequently a random variable will be encountered in its vicinity. The variance σ_X^2, defined as the expected value of the function $g(X) = (X - \mu_X)^2$, is a population parameter that quantifies this concept. The variance can also be looked upon as an operator that is defined in terms of another operator (the expectation operator) as follows:

$$\sigma_X^2 = \text{Var}[X] = E[(X - \mu_X)^2] \ ,$$

$$= \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) dx \qquad (2.3\text{-}12)$$

where $\text{Var}[X]$ represents an operator that operates on X. The intuitive sense of σ_X^2 is that it is the sum of the frequency weighted deviations, which have been squared, from the mean. As such, σ_X^2 represents the amount of dispersion of the random variable about the mean: when σ_X^2 is relatively large, then a random variable is less likely to have values in the

immediate vicinity of the mean. The standard deviation σ_X is simply the square root of the variance: $\sigma_X = (\mathrm{Var}[X])^{1/2}$.

By exercising the linear property of the expectation operator, equation 2.3-12 can be expressed in an alternate form:

$$\sigma_X^2 = \mathrm{Var}[X] = E[(X - \mu_X)^2] = E[X^2 - 2X\mu_X + \mu_X^2],$$

$$= E[X^2] - \mu_X^2. \tag{2.3-13}$$

The variance operator, like the expectation operator, can be generalized to operate on any function $g(X)$:

$$\mathrm{Var}[g(X)] = E[g^2(X)] - (E[g(X)])^2. \tag{2.3-14}$$

The variance operator, however, is not a linear operator, as demonstrated with the function $g(X) = a + bX$:

$$\mathrm{Var}[g(X)] = E[(a + bX)^2] - (E[a + bX])^2,$$

$$= b^2 E[X^2] - b^2 \mu_X^2 = b^2 \sigma_X^2, \tag{2.3-15}$$

because $E[a + bX] = a + b\mu_X$. By letting $b = 0$ in the above example, one can demonstrate that the variance of a constant, as expected, is zero.

When the standard deviation is normalized by the mean of the random variable ($\mu_X \neq 0$), it is referred to as the coefficient of variation V_X:

$$V_X = \sigma_X / \mu_X. \tag{2.3-16}$$

Estimators for this population parameter, as well as the variance and standard deviation, are discussed in a later section of this report.

As an example of an application of the variance operator, consider an application on the standard normal random variable $Z = (X - \mu_X)/\sigma_X$:

$$\mathrm{Var}[(X - \mu_X)/\sigma_X] = \mathrm{Var}[X]/\sigma_X^2 = 1 \tag{2.3-17}$$

which results by analogy with equation 2.3-15. Thus, if X is a normal random variable with mean μ_X and variance σ_X^2, then Z is a zero-mean random variable with a variance of unity, which is commonly denoted $N(0,1)$.

Problem 2.3-2

a. For $f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$

i. Plot $f(x)$.
ii. Derive and plot $F(x)$.
iii. Calculate $E[X]$ and $\mathrm{Var}[X]$.

b. An estimator of the variance σ_X^2 of the random variable X can be developed directly from equation 2.3-12. First, $f(x)dx$ is estimated by f_i^* of equation 2.2-2. Then x is replaced by $\bar{x}_i$, the center of each class interval corresponding to f_i^*. Finally, μ_X is estimated by $\bar{x}$ from equation 2.3-1. Then

$$s_X^{*2} = \sum_{i \le x_m / \Delta x} (\bar{x}_i - \bar{x})^2 f_i^*$$

gives an estimate of σ_X^2. Apply this estimator to the log-transmissivity data of table 2.2-3.

2.4 Jointly Distributed Random Variables

The investigator frequently encounters the problem that he or she has to deal with two (or more) random variables in the same probability statement. As an example, in the case of random variables X and Y, where X and Y are possibly correlated, one might desire the probability that X is less than or equal to a, and Y is less than or equal to b. If the investigator should know the form of the joint probability density function $f(x,y)$ for these two random variables, then this probability statement is definable:

$$P(X \le a \text{ and } Y \le b)$$

$$= \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)\,dy\,dx = F(a,b) \tag{2.4-1}$$

where $F(a,b)$ is the equivalent cumulative distribution function. (The statement $P(X \le a$ and $Y \le b)$ is also denoted frequently as $P(X \le a, Y \le b)$; the more explicit form will be used in this discussion.) As in the univariate case, it is required that the mass under the joint probability density function equal unity:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1. \tag{2.4-2}$$

The concept of joint probability density functions applies to any number of multiple random variables; the following discussion, however, is largely restricted to the bivariate case.

As an example of an experiment yielding jointly distributed random variables, consider the results from a simple nonsteady-state pumping test of a confined aquifer: When the Theis equation is used to evaluate data from these tests, information concerning the storativity and transmissivity of the aquifer results. Indeed, we can easily imagine that these quantities are random variables, varying from location to location in response to the local distribution of materials composing the aquifer. More important, however, would be the manner in which they vary with regard to each other: Should the clay content of the aquifer increase at some point, it might be expected that the transmissivity will decrease while the storativity, reflecting the compressibility of the aquifer, would increase. Thus, quite possibly these quantities, with regard to the aquifer in question, could be treated as jointly distributed random variables which are, in some manner, interdependent.

Assume for the moment that we have determined the form of the joint density function of storativity and transmissivity. For argument's sake, let X represent the transmissivity random variable (or its logarithmic transformation) and Y represent the storativity (or a functional transformation thereof) and then denote the joint density as $f(x,y)$. Now assume that we are interested in the probability that X is less than or equal to a, regardless of the value of Y; that is, we wish to evaluate the probability that our measure of the transmissivity will take on a specific range of values, whereas the exact value of storativity is unimportant to us. For our probability statement regarding X to be meaningful, all values of Y which influence the joint density function must be taken into consideration, for different values of Y would surely influence a statement on X alone. To obtain the total contribution of Y to the joint density function, we allow that Y may take on any value in the interval $(-\infty, \infty)$ and write our probability statement as

$$P(X<a \text{ and } -\infty<Y<\infty)=\int_{-\infty}^{a}\int_{-\infty}^{\infty} f(x,y)dydx$$

$$=\int_{-\infty}^{a} f_X(x)dx \qquad (2.4-3)$$

in which the evaluation of the inner integral with respect to y results in a function $f_X(x)$ that meets all requirements to be a probability density function. Thus, in general, univariate density functions can be recovered from joint density functions by integration, and this integration has the effect of summing the total contribution of one random variable in the bivariate joint density onto the axis of the other variate, the second variate giving the relative frequency of occurrence of the event in question. These densities are referred to as <u>marginal probability density functions</u> and, with respect to the bivariate joint density $f(x,y)$, they are defined as

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy , \qquad (2.4-4)$$

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx , \qquad (2.4-5)$$

where $f_X(x)$ is the marginal density for the X random variable and similarly $f_Y(y)$ for the Y random variable. The marginal-density concept is easily extended to multiple random variables when they are jointly distributed.

2.4.1 Expectation of Jointly Distributed Random Variables

The expectation operator for jointly distributed random variables is defined in the same manner as in the univariate case. Thus, if X and Y are jointly distributed, and $g(X, Y)$ is a function of these two random variables, then a general definition of the expectation operator is

$$E[g(X, Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dydx . \qquad (2.4-6)$$

If on the other hand, we desire the expected value of $h(X)$, which is a function of X only, we set $g(x,y)$ equal to $h(x)$ and proceed as in equation 2.4-6). The result,

$$E[h(X)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)f(x,y)dydx$$

$$=\int_{-\infty}^{\infty} h(x)f_X(x)dx \qquad (2.4-7)$$

shows that, in such cases, finding the expected value reduces to finding the marginal density and integrating. By letting $h(X)$ equal X, one realizes that the mean μ_X is equal to the integral of the product of x and the marginal density $f_X(x)$, as might be expected.

Consider the case where $g(X, Y)$ equals the product $(X-\mu_X)(Y-\mu_Y)$. The expected value of this product gives an indication of how X and Y vary together. If the absolute value of the expected value of this product is exceptionally large, then one would expect that X and Y are highly correlated. This expected value of X and Y is referred to as the covariance of X and Y, and is denoted $\text{Cov}[X, Y]$ or σ_{XY}:

$$\text{Cov}[X, Y] = E[(X-\mu_X)(Y-\mu_Y)]$$

$$= E[XY] - \mu_Y \mu_X . \qquad (2.4\text{-}8)$$

Note that the covariance of X with itself is $\text{Cov}[X, X] = \text{Var}[X]$.

Returning to the example of transmissivities and storativities of the previous section, we see that the covariance provides a measure of the degree of interdependence between random variables. That is, because X and Y are both random, we would not expect observations of X and Y to show a perfect relationship; rather, the relationship will be clouded with noise. Because the expected value implies a frequency-weighted average of the function in question, and because the frequency distribution will reflect the amount of relationship between X and Y, summing the product of these frequency weights with $(X-\mu_X)(Y-\mu_Y)$ over the total variate space will give the average relationship between X and Y. It will be demonstrated in the next section that if X, the measure of transmissivity, and Y, the measure of storativity, were independent, then the covariance would theoretically be zero. However, if our intuition is correct, we would not expect this; rather we might expect, should the aquifer have a rather high clay content, that the two variables will be negatively correlated.

If the covariance is normalized with the standard deviations of the two random variables, then it is referred to as the correlation coefficient ρ_{XY}:

$$\rho_{XY} = \text{Cov}[X, Y] / (\sigma_X \sigma_Y) . \qquad (2.4\text{-}9)$$

The correlation coefficient, as a measure of the linear relationship between X and Y, has the property that its absolute value is less than or equal to unity:

$$|\rho_{XY}| \leq 1 . \qquad (2.4\text{-}10)$$

That is, when X and Y are precisely linearly related, then $|\rho_{XY}|$ will equal unity. If there is no relationship between X and Y, as shown in the next section, $\text{Cov}[X, Y]$ and therefore ρ_{XY} will be zero. This property is demonstrated in appendix 2.11.1, but this appendix requires some knowledge of the next section.

2.4.2 Independent Random Variables

Two random variables X and Y are said to be independent if, for all a and b,

$$P(X \leq a \text{ and } Y \leq b)$$

$$= P(X \leq a) P(Y \leq b)$$

$$= \int_{-\infty}^{a} \int_{-\infty}^{b} f_X(x) f_Y(y) dy dx , \qquad (2.4\text{-}11)$$

where $f_X(x)$ and $f_Y(y)$ are the densities of X and Y, respectively. Equation 2.4-11 implies that the joint density function of two independent random variables is the product of their individual densities, that is,

$$f(x, y) = f_X(x) f_Y(y) . \qquad (2.4\text{-}12)$$

Of course, an event corresponding to $X \leq a$ and $Y \leq b$ would be expected to occur with equal or less frequency than an event corresponding to either $X \leq a$ or $Y \leq b$ separately. Only in the case of a complete lack of dependence between these events can we say that $P(X \leq a \text{ and } Y \leq b) = P(X \leq a) P(Y \leq b)$. This is a somewhat intuitive result that has already been used in connection with the two-dice experiment; if X is an outcome of the first die and Y the second, then $P(X = 1 \text{ and } Y = 2) = P(X = 1) P(Y = 2) = 1/36$.

A random sample is, ideally, a collection of independent random variables. That is, prior to their observation, each element of a random sample is a random variable; its value is not known until after the observation process is

completed. These outcomes should not have any interdependence which might affect the sample density. This generally requires careful design of the experiment from which the observations result so that all $X_i, i=1, ..., n$, are independent.

The question of independence of two random variables X and Y has important implications on their covariance, for if X and Y are independent, then

$$\mathrm{Cov}[X, Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_X(x) f_Y(y) dx dy$$

$$= \int_{-\infty}^{\infty} (x - \mu_X) f_X(x) dx \int_{-\infty}^{\infty} (y - \mu_Y) f_Y(y) dy$$

$$= E[X - \mu_X] E[Y - \mu_Y] = 0 . \qquad (2.4\text{-}13)$$

However, if the covariance of two random variables is zero, it does not necessarily follow that they are independent. One may only suspect that independence is the cause of a zero covariance.

2.4.3 Conditional Probabilities

The marginal probability density function, as developed in equation 2.4-3, can be considered to be a special case of a more general concept referred to as conditioning. Generally speaking, a multivariate probability statement is subject to conditioning when a subset of the random variables pertaining to an experiment falls under some restriction, causing the remaining variables to be conditioned by this restriction. In the case of the marginal density function, we examined the probability that X is less than a, given that Y can take on any value in the interval $(-\infty, \infty)$. Thus, the restriction that Y take on a specific set of values conditions the probability that X is less than a. Formally, we state this as

$$P(X < a \mid -\infty < Y < \infty) .$$

In general, the restriction can be applied to any interval (b,c), where $b < c$, and need not be limited to the interval $(-\infty, \infty)$. However, as in the case of the marginal density, the variable or variables subject to restriction are effectively removed from the probability statement; the variable or variables being conditioned are the ones over which the frequency of occurrence of a specific event may be questioned.

When an experiment which results in a bivariate random variable is conditioned over a range other than $(-\infty, \infty)$, a reduction of the potential sample space available to the experiment results. In the previous example of jointly varying transmissivities X and storativities Y, if we were interested in the conditioned results that X is less than a, given that we are only interested in a specific range of values (b,c) for storativities, then the specific value which Y takes on does not interest us, as long as it falls between b and c. One could proceed as in equation 2.4-3 to evaluate this probability, except for an obvious pitfall: The resulting probability statement over X, where X can take on any value less than a, would not necessarily have the property of cumulative distribution functions noted in equation 2.2-11. That is, as X is the remaining active random variable in the probability statement, its probability of occurrence over the interval $(-\infty, \infty)$ should be unity:

$$\lim_{a \to \infty} P(X < a \mid b < y < c) = 1 .$$

However, by restricting Y to a specific interval (b,c), then as a goes to infinity, an integral of the form of equation 2.4-3 will most probably have a lesser value than unity when the inner integral over y is restricted to a range of something less than $(-\infty, \infty)$. Thus, an integration with the form of equation 2.4-3 alone will not produce a form suitable to serve as a cumulative distribution function for the conditioned variable X.

So that a probability statement resulting from conditioning has the limiting value of unity, these statements must be appropriately normalized. If, as in the bivariate case, we desire $P(X < a \mid b < Y < c)$, then we must normalize by $P(-\infty < X < \infty$ and $b < Y < c)$; that is,

$$P(X < a \mid b < Y < c) = \frac{P(X < a \text{ and } b < Y < c)}{P(-\infty < X < \infty \text{ and } b < Y < c)}$$

$$= \frac{\int_{-\infty}^{a} \int_{b}^{c} f(x,y) dy dx}{\int_{-\infty}^{\infty} \int_{b}^{c} f(x,y) dy dx}$$

$$= \frac{\int_{-\infty}^{a} \int_{b}^{c} f(x,y) dy dx}{\int_{b}^{c} f_Y(y) dy} . \qquad (2.4\text{-}14)$$

Thus, the <u>conditional probability density function</u>, $f(x|b<Y<c)$, for the conditioned random variable X may be defined as

$$f(x|b<Y<c)=\int_b^c f(x,y)dy / \int_b^c f_Y(y)dy \quad , \quad (2.4\text{-}15)$$

which of course gives the limiting value of unity when integrated with respect to x over the interval $(-\infty, \infty)$. Note that when $b=-\infty$ and $c=\infty$, then $f(x|b<Y<c)=f_X(x)$, as indicated by the previous discussion of marginal densities.

Remarkably, the conditional density exists even when the restriction is that, in the example of the bivariate case, Y take on a specific value. To see this easily, consider $P(X<a|Y=c)$; then equation 2.4-15 may be written as

$$f(x|Y=c)=\lim_{\delta\to 0}\frac{\displaystyle\int_c^{c+\delta} f(x,y)dy}{\displaystyle\int_c^{c+\delta} f_Y(y)dy}$$

$$=\lim_{\delta\to 0}\frac{[F(x,c+\delta)-F(x,c)]/\delta}{[F_Y(c+\delta)-F_Y(c)]/\delta}$$

$$=\frac{dF(x,y)/dy}{dF_Y(y)/dy}\bigg|_{y=c}$$

$$=\frac{f(x,c)}{f_Y(c)} \quad . \quad (2.4\text{-}16)$$

Thus, we may recover the density function for X for any particular slice, $Y=c$, through the joint density function $f(x,y)$. If $f(x,y)$ were defined for the example of transmissivity and storativity random variables, equation 2.4-16 would enable us to predict the probability of events concerning transmissivity X for any given value of storativity Y.

The student should also note that some remarkable simplifications result if X and Y are independent random variables. That is, if X and Y are independent, then from equations 2.4-12 and 2.4-14 we see that

$$P(X<a|b<Y<c)=P(X<a) \quad . \quad (2.4\text{-}17)$$

Indeed, this is yet another way in which we can define independence of random variables.

The following problem is intended to familiarize the student with the concept of conditioning; it is not intended to be rigorous. The key to understanding conditioning, especially for discrete random variables, is to understand how it restricts the sample space and realize that the probability of occurrence of an event which contains the entire remaining sample space must be unity.

Problem 2.4–1

a. Given two dice that are thrown sequentially, what is the probability that the first is a three and the second is a two? That is,
  $P(X=3 \text{ and } Y=2)$?

b. What is the probability that the sum of the dice is five? That is,
  $P(X+Y=5)$?

c. Given that the first die is three, what is the probability that the second is two? That is,
  $P(Y=2|X=3)$?

d. Given that the first die is three, what is the probability that the sum of the two dice is five? That is,
  $P(X+Y=5|X=3)$?

e. Given that the first die is three, what is the probability that the sum of the two dice is less than or equal to five? That is,
  $P(X+Y\leq 5|X=3)$?

Parts c, d, and e are conditional probability statements; that is, the probability statement is conditioned by prior information.

2.4.4 Variance of a Column Vector

Our purpose in this section is to develop a representation for the variance of a column vector. As a vehicle to this end, consider the linear equation

$$Y=a_1X_1+a_2X_2+a_3X_3 \quad (2.4\text{-}18)$$

where Y, X_1, X_2, and X_3 are random variables and a_1, a_2 and a_3 are constants. The variance of Y is

$$\mathrm{Var}[Y]=E\{[(a_1X_1+a_2X_2+a_3X_3)$$

$$-E(a_1X_1+a_2X_2+a_3X_3)]^2\}$$

$$=a_1^2\sigma_{X_1}^2+a_2^2\sigma_{X_2}^2+a_3^2\sigma_{X_3}^2$$

$$+2a_1a_2\sigma_{X_1X_2}+2a_1a_3\sigma_{X_1X_3}$$

$$+2a_2a_3\sigma_{X_2X_3} \qquad (2.4\text{-}19)$$

where correlations between X_1, X_2, and X_3 have been allowed for. A vector representation for equation 2.4-19 is[1]

$$\mathrm{Var}[Y]=E[Y^2]-(E[Y])^2=E[\underline{a}\,\underline{X}\,\underline{X}^T\underline{a}^T]$$

$$-E[\underline{a}\,\underline{X}]E[\underline{a}\,\underline{X}] \qquad (2.4\text{-}20)$$

where

$$\underline{a}=[a_1,a_2,a_3]\text{ and }\underline{X}=\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix} \quad .$$

Because expectation is a linear operator, the right side of equation 2.4-20 can be expressed as

$$E[\underline{a}\,\underline{X}\,\underline{X}^T\underline{a}^T]-E[\underline{a}\,\underline{X}]E[\underline{X}^T\underline{a}^T]$$

$$=\underline{a}E[(\underline{X}-E[\underline{X}])(\underline{X}-E[\underline{X}])^T]\underline{a}^T \quad (2.4\text{-}21)$$

where $\underline{a}\,\underline{X}=\underline{X}^T\underline{a}^T$. The expected value of a matrix is the matrix of expected values of each element. Thus,

$$E[(\underline{X}-E(\underline{X}))(\underline{X}-E(\underline{X}))^T]=$$

$$\begin{bmatrix}\sigma_{X_1}^2 & \sigma_{X_1X_2} & \sigma_{X_1X_3}\\ \sigma_{X_1X_2} & \sigma_{X_2}^2 & \sigma_{X_2X_3}\\ \sigma_{X_1X_3} & \sigma_{X_2X_3} & \sigma_{X_3}^2\end{bmatrix} \qquad (2.4\text{-}22)$$

This matrix is defined to be the variance of a 3×1 column vector $\underline{X}$, and allows one to express equation 2.4-19 in matrix notation as

$$\mathrm{Var}[Y]=\underline{a}\,\mathrm{Var}[\underline{X}]\underline{a}^T \quad . \qquad (2.4\text{-}23)$$

---

[1]Throughout this text singly underlined symbols represent vectors and doubly underlined symbols represent matrices.

If the variances $\sigma_{X_1}^2$, $\sigma_{X_2}^2$, and $\sigma_{X_3}^2$ are all equal, the matrix 2.4-22 becomes

$$\mathrm{Var}[\underline{X}]=\begin{bmatrix}1 & \rho_{X_1X_2} & \rho_{X_1X_3}\\ \rho_{X_1X_2} & 1 & \rho_{X_2X_3}\\ \rho_{X_1X_3} & \rho_{X_2X_3} & 1\end{bmatrix}\sigma^2 \quad (2.4\text{-}24)$$

where $\rho_{X_iX_j}$ is the correlation coefficient for X_i and X_j, and σ^2 is the common variance. A further reduction in equation 2.4-22 occurs if X_1, X_2, and X_3 are uncorrelated, causing the correlation coefficients in equation 2.4-24 to be zero. In this case,

$$\mathrm{Var}[\underline{X}]=\underline{I}\sigma^2 \qquad (2.4\text{-}25)$$

where $\underline{I}$ is a 3×3 identity matrix. These forms have practical importance in regression.

Problem 2.4-2

a. Carry out the expectation indicated and show that equation 2.4-19 holds.
b. Demonstrate that equation 2.4-21 holds and that

$$\mathrm{Var}[\underline{a}\,\underline{X}]=\underline{a}\,\mathrm{Var}[\underline{X}]\underline{a}^T \quad .$$

c. Let $Y_i=\underline{a}_i\underline{X}$, where $\underline{a}_i=[a_{i1},a_{i2},a_{i3}]$, and $\underline{X}$, defined as in equation 2.4-20, is a column vector of random variables. Further, let $\underline{Y}=\underline{A}\underline{X}$, where

$$\underline{Y}=\begin{bmatrix}Y_1\\Y_2\\.\\.\\.\\Y_p\end{bmatrix}\quad\text{and}\quad\underline{A}=\begin{bmatrix}\underline{a}_1\\\underline{a}_2\\.\\.\\.\\\underline{a}_p\end{bmatrix} \quad .$$

That is, $\underline{A}$ is a $p\times3$ matrix composed of the row vectors $\underline{a}_i$, $i=1,...,p$. Show that $\mathrm{Var}[\underline{Y}]=\mathrm{Var}[\underline{A}\underline{X}]=\underline{A}\mathrm{Var}[\underline{X}]\underline{A}^T$. (Hint: Equation 2.4-22 still defines the variance of a column vector;

$$\sigma_{Y_iY_j}=E[Y_iY_j]-E[Y_i]E[Y_j]$$

$$=E[\underline{a}_i\underline{X}\,\underline{X}^T\underline{a}_j^T]-E[\underline{a}_i\underline{X}]E[\underline{X}^T\underline{a}_j^T].)$$

2.5 Estimators of Population Parameters

A <u>statistic</u> is defined as any computation from a random sample resulting in a specific value. As such, a statistic is considered to be a random variable, since it is highly probable that the computed value would change from random sampling to random sampling. Note that this definition precludes that a statistic contain any unknown parameters. Estimators of population parameters are considered to be statistics and, therefore, random variables. Consider equation 2.3-1 as an estimator of the mean:

$$\bar{x} = \sum_{i \leq x_m/\Delta x} f_i^* \bar{x}_i \; . \qquad (2.5\text{-}1)$$

The estimated frequency f_i^* is computed from values of observations x_i originating from a random sampling of the sample space. However, prior to sampling, a random sample is merely an abstract collection of random variables X_i, $i=1,...,n$. Any function of random variables, as equation 2.5-1 would be prior to sampling, is also a random variable, perhaps having a completely different distribution than those individuals composing the collection.

Our discussion of statistics will largely be from the a priori viewpoint; that is, in the case of equation 2.5-1, $\bar{x}$ is the value of the random variable $\bar{X}$, which is an estimator for the population mean, as developed from some arbitrary random sample.

2.5.1 Mean Estimator

As an estimator for the population mean, equation 2.5-1, in addition to being cumbersome to compute, has the debility that it is dependent upon an arbitrary selection of a class interval. That is, since f_i^* is dependent upon Δx, the value of $\bar{X}$ will depend upon the choice of Δx used in the computation. As a means of pursuing this problem, assume that we have at our disposal a random sample consisting of n observations, and at some point their distribution appears as in figure 2.5-1. Because Δx is arbitrary, it can be reduced to δ, the minimum of all differences in neighboring values of the random

variable. Then f_i^* would take on only two values, $1/n$ or 0, and would have the ragged saw-toothed shape shown in figure 2.5-2. The teeth in figure 2.5-2 will be concentrated in regions where f_i^* in figure 2.5-1 is larger. Of course, that a repeat value of a random variable could occur is highly unlikely, as the probability of such an occurrence is essentially zero for a continuous random variable (see equation 2.2-14). By using the class interval δ of figure 2.5-2 in our computation, we see that the problem of estimating frequency weights for $\bar{x}_i$, the central location in each class interval, from a discrete data set has essentially been removed, as these weights now take on only two specific values for this and any other smaller class interval.

The use of the smaller class interval, δ, is expected to produce a better estimator of the population mean because a value of $\bar{X}$ would contain less measurement error associated with the arbitrary selection of the class interval. It is still cumbersome, however, to calculate the central value $\bar{x}_i$ of these possibly very small class intervals, especially when one considers that many do not contribute to the estimator. We ask ourselves if it is not possible to use the observations x_i in their place. By reducing Δx even further until every observation is isolated in the center of its own infinitesimally small class interval, in which case f_i^* would remain at the $1/n$ level, observations x_i can be used in place of $\bar{x}_i$ in equation 2.5-1 without significantly altering the basis of our estimator. Using future observations X_i of the process in place of central-interval values, $\bar{x}_i$, an estimator based on an infinitesimally small interval would appear as

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (2.5\text{-}2)$$

where n is the size of the random sample, and X_i, $i=1,...,n$, is the collection of random variables from the random sample. Equation 2.5-2 is the preferred estimator for the population mean.

A sample statistic is said to be <u>unbiased</u> if its expected value is equal to the population parameter that it estimates. Consider the expected value of the sample mean, derived from random variables X_i, $i=1,...,n$. Since $E[X_i]=\mu_X$,
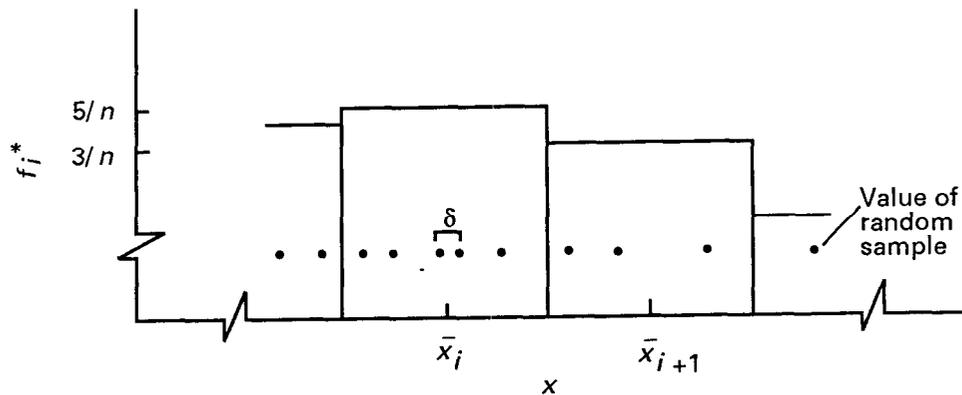
Figure 2.5-1
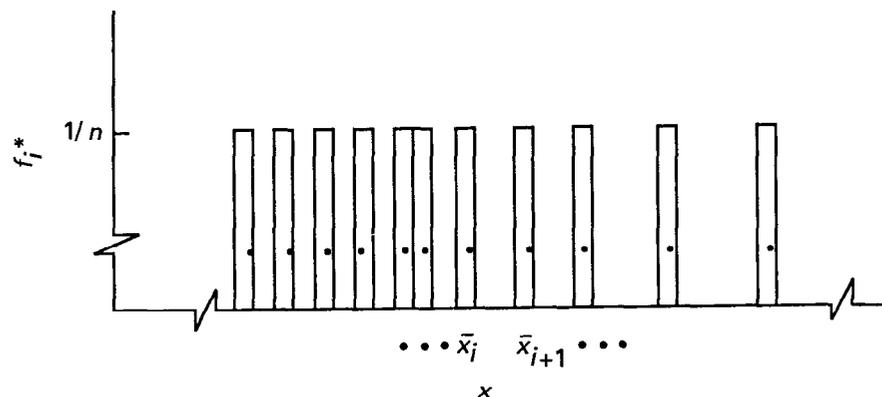
Figure 2.5-2

$$E[\bar{X}] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \sum_{i=1}^{n} \mu_X = \mu_X \ . \qquad (2.5\text{-}3)$$

Hence, $\bar{X}$ is an unbiased estimator of μ_X. Although examining an estimator for unbiased qualities is important, it does not necessarily insure that the estimator is the most efficient (or best) in the sense that the variance of the estimator is the smallest. It is, however, an important quality, and the variance estimator is examined for this quality in the next section.

Problem 2.5-1

a. Recompute the sample mean for the data set in problem 2.2-2 using equation 2.5-2 as the mean estimator. How does this result vary from that of problem 2.3-1?

How do you, in light of equation 2.2-14, explain the repeat values in the data set (note that this data set represents a random sampling of a continuous random variable)?

b. With regard to a large regional aquifer, well data such as that in table 2.2-2 represent point estimates of transmissivities. The best estimate of the effective transmissivity (the one to use in modeling the flow field) is generally considered to be the geometric mean of these point estimates. The statistic for the geometric mean is defined as

$$\bar{T}_g = (T_1 \cdot T_2 \cdot T_3 \cdot \cdots \cdot T_n)^{1/n}$$

where T_i, $i=1,...,n$, is a random sample from the sample space of the T random variable. Letting $X_i = \log_{10} T_i$, we see that

$$\log_{10}[\bar{T}_g]= \frac{1}{n}\log_{10}(T_1 \cdot T_2 \cdot T_3 \cdots \cdots T_n)$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \quad .$$

Therefore,

$$\bar{T}_g = 10^{\bar{X}} \quad .$$

What is the geometric mean of the transmissivity data in table 2.2-2?

As a measure of the dispersion about the geometric mean, one could use the estimator

$$D_g = 10^{(\bar{X} \pm S_x)} \quad .$$

What is the dispersion D_g about the geometric mean? Use results of problem 2.3-2, part b, as values for $\bar{X}$ and S_X^2. Considering the dispersion, how do you feel about $\bar{t}_g$ being the effective transmissivity of the carbonate rocks of central Pennsylvania?

2.5.2 Variance Estimator

As an estimator of the variance σ_X^2, consider using an estimator S_X^{*2} whose value s_X^{*2} is calculated from the equation

$$s_X^{*2} = \sum_{i \le x_n/\Delta x} f_i^* (\bar{x}_i - \mu_X)^2 \quad , \tag{2.5-4}$$

which is analogous to equation 2.3-12 for the population parameter. If the class interval is taken to be small enough so as to isolate every future observation X_i in a class interval, then equation 2.5-4 can be rewritten in terms of random observations X_i, $i=1,...,n$, as

$$S_X^{*2} = \frac{1}{n}\sum_{i=1}^{n} (X_i - \mu_X)^2 \quad , \tag{2.5-5}$$

because $f_i^* = 1/n$ and $X_i = x_i$ prior to sampling. When the underlying population X_i, $i=1,...,n$, is normally distributed, it can be shown that equation 2.5-5 is the most efficient, unbiased

estimator of the variance σ_X^2 in the sense that its variance is the least of all possible unbiased estimators for σ_X^2.

On occasion, the population mean μ_X can be determined from other considerations, as was done in the titration experiment in section 2.2.5. However, usually μ_X is also unknown, requiring that μ_X be replaced by $\bar{X}$ in equation 2.5-5:

$$S_X^{*2} = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2 \quad . \tag{2.5-6}$$

To test whether equation 2.5-6 is an unbiased estimator of σ_X^2, the expected value of S_X^{*2} is determined. The actual mechanics of this operation are presented in appendix 2.11.2; only the result is presented here:

$$E(S_X^{*2}) = \frac{n-1}{n}\sigma_X^2 \quad . \tag{2.5-7}$$

Thus, S_X^{*2} is a biased estimator of σ_X^2. To produce an unbiased estimator of σ_X^2, S_X^{*2} is multiplied by the ratio $n/(n-1)$:

$$S_X^2 = \frac{n}{n-1} S_X^{*2} = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2 \quad . \tag{2.5-8}$$

This estimator is unbiased but less efficient than equation 2.5-6. However, it is the preferred estimator for small sample sizes.

Heuristically, one can argue that this adjustment to the estimator is necessary, because the population mean μ_X is being estimated by the sample statistic $\bar{X}$. The sample mean will be located at the centroid of the random sample, regardless of whether its value is near that of the population mean. Thus, an equation that estimates the variance about this centroid will produce a smaller value than if the estimate were made about the population mean. The adjustment, then, merely compensates for the smaller deviates produced by using $\bar{X}$ in place of μ_X.

Equation 2.5-8 can be rewritten, with the aid of some algebraic manipulation, to produce a slightly more useful form for hand calculations:

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - 2\bar{X} \sum_{i=1}^{n} X_i + n\bar{X}^2 \right]$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \right] . \qquad (2.5\text{-}9)$$

The estimator for the standard deviation is taken to be the square root of the variance estimator. Values x_i, $i=1,...,n$, obtained by sampling the population of X randomly, are used in place of X_i in equation 2.5–9 to obtain a value s_X^2 for the sample statistic S_X^2.

2.5.3 Estimator of Correlation Coefficient

In a manner analogous to the variance, an estimator for the covariance, and therefore the correlation coefficient, can be derived. Let R_{XY} represent the estimator for the correlation coefficient ρ_{XY}; then, for paired data,

$$R_{XY} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \right]^{\frac{1}{2}}} \qquad (2.5\text{-}10)$$

or, provided that S_X and S_Y originate from the paired data,

$$R_{XY} = \frac{1}{(n-1)S_X S_Y} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) , \qquad (2.5\text{-}11)$$

which can be written, for purposes of hand calculation, as

$$R_{XY} = \frac{1}{(n-1)S_X S_Y} \left(\sum_{i=1}^{n} X_i Y_i - n\bar{Y}\bar{X} \right) \qquad (2.5\text{-}12)$$

where S_X and S_Y are calculated by taking the square root of either equations 2.5–8 or 2.5–9. The actual value r_{XY} of R_{XY} is obtained by

using values x_i, $i=1,...,n$, from a random sample in place of X_i.

2.5.4 Summary

In summary, population parameters and equivalent sample statistics can be tabulated as follows:

Population parameter	Sample statistic
μ	$\bar{X} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$
σ_X^2	$S_X^2 = \dfrac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \right]$
V_X	$C_X = S_X / \bar{X}$
ρ_{XY}	$R_{XY} = \dfrac{1}{(n-1)S_X S_Y} \left(\sum_{i=1}^{n} X_i Y_i - n\bar{Y}\bar{X} \right)$

These estimators can also be stated in matrix form. For instance, let $d_i = x_i - \bar{x}$; then a value for S_X^2 is

$$s_X^2 = \frac{1}{n-1} \underline{d}^T \underline{d} \qquad (2.5\text{-}13)$$

where $\underline{d}$ is a column vector of deviates and $\underline{d}^T$ is its transpose. If $e_i = y_i - \bar{y}$, then a value for R_{XY} is

$$r_{XY} = \frac{\underline{e}^T \underline{d}}{(n-1)s_X s_Y} . \qquad (2.5\text{-}14)$$

Forms similar to equations 2.5–13 and 2.5–14 are commonly encountered in linear regression.

Problem 2.5–2

Using the following data set, calculate the sample mean, variance, and standard deviation of both dissolved solids and specific conductance; then calculate their correlation coefficient.

Specific conductance and dissolved solids data for
wells in carbonate rocks of Maryland

[From Nutter, 1973, p. 63–68]

Specific conductance (μmho/cm)	Dissolved solids (ppm)
278	257
1,120	610
533	338
723	458
462	264
1,030	562
357	231
304	175
469	268
641	388
969	638
876	532
721	405
895	610
501	304
323	171
310	201
1,230	736
504	290
319	208
704	464
1,130	688
600	342

2.6 Transformation of Random Variables

As we have noted previously, statistics are combinations of random variables and, as such, must be random variables themselves. If the population from which the random sample is selected can be identified, then one can frequently identify the probability density functions of statistics, which are estimators of the population parameters. If a density function is identified, then one should be able to develop criteria for testing the accuracy of these estimators. With these objectives in mind, we proceed to identify density functions that result from the several types of transformations that produce statistics.

Before proceeding with this identification process, we make note of two general results from expectation which are applicable to all random variables, regardless of their distribution. In general, if $X_1, X_2,...,X_n$ are independent variables with identical mean μ_X and identical standard deviation σ_X, then

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (2.6\text{-}1)$$

is also a random variable with mean

$$\mu_{\bar{X}} = E[\bar{X}] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \mu_X \qquad (2.6\text{-}2)$$

and variance

$$\sigma_{\bar{X}}^2 = \text{Var}[\bar{X}] = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}[X_i] = \sigma_X^2/n \ . \quad (2.6\text{-}3)$$

Equation 2.6–2 was used previously to show that $\bar{X}$ is an unbiased estimator of μ_X, and equation 2.6–3 is demonstrated more fully in appendix 2.11.2. Note that equation 2.6–3 only succeeds because $\text{Cov}[X_i, X_j] = 0$, $i \neq j$; that is, the X_i's are independent.

The square root of equation 2.6–3, $\sigma_{\bar{X}}$, is also known as the standard error of $\bar{X}$. The standard deviation of any statistical measure is referred to as the standard error of that statistic.

2.6.1 Sum of Independent Normal Random Variables

Let X_1 and X_2 be independent normal random variables, X_1 with mean zero and variance one ($N(0,1)$) and X_2 with mean zero and variance k ($N(0,k)$). How, then, is their sum distributed? To answer this question, consider

$$P(Y \leq y) = P(X_1 + X_2 \leq y) \ ,$$

where $Y = X_1 + X_2$. By noting that $P(X_1 + X_2 \leq y) = P(X_1 \leq y - X_2$ and $-\infty \leq X_2 \leq \infty)$, comparison with equation 2.4–11 shows that

$$F_Y(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y-x_2} \frac{e^{-x_1^2/2}}{\sqrt{2\pi}} \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_1 dx_2 \ ,$$

$$= \int_{-\infty}^{\infty} F_{X_1}(y-x_2) \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_2 \ . \qquad (2.6\text{-}4)$$

To find the probability density function of Y, we differentiate $F_Y(y)$ with respect to y; that is

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \int_{-\infty}^{\infty} \frac{d}{dy} F_{X_1}(y-x_2) \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_2$$

$$= \int_{-\infty}^{\infty} \frac{e^{-(y-x_2)^2/2}}{\sqrt{2\pi}} \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_2$$

$$= \frac{1}{2\pi\sqrt{k}}$$

$$\cdot \int_{-\infty}^{\infty} \exp[-(y^2 - 2yx_2 + (k+1)x_2^2/k)/2] dx_2 \quad (2.6\text{-}5)$$

which, after some algebraic manipulation, yields

$$f_Y(y) = \frac{\exp[-y^2/(2k+2)]}{2\pi\sqrt{k}}$$

$$\cdot \int_{-\infty}^{\infty} \exp\left[-\left(x_2\sqrt{\frac{k+1}{2k}} - y\sqrt{\frac{k}{2k+2}}\right)^2\right] dx_2. \quad (2.6\text{-}6)$$

By letting $u = \sqrt{2}\left(x_2\sqrt{\frac{k+1}{2k}} - y\sqrt{\frac{k}{2k+2}}\right)$, then

$$f_Y(y) = \frac{\exp[-y^2/(2k+2)]}{\sqrt{2\pi(k+1)}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2} du$$

$$= \frac{\exp[-y^2/(2k+2)]}{\sqrt{2\pi(k+1)}} \quad (2.6\text{-}7)$$

which follows from equation 2.2–11. Thus, the sum of two independent zero-mean normal random variables, one with variance unity and the other with variance k, is a normal random variable with variance $k+1$; that is, $N(0,k+1)$.

If, in the previous problem, k were to equal one, then we see that the sum of two $N(0,1)$ independent random variables is a $N(0,2)$ random variable. By adding yet another independent $N(0,1)$ random variable to the previous two $N(0,1)$ random variables, induction tells us that a $N(0,3)$ random variable results. Thus, in general, the sum of n independent $N(0,1)$ random variables results in a $N(0,n)$ random variable.

We are now in a position to determine the distribution of the statistic $\bar{X}$, as shown in equation

2.6–1, if $\bar{X}$ is determined from a random sample in which all the observations X_i, $i=1,...,n$, are independent normal random variables with common mean μ_X and variance σ_X^2; that is, $N(\mu_X, \sigma_X^2)$. We note from equation 2.6–3 that $\bar{X}$ has the standard deviation $\sigma_X/\sqrt{n}$. If we standarize $\bar{X}$ by its mean and standard deviation, and multiply this result by $\sqrt{n}$, then

$$\sqrt{n}\left(\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}}\right) = \sqrt{n}\left(\frac{\sum_{i=1}^{n}X_i - n\mu_X}{\sqrt{n}\sigma_X}\right) = \sum_{i=1}^{n}\left(\frac{X_i-\mu_X}{\sigma_X}\right) \quad (2.6\text{-}8)$$

results. We see that this new statistic is the sum of n normal random variables, each with mean zero and variance one. From the previous paragraph, equation 2.6–8 must be a normal random variable with mean zero and variance n. To obtain a random variable with mean zero and variance unity, one would divide equation 2.6–8 by the square root of n. By inspection, then, the quantity $(\bar{X}-\mu_X)/(\sigma_X/\sqrt{n})$ must be a standard normal random variable, and $\bar{X}$ must be normal with mean μ_X and variance σ_X^2/n. Thus, if one knew that a random sample were composed of normal random variables with a particular mean and variance, then one could investigate the probability that a future determination of the sample mean could take on a particular range of values.

2.6.2 The Chi-Square Distribution

We are frequently concerned with the square of a random variable and may wish to know its density function. Assuming that the random variable X is normally distributed with mean zero and variance one ($N(0,1)$), we may inquire as to the nature of the distribution of its square, $Y=X^2$. Proceeding as in the previous section, we find the cumulative distribution of Y:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \quad . \quad (2.6\text{-}9)$$

By taking the derivative of $F_Y(y)$, one finds the density function of Y:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{\sqrt{2\pi}} \frac{e^{-y/2}}{\sqrt{y}}, \; y \geq 0 \quad (2.6\text{-}10)$$

which is the chi-square density function with 1 degree of freedom.

Chi-square random variables have a useful additive property similar to that exhibited by independent normals. Namely, if Y_1 and Y_2 are independent chi-square random variables with degrees of freedom ν_1 and ν_2, then $Y_1 + Y_2$ is a chi-square random variable with degrees of freedom $\nu_1 + \nu_2$. Consequently, if $Y_1, Y_2, ..., Y_n$ are independent chi-square random variables each with 1 degree of freedom, then $Y_1 + Y_2$ is a chi square with 2 degrees of freedom, $(Y_1 + Y_2) + Y_3$ is a chi square with 3 degrees of freedom and, in general, $\sum_{i=1}^{n} Y_i$ is a chi square with n degrees of freedom. Values for the cumulative distribution function of the chi-square distribution with ν degrees of freedom are to be found in table 2.10-2.

If X_i, $i=1,...,n$, are independent normal random variables, each with mean μ_X and variance σ_X^2, then $\sum_{i=1}^{n} ((X_i - \mu_X)/\sigma_x)^2$ must be a chi-square random variable with n degrees of freedom. This follows from the previous argument by letting $Y_i = (X_i - \mu_X)^2/\sigma_X^2$ and noting that Y_i is the square of $N(0,1)$ random variable. Furthermore, because

$$\sum_{i=1}^{n} \frac{(X_i - \mu_X)^2}{\sigma_X^2} = n\frac{(\bar{X} - \mu_X)^2}{\sigma_X^2} + \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma_X^2}, \; (2.6\text{-}11)$$

the statistic S_X^2 can be written in terms of this sum as

$$\frac{(n-1)S_X^2}{\sigma_X^2} = \sum_{i=1}^{n} \frac{(X_i - \mu_X)^2}{\sigma_X^2} - \frac{(\bar{X} - \mu_X)^2}{(\sigma_X/\sqrt{n})^2} \quad . \quad (2.6\text{-}12)$$

Under the condition that the underlying population is independent and normal, it was demonstrated in section 2.6.1 that $\bar{X}$ is normal with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$. Thus, $(\bar{X} - \mu_X)^2/(\sigma_X^2/n)$, under this condition, is chi square with 1 degree of freedom. One might reasonably expect, then, that

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2(\nu) \quad\quad (2.6\text{-}13)$$

is a chi-square random variable with $\nu = n-1$ degrees of freedom, which is indeed the case when the underlying population of X_i's are independent normal random variables.

2.6.3 The F Distribution

The density function for the ratio of two independent chi-square random variables can be calculated rather easily by the method used in the previous sections. However, because we have little need of the actual form of this density function, known as the F distribution, we relieve the student of working through the actual calculation if he or she will accept the following statement: If X_1 is a chi-square random variable with ν_1 degrees of freedom, and X_2 is a chi-square random variable with ν_2 degrees of freedom, and X_1 and X_2 are independent, then

$$\frac{X_1/\nu_1}{X_2/\nu_2} \sim F(\nu_1, \nu_2) \quad\quad (2.6\text{-}14)$$

defines the F distribution with ν_1 and ν_2 degrees of freedom.

Table 2.10-3 is a tabulation of values of $F(\nu_1, \nu_2)$ which satisfies the probability statement

$$P(F(\nu_1, \nu_2) \leq F_\alpha(\nu_1, \nu_2)) = 1 - \alpha \quad , \quad (2.6\text{-}15)$$

where α equals 0.05; the meaning of equation 2.6-15 is illustrated in figure 2.6-1. Note that the reciprocal of an entry $F_\alpha(\nu_1, \nu_2)$ in table 2.10-3 is equal to $F_{1-\alpha}(\nu_2, \nu_1)$. That is, if equation 2.6-15 holds, then

$$P(F(\nu_2, \nu_1) > F_\beta(\nu_2, \nu_1))$$
$$= P(1/F(\nu_2, \nu_1) \leq 1/F_\beta(\nu_2, \nu_1))$$
$$= P(F(\nu_1, \nu_2) \leq 1/F_\beta(\nu_2, \nu_1))$$
$$= \beta \quad , \quad\quad (2.6\text{-}16)$$

as, by equation 2.6-14, $1/F(\nu_2, \nu_1)$ is an $F(\nu_1, \nu_2)$ random variable. By comparing equation 2.6-15 with the third line in equation 2.6-16, we see that, when β equals $1-\alpha$,

$$F_\alpha(\nu_1, \nu_2) = \frac{1}{F_{1-\alpha}(\nu_2, \nu_1)} \quad . \quad (2.6\text{-}17)$$
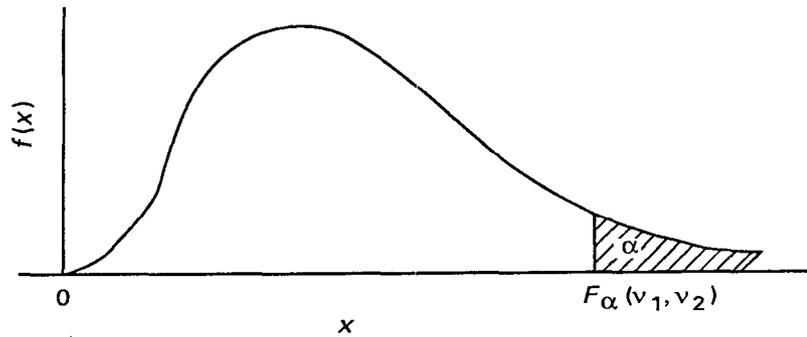
Figure 2.6-1

Thus, if we wish to evaluate $F_{1-\alpha}(n_1, n_2)$ for the statement

$$P(F(n_1,n_2) \leq F_{1-\alpha}(n_1,n_2)) = \alpha \qquad (2.6\text{-}18)$$

where α is the relative mass indicated in figure 2.6-2, then we need only find $F_\alpha(\nu_1,\nu_2)$, where $\nu_1 = n_2$ and $\nu_2 = n_1$, in a table of values for the F distribution and calculate its reciprocal.

As an example of a practical statistic associated with the $F(\nu_1,\nu_2)$ random variable, consider two random samples of size n_1 and n_2, which have been selected from two normal populations with variances σ_1^2 and σ_2^2, respectively. Let

$$X_1 = (n_1-1)S_1^2/\sigma_1^2 \qquad (2.6\text{-}19)$$

and

$$X_2 = (n_2-1)S_2^2/\sigma_2^2 \qquad (2.6\text{-}20)$$

where S_1^2 and S_2^2 are sample variances that are independent, because they originate from separate random samples. From equations 2.6-13 and 2.6-14, we see that

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(\nu_1,\nu_2) \qquad (2.6\text{-}21)$$

is an $F(\nu_1,\nu_2)$ random variable with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom.

If σ_1^2 were to equal σ_2^2, then equation 2.6-21 would undergo an obvious simplification. As a case in point, consider the ratio of $(\bar{X}-\mu_X)^2/(\sigma_X^2/n)$, which is the square of a $N(0,1)$ random variable, and $(n-1)S_X^2/\sigma_X^2$, which is a

$\chi^2(n-1)$ random variable, where $\bar{X}$ and S_X^2 are statistics developed from the same population. One can show (rather arduously) that $\bar{X}$ and S_X^2 are independent, even though they originate from the same random sample. Thus,

$$\frac{(\bar{X}-\mu_X)^2/(\sigma_X^2/n)}{S_X^2/\sigma_X^2} = \frac{(\bar{X}-\mu_X)^2}{S_X^2/n} \sim F(1,n-1) . \quad (2.6\text{-}22)$$

The square root of equation 2.6-22 is also known as a T random variable with $n-1$ degrees of freedom. However, as the T random variable is, in general, equal to the square root of an $F(\nu_1, \nu_2)$ random variable with $\nu_1 = 1$, no additional time will be devoted to it.

Problem 2.6-1

Residuals ε from a titration experiment (section 2.2.5) have the following values in moles of acid:

$$-0.011, \quad +0.003, \quad +0.004, \quad -0.01, \quad +0.005,$$

$$+0.014, \quad +0.004, \quad +0.001, \quad -0.01, \quad +0.003.$$

Calculate $\bar{\varepsilon}$ and s_ε^2 from this random sample. Assume $\mu_\varepsilon = 0$; from equation 2.6-22 derive the probability statement

$$P\left(-\sqrt{F_\alpha(1, n-1)} \leq \bar{\varepsilon}/(S_\varepsilon/\sqrt{n}) \leq \sqrt{F_\alpha(1, n-1)}\right) = 1-\alpha.$$

(Hint: $a^2 \leq b$ is equivalent to $-\sqrt{b} \leq a \leq \sqrt{b}$). Find the interval corresponding to this statement when $\alpha = 0.05$ (that is, go to table 2.10-3 and find $F_\alpha(1,n-1)$ and then calculate $-\sqrt{F_\alpha(1,n-1)}$ and $\sqrt{F_\alpha(1,n-1)}$). How does the value of the statistic
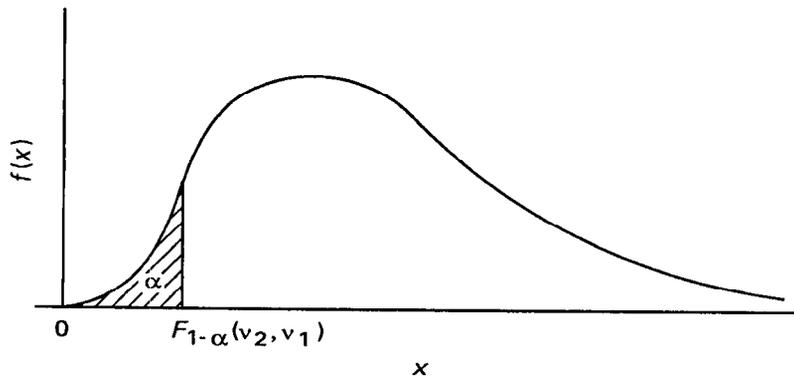
Figure 2.6-2

$\bar{\varepsilon}/(S_\varepsilon/\sqrt{n})$, calculated from the above random sample, compare with this interval? Would you expect this result? Would it bother you if the value fell outside the interval?

2.7   Central Limit Theorem

An interesting and difficult-to-prove theorem of statistics and probability, known as the Central Limit Theorem, concerns the sum of random variables:

Let $X_1, X_2,..., X_n$ be a sequence of identically distributed, independent random variables each with mean μ_X and variance σ_X^2. Then, the distribution of

$$\frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}}$$

tends to a standard normal random variable as n goes to infinity. That is,

$$\lim_{n \to \infty} P\left(\frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}} \leq a\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-t^2/2} dt \qquad (2.7\text{-}1)$$

regardless of the distribution of X_i, $i=1,...,n$.

The key to understanding the impact of this theorem is to realize that the underlying distribution of the X_i random variables can be that of any random variable. Thus, for instance, a chi-square random variable with n degrees of freedom is the sum of other chi-square random variables, and, if n becomes large enough, the chi-square random variable approaches the normal random variable (in fact, the normal tables are used to approximate the chi-square random variable for large values of n).

Another result of this theorem concerns the robustness of some of the distributions developed in the previous section. In particular, if the underlying distribution of X_i's was not normal in equation 2.6-22, this statistic would still be approximately an $F(1,n-1)$ random variable, provided the sample size n were large. The argument for this statement proceeds as follows: Because the numerator of equation 2.6-22 is, when n is large, the square of an approximately normal random variable (by the central limit theorem), it will tend to be a chi-square distributed random variable with 1 degree of freedom. The denominator, on the other hand, will approach unity for large n, because another law of probability dictates that, as n becomes large, S_X^2 approaches σ_X^2 (this phenomenon occurs regardless of the underlying distribution). The net result is that, regardless of the distribution of the X_i random variables, equation 2.6-22 tends, for large n, to be a chi-square distributed random variable with 1 degree of freedom. However, one can show that, for large ν_2, the $F(\nu_1,\nu_2)$ random variable (equation 2.6-14) tends to a chi-square random variable whose value has been diminished by a factor of $1/\nu_1$. Thus, regardless of the underlying distribution of the X_i random variables, we say that equation 2.6-22 behaves asymptotially as an $F(1,n-1)$ random variable when n is large, as both equations 2.6-22 and 2.6-14 have the same distribution for the limiting case where the degrees of freedom in the denominator become large.

2.8 Confidence Limits

We have already noted that statistics are random variables themselves. Now we wish to use the information developed in the previous sections concerning the form of these random variables to make a statement about the reliability of these statistics as estimators. We attempt to define an interval, based upon the statistic, such that a certain percentage of all such intervals, as constructed from different random samples, contain the population parameter that the statistic is thought to estimate. For example, if 5/6 of all possible intervals constructed from repeated sampling contain the population parameter Θ, then there is a probability of 5/6 that the interval we construct from any given random sample actually contains Θ (see figure 2.8-1).

As an example of the interval-construction process, consider the statistic $\bar{X}$ and the population parameter μ_X. We know, from the central limit theorem, that this statistic is approximately normally distributed with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$ and that

$$\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}} \sim N(0,1) \qquad (2.8-1)$$

is approximately true. Of course, when $\bar{X}$ is based on a random sampling of a normal population, then equation 2.8-1 is exactly true. This standard normal random variable will be used to devise a $(1-\alpha)100\%$ <u>confidence interval</u> for

μ_X. This objective is achieved by first looking at the probability statement

$$P(-N_{\alpha/2}(0,1){\leq}N(0,1){\leq}N_{\alpha/2}(0,1))=1-\alpha \quad (2.8-2)$$

and finding the values $\pm N_{\alpha/2}(0,1)$ which correspond to $1-\alpha$. This probability statement says that, $(1-\alpha)100\%$ of the time, a value of $N(0,1)$, obtained from a repetition of the experiment, will fall between $-N_{\alpha/2}(0,1)$ and $N_{\alpha/2}(0,1)$. Assuming that σ_X is known, and with a little help from equation 2.8-1, equation 2.8-2 can be rewritten as

$$P\left(\bar{X}-N_{\alpha/2}(0,1)\,\frac{\sigma_X}{\sqrt{n}} \leq \mu_X \leq \bar{X}+N_{\alpha/2}(0,1)\,\frac{\sigma_X}{\sqrt{n}}\right)$$

$$=1-\alpha \ . \qquad (2.8-3)$$

This probability statement says that, $(1-\alpha)100\%$ of the time, the interval $(\bar{X}-N_{\alpha/2}(0,1)\sigma_X/\sqrt{n},\ \bar{X}+N_{\alpha/2}(0,1)\sigma_X/\sqrt{n})$ constructed with a value of $\bar{X}=\bar{x}$ from a particular random sample will contain μ_X. Thus,

$$\bar{x}-N_{\alpha/2}(0,1)\,\frac{\sigma_X}{\sqrt{n}} \leq \mu_X \leq \bar{x}+N_{\alpha/2}(0,1)\,\frac{\sigma_X}{\sqrt{n}} \ (2.8-4)$$

is a $(1-\alpha)100\%$ confidence interval for a random sample of size n, whose variance is known and whose sample mean $\bar{x}$ can be calculated. The investigator would be able to say that the probability is $1-\alpha$ that this interval contains μ_X; however, in interpreting this statement, one
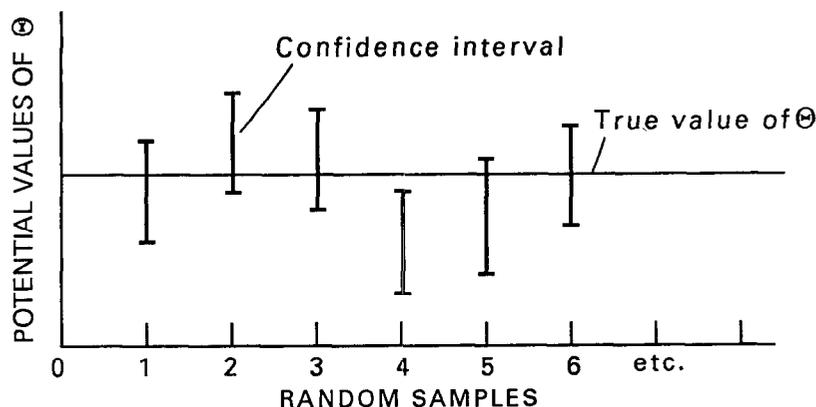
Figure 2.8-1

should realize that the interval is random, not μ_X.

As an actual numeric example of an application of equation 2.8-4, consider the following data: $\sigma_X = 0.3$, $\bar{x} = 2.6$, $n = 36$. Find a 95% confidence interval. From table 2.10-1, for $\alpha/2 = 0.025$, we see that $N_{\alpha/2}(0,1) = 1.96$. Hence a 95% confidence interval is

$$2.6 - (1.96)(0.3/\sqrt{36}) \leq \mu_X \leq 2.6 + (1.96)(0.3/\sqrt{36})$$

or

$$2.50 \leq \mu_X \leq 2.70 \ .$$

Thus, as 95 of 100 intervals so constructed contain the mean, there is a 95% probability that this one contains μ_X.

If σ_X is not known, equation 2.8-4 cannot be used. However, if the underlying population is nearly normal, then σ_X^2 can be estimated by S_X^2 as discussed in section 2.6.2 and one can use either the T distribution (with $n-1$ degrees of freedom) or the F distribution as given in equation 2.6-22 to make an appropriate probability statement that can be converted to an interval on μ_X.

Problem 2.8-1

    a. Seven gold assays from stockpiled ore are: 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 grams per metric ton. Find a 95% confidence interval for the mean grade of the ore assuming an approximate normal distribution (hint: use equation 2.6-22; why?).

    b. Write an interpretation of this interval.

2.9 Hypothesis Testing

Assume that you have determined by a method which you consider to be very good that a population parameter Θ should take on a particular range of values. On the other hand, another independent source suggests that the parameter Θ should take on a value b, which lies outside this range. This discrepancy is disconcerting, and you need some method of testing this independent estimate of Θ. You construct a hypothesis, referred to as the null hypothesis,

H_0, that b is the true value of the parameter; symbolically this may be stated:

$$H_0:\Theta = b.$$

Ideally, of course, you wish to reject this hypothesis, but a procedure is needed whereby you can approach the problem objectively. When a random sample is available to the investigator, hypothesis tests can provide this procedure.

If, from a random sample $X_1, X_2,...,X_n$, a test statistic ψ can be constructed which, in some manner, is a measure of Θ, then often a statistical method can be devised to test the probable veracity of the null hypothesis. It is assumed that the distribution of the test statistic is known under the assumption of the null hypothesis or at least can be approximated. Of course, by definition a statistic cannot contain any unknown parameters. Any population parameters which it may contain must be known either by hypothesis or some other means; otherwise ψ would cease to be a statistic. The statistical test will consist of finding a critical interval with a low probability of occurrence under the null hypothesis such that, should a value of ψ as determined from a random sample fall into this interval, the null hypothesis would be rejected and the alternate hypothesis H_1, which usually consists of one of the following, would be accepted:

$$H_1:\Theta > b$$

$$H_1:\Theta < b$$

$$H_1:\Theta \neq b.$$

The alternate hypothesis chosen depends on the nature of the test. A method of intelligently selecting critical intervals must be devised before the test can be completed, for not any arbitrary interval with a small probability of occurrence will do.

2.9.1 Type I Error

In hypothesis-testing procedures, one is ultimately concerned with the possibility of rejecting the null hypothesis when it is true. The objective of hypothesis testing is to make as

small as possible the probability of committing this error, referred to as a type I error. That is, the probability statement

$$P(\text{reject } H_0|H_0 \text{ true})=\alpha \qquad (2.9-1)$$

is constructed, and α, the level of significance of the test, is chosen as small as the investigator deems reasonable. For continuous random variables, the probability α must be associated with some interval about the test statistic ψ, the statistic fulfilling the requirements of the null hypothesis. Generally speaking, the test statistic ψ will contain an estimator $\hat{\Theta}$ of the population parameter Θ. If $\hat{\Theta}$, as evaluated from some arbitrary random sampling of the experiment, were to have a value close to b, the assumed value of Θ under the null hypothesis, we would not expect to reject the null hypothesis. Rather, only when this value of $\hat{\Theta}$ was distant from b would the null hypothesis be rejected. Thus, the logical choice of an interval in ψ would be one in which all possible values of $\hat{\Theta}$ used in the calculation of ψ would be as distant as possible from b. When the distribution of ψ has infinite tails, then this procedure will cause the interval to include one or both tails, depending on the nature of the test. This interval, whose exact starting and (or) ending point(s) will be determined by the significance level α of the test, will correspond to the critical region where H_0 will be rejected should a calculated value of ψ fall into this region. In most cases, this procedure will cause the critical interval to obtain its maximum length at the chosen significance level α.

When transforming equation 2.9-1 into a probability statement over the test statistic ψ, it is often preferable to first consider the impact of the alternate hypothesis on $\hat{\Theta}$. Consider again the null hypothesis where $\Theta=b$; then, should $\Theta>b$ properly represent the alternate hypothesis, it is useful to consider that, heuristically if not exactly, the probability of committing a type I error can be stated $P(\hat{\Theta}>a|H_0)$, where a is some value of $\hat{\Theta}$ such that $b<a<\infty$, thus giving one an understanding that $\hat{\Theta}$ must take on, relatively, a large positive value in order for H_0 to be rejected. For this alternate hypothesis, a more accurate statement of equation 2.9-1 usually takes the form

$$P(\psi>c|\Theta=b)=\alpha , \qquad (2.9-2)$$

as the distribution of ψ is always assumed to be known.

To complete the above test, a value for c corresponding to α is obtained from a table of cumulative probabilities. Values of ψ less than c correspond to a region where the probability of committing a type I error may not be small. Therefore, if a value of ψ is less than c, we are forced to accept the null hypothesis to avoid committing a type I error. If this value is larger than c, then the probability of committing this error is considered small, and we can confidently reject H_0 at the α significance level.

Examples of hypothesis testing, which should clarify the actual mechanics of the procedure, are presented subsequently; however, before proceeding to these examples, note that we are frequently required to play the role of the devil's advocate in hypothesis testing. Often, we really desire to test the acceptability of a hypothesized value of a parameter. To accomplish this task, we first attempt to reject this value by making it the subject of the null hypothesis. If we cannot reject the null hypothesis, then we must admit that the hypothesized value is indeed a candidate for the true value of the parameter in question.

2.9.2 One-Tailed Test

As an example of developing the probability statement associated with equation 2.9-2, assume that we wish to test the hypothesis that the mean of a population is μ_0, versus the alternate hypothesis that the population mean is greater than μ_0 (assume that the standard deviation σ_X is known):

$$H_0:\mu_X=\mu_0$$

versus

$$H_1:\mu_X>\mu_0.$$

This test is referred to as a one-tailed test because the alternate hypothesis only allows for a mean greater than that indicated by H_0.

A random variable is needed whereby we may build a probability statement around the type

I error. Assume that data in the form of a random sample $X_1, X_2, ..., X_n$ from the population exist; a natural random variable for this purpose would be the estimator of the mean $\bar{X}$. As $\bar{X}$ is an estimator of μ_X, and as the alternate hypothesis presupposes that μ_X is large, one would reject H_0 if a value of $\bar{X}$, as determined from a random sample, were significantly larger than μ_0. From the central limit theorem, assume that $\bar{X}$ is approximately normally distributed with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$. Equation 2.9–2 can be represented in terms of the statistic $\bar{X}$ as

$$P(\bar{X} > a | H_0) = \alpha . \qquad (2.9\text{–}3)$$

Although the distribution of $\bar{X}$ is known, a statistic which will allow us to incorporate the null hypothesis that $\mu_X = \mu_0$ is needed. A statistic meeting this requirement and for which values of all the parameters can be supplied is $(\bar{X} - \mu_X)/(\sigma_X/\sqrt{n})$. With this test statistic, equation 2.9–3 can be restated as

$$P\left(\frac{\bar{X} - \mu_x}{\sigma_X/\sqrt{n}} > c \,\middle|\, \mu_X = \mu_0\right) = P\left(\frac{\bar{X} - \mu_0}{\sigma_X/\sqrt{n}} > N_\alpha(0,1)\right) = \alpha$$

$$(2.9\text{–}4)$$

where μ_0 is used in place of μ_X to satisfy the null hypothesis. Note that under the null hypothesis $(\bar{X} - \mu_0)/(\sigma_X/\sqrt{n})$ is a normal random variable with mean zero and variance unity; thus, $N_\alpha(0,1)$ becomes the lower limit c of the critical region for this test.

All possible values of the test statistic greater than $N_\alpha(0,1)$, where α is the level of significance of the test, constitute the critical region where H_0 would be rejected. In other words, $N_\alpha(0,1)$ is the critical value, corresponding to the limit c in equation 2.9–2, which determines whether we accept or reject the null hypothesis. If a value of the test statistic is greater than $N_\alpha(0,1)$, we would reject H_0 at the α significance level. If the value were less, then we would be forced to accept the null hypothesis for fear of making a type I error.

As a sample application of this procedure, consider the data used to construct the confidence interval at the end of section 2.8: $\sigma_X = 0.3$, $\bar{x} = 2.6$ and $n = 36$. We are told that the

population mean is really zero, a statement that seems rather dubious to us as we believe it to be some positive real number. We set the null hypothesis that μ_X is indeed zero, $H_0{:}\mu_X = 0$, and hope that we can confidently disallow it. Our alternate hypothesis consists of our own belief, $H_1{:}\mu_X > 0$. As we wish to be very sure that we do not commit a type I error, we set the level of significance of our test at $\alpha = 0.025$. We determine the critical value of our test statistic from table 2.10–1: $N_\alpha(0,1) = 1.96$. We evaluate the test statistic under the assumption of the null hypothesis: $\bar{x}/(\sigma_X/\sqrt{n}) = 52$. Because this value of the test statistic is considerably larger than the critical value, we reject H_0 at the 0.025 significance level, realizing that, although we may have committed a type I error, it is highly unlikely.

2.9.3 Two-Tailed Test

Suppose that σ_X^2 is unknown, but we are given a random sample $X_1, X_2, ..., X_n$ from a normal population. We wish to test the hypothesis

$$H_0{:}\mu_X = \mu_0$$

versus

$$H_1{:}\mu_X \neq \mu_0$$

at a significance level α. This is referred to as a two-tailed test: We reject H_0 if a measure of μ_X is either significantly greater or less than μ_0.

To construct this test, recall the statistic from equation 2.6–22:

$$\frac{(\bar{X} - \mu_X)^2}{S_X^2/n} \sim F(1, n-1) \qquad (2.9\text{–}5)$$

which is the $F(1,n-1)$ random variable. This statistic fulfills our requirement for a test statistic; it can be used to satisfy the null hypothesis, and the remaining statistics or parameters are either known or can be evaluated from a random sample. Now consider the probability of a type I error:

$$P(\text{reject } H_0 | H_0 \text{ true})$$

$$= P(\bar{X} < a | H_0) + P(\bar{X} > b | H_0)$$

$$= \alpha , \qquad (2.9\text{–}6)$$

where two critical values are necessary as it is possible to reject the null hypothesis if a value of $\bar{X}$ is either larger or smaller than μ_0. In terms of the test statistic, under the condition that the null hypothesis holds, we see that

$$P\left(\left|\frac{\bar{X}-\mu_0}{S_X/\sqrt{n}}\right|^2 > F_\alpha(1,\ n-1)\right)$$

$$=P\left(\frac{\bar{X}-\mu_0}{S_X/\sqrt{n}} < -\sqrt{F_\alpha(1,n-1)}\right.$$

$$\text{or } \left.\frac{\bar{X}-\mu_0}{S_X/\sqrt{n}} > \sqrt{F_\alpha(1,n-1)}\right)$$

$$=P\left(\frac{\bar{X}-\mu_0}{S_X/\sqrt{n}} < -\sqrt{F_\alpha(1,n-1)}\right)$$

$$+P\left(\frac{\bar{X}-\mu_0}{S_X/\sqrt{n}} > \sqrt{F_\alpha(1,n-1)}\right)$$

$$=\alpha \qquad\qquad (2.9\text{-}7)$$

which is equivalent to equation 2.9–6.

To complete the test, we need only to evaluate the test statistic with a random sample. If $(\bar{x}-\mu_0)^2/(s_X^2/n)$ be greater than $F_\alpha(1,n-1)$, we would reject the null hypothesis at the α significance level.

2.9.4 Type II Error

A test statistic is commonly selected for its ability to determine the probability of committing a type II error, as well as a type I error. A type II error is committed by accepting the null hypothesis when the alternate is true. By calculating the probability that the test statistic does not fall in the critical region, given that Θ takes on any value other than that assumed under the null hypothesis, the probability β of committing this error can be evaluated. Thus, for tests indicated previously, β is a continuous function of possible values of the population parameter Θ, other than the value b assumed under the null hypothesis. For a critical region corresponding to a given α, a good test statistic

should produce small values of β for hypothetical values of Θ rather distant from b. However, β should increase sharply in value as possible values of Θ approach b and obtain a value as close to one as feasible in the immediate vicinity of b. Means are available for determining test statistics which, for certain tests, excel at the above characteristics, but a presentation of these methods is beyond the scope of this course. In most cases, a statistic which contains an estimator of the population parameter being tested and for which all other parameters are either known, or estimators of said parameters are contained in the statistic, will suffice as a test statistic; however, it may not be the best test statistic.

Note that if α, the probability of committing a type I error, were made extremely small, then the null hypothesis would almost always be accepted. At first glance, one would assume that something was amiss in the hypothesis testing procedure, as it is apparently possible to bias the test by selecting an extreme value for α. However, when the value of α is decreased, the probability of committing a type II error, $\beta(\Theta)$, is increased for all values Θ. Thus, an investigator who seeks to avoid committing a type I error by intentionally selecting a small value for α runs an increased risk of committing a type II error, which is equally as damaging. If need be, a plot of $\beta(\Theta)$ can be made for various hypothetical values of α and Θ; this can often be a rather complicated task. A rule-of-thumb value for α is 0.05, which appears to serve hypothesis test users well in most cases.

2.9.5 Summary of Method

To summarize, the steps for testing a hypothesis concerning a population parameter Θ are:
1. Define the null hypothesis $H_0{:}\Theta=\Theta_0$.
2. Decide upon the nature of the test; that is, $H_1{:}\Theta<\Theta_0$, $H_1{:}\Theta>\Theta_0$ or $H_1{:}\Theta \neq \Theta_0$.
3. Choose a level of significance α.
4. Select an appropriate test statistic and establish the critical region.
5. Compute the value of the statistic from a random sample of size n.
6. Draw conclusion of test: reject H_0 if the statistic has a value in the critical region; otherwise accept H_0.

Problem 2.9-1

a. Set up problem 2.6-1 as a hypothesis test (do not complete the test).

b. Given two random samples from independent normal populations with the following sample statistics:

Statistic	Random sample 1	Random sample 2
n	25	16
$\bar{x}$	82	78
s_x	8	7

test the following hypotheses at a significance level of $\alpha = 0.05$:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1 .$$

c. An outside source informs you that the stockpiled ore of problem 2.8-1 actually only assays an average of 9.8 grams per metric ton. Can you refute this claim at a significance level of 0.05? (Construct a hypothesis test for this purpose.)

2.10 Tables of Probability Distributions

[All tables modified from Walpole and Myers (1972), with permission from the publisher]

Table 2.10-1

Areas, $1-\alpha$, Under the Normal Curve

$N_\alpha(0,1)$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0 09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0 0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0 0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0 0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0011	0.0010
-2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0 0016	0.0015	0 0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0 0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0 0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0 0078	0.0075	0.0073	0 0071	0.0069	0.0068	0.0066	0.0064
-2.3	0 0107	0 0104	0.0102	0 0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0 0136	0 0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0 0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0 0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0352	0.0344	0 0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0 0427	0 0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0 0526	0.0516	0.0505	0.0495	0.0485	0.0475	0 0465	0.0455
-1.5	0.0668	0 0655	0 0643	0.0630	0.0618	0 0606	0.0594	0.0582	0.0571	0 0559
-1.4	0.0808	0.0793	0.0778	0.0764	0 0749	0.0735	0.0722	0.0708	0.0694	0.0681
-1.3	0 0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0 1093	0.1075	0.1056	0.1038	0 1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0 1170
-1.0	0.1587	0.1562	0 1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0 1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0 2090	0.2061	0.2033	0.2005	0 1977	0.1949	0 1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0 2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0 3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0 4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0 4286	0.4247
-0.0	0 5000	0.4960	0 4920	0.4880	0.4840	0.4801	0.4761	0 4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0 5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0 5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0 6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0 4	0.6554	0 6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0 7123	0.7157	0 7190	0.7224
0 6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0 7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0 7881	0.7910	0 7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0 9	0.8159	0.8186	0 8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0 8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0 8729	0.8749	0.8770	0.8790	0 8810	0.8830
1 2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0 9147	0.9162	0 9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0 9370	0.9382	0.9394	0 9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0 9474	0.9484	0.9495	0 9505	0 9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0 9582	0 9591	0.9599	0.9608	0 9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0 9671	0.9678	0.9686	0 9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2 1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2 2	0 9861	0.9864	0 9868	0.9871	0.9875	0.9878	0 9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0 9906	0.9909	0.9911	0 9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0 9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0 9969	0.9970	0.9971	0.9972	0.9973	0.9974
2 8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3 0	0.9987	0.9987	0.9987	0.9988	0.9988	0 9989	0.9989	0.9989	0.9990	0.9990
3 1	0 9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0 9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0 9997	0.9997	0.9998

Table 2.10-2

Critical Values of the Chi-Square Distribution

ν	α							
	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	3.841	5.024	6.635	7.879
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

2.10   Tables of Probability Distributions—Continued

Table 2.10-3

$$F_\alpha(\nu_1, \nu_2)$$
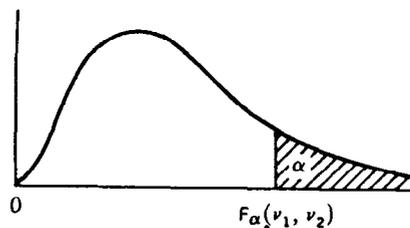
Critical Values of the F Distribution

$$F_{0.05}(\nu_1, \nu_2)$$

ν_2	ν_1								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Reproduced from Table 18 of *Biometrika Tables for Statisticians*, Vol. I, by permission of E. S. Pearson and the Biometrika Trustees.

Table 2.10-3 continued

$$F_{0.05}(\nu_1, \nu_2)$$

ν_2	ν_1									
	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

2.11   Appendices

2.11.1   Correlation of Two Linearly Related Random Variables

Consider a linear relationship between two random variables X and Y such that

$$Y = a + bX + \varepsilon \qquad (2.11-1)$$

where ε represents a zero-mean random error (independent of X). Then

$$\sigma_Y^2 = b^2 \sigma_X^2 + \sigma_\varepsilon^2 \qquad (2.11-2)$$

because $\sigma_{X\varepsilon} = E[(X - \mu_X)\varepsilon] = 0$. By direct calculation of σ_{XY} from equation 2.11-1, one obtains

$$\begin{aligned}
\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y)] \\
&= E[(X - \mu_X)(b(X - \mu_X) + \varepsilon)] \\
&= b\sigma_X^2 \quad .
\end{aligned} \qquad (2.11-3)$$

Upon squaring both sides of equation 2.11-3 and dividing by $\sigma_X^2 \sigma_Y^2$, one obtains

$$\rho_{XY}^2 = b^2 \frac{\sigma_X^2}{\sigma_Y^2} \qquad (2.11-4)$$

which, from equation 2.11-2, can be put in the form

$$\rho_{XY}^2 = 1 - \frac{\sigma_\varepsilon^2}{\sigma_Y^2} \quad . \qquad (2.11-5)$$

Again from equation 2.11-2, it is seen that

$$b^2 \sigma_X^2 = \sigma_Y^2 - \sigma_\varepsilon^2 \qquad (2.11-6)$$

and as $b^2 \sigma_X^2$ is a nonnegative quantity, σ_Y^2 must be greater than or equal to σ_ε^2. This shows that, for a linear relationship, ρ_{XY}^2 is either less than unity or equal to one if σ_ε^2 is equal to zero.

2.11.2   Expected Value of Variance Estimator

The sample statistic S_X^{*2} is defined as

$$S_X^{*2} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \quad . \qquad (2.11-7)$$

By taking expected values of both sides, one sees that

$$\begin{aligned}
E\left[S_X^{*2}\right] &= \frac{1}{n} E\left[\sum_{i=1}^{n} (X_i - \bar{X})^2\right] \\
&= \frac{1}{n} E\left[\sum_{i=1}^{n} (X_i - \mu_X - (\bar{X} - \mu_X))^2\right] \\
&= \frac{1}{n} \left[\sum_{i=1}^{n} E[(X_i - \mu_X)^2] - E[n(\bar{X} - \mu_X)^2]\right]
\end{aligned} \qquad (2.11-8)$$

where use is made of the fact that $\sum_{i=1}^{n}(X_i - \mu_X) = n(\bar{X} - \mu_X)$. Now the second expected value in equation 2.11-8 becomes

$$\begin{aligned}
E[n(\bar{X} - \mu_X)^2] &= \frac{1}{n} E\left[\sum_{i=1}^{n}(X_i - \mu_X) \sum_{j=1}^{n}(X_j - \mu_X)\right] \\
&= \frac{1}{n} E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - \mu_X)(X_j - \mu_X)\right] \\
&= \frac{1}{n} \sum_{i=1}^{n}\sum_{j=1}^{n} \text{Cov}[X_i, X_j] \quad .
\end{aligned} \qquad (2.11-9)$$

As X_i and X_j are randomly selected and therefore independent,

$$\text{Cov}[X_i, X_j] = \begin{cases} \sigma_X^2 & i = j \\ 0 & i \neq j \end{cases} \quad . \qquad (2.11-10)$$

Thus, equation 2.11-9 becomes

$$E[n(\bar{X} - \mu_X)^2] = \sigma_X^2 \qquad (2.11-11)$$

which allows us to write equation 2.11-8 as

$$\begin{aligned}
E\left[S_X^{*2}\right] &= \frac{1}{n}\left[\left(\sum_{i=1}^{n}\sigma_X^2\right) - \sigma_X^2\right] \\
&= \frac{n-1}{n}\sigma_X^2
\end{aligned} \qquad (2.11-12)$$

and demonstrates the desired result. Note that equation 2.11-11 also demonstrates that

$$\sigma_{\bar{X}}^2 = \sigma_X^2/n \qquad\qquad (2.11-13)$$

as $\sigma_{\bar{X}}^2 = E[(\bar{X} - \mu_X)^2]$.

References Cited

Nutter, Lawrence, 1973, Hydrogeology of the carbonate rocks, Fredrick and Hagerstown Valleys, Maryland: Maryland Geological Survey Report of Investigations No. 19, 70 p.

Siddiqui, S.H., 1969, Hydrogeologic factors influencing well yields and aquifer hydraulic properties of folded and faulted carbonate rocks in Central Pennsylvania: State College, Pennsylvania State University, unpublished Ph.D. thesis, 502 p.

Walpole, R.E. and Myers, R.H., 1972, Probability and statistics for engineers and scientists: New York, Macmillan, 506 p.

Additional Reading

Benjamin, J.R., and Cornell, A., 1970, Probability, statistics, and decision for civil engineers: New York, McGraw-Hill, 684 p.

3   Regression Solution of Modeling Problems

3.1   Introduction and Background

Ground-water flow models are members of a class of models known as mathematical models, in which the physical model of the flow system is replaced by mathematical expressions containing mathematical variables, parameters, and constants (Krumbein and Graybill, 1965, p. 15). Mathematical models always involve simplification of the actual (true) physical system. Krumbein and Graybill (1965, p. 15) argued that mathematical models can be classified into several types, including deterministic models, statistical models, and stochastic-process models.

A deterministic model is one in which the dependent variable(s) can be exactly computed from an expression involving independent variables, parameters, and constants. Note that deterministic models do not have to be physically based but may instead be completely empirical. The classical and inverse flow models discussed in section 1 are of the physically based deterministic type. In contrast, a statistical model is a deterministic model that has one or more random components added. These random components frequently involve measurement or other errors, but may involve separate sources of random variability as well. Incorporation of the errors in both observed heads and estimated parameters discussed in section 1 converts the deterministic flow model into a statistical model.

The term "stochastic model" can be considered to be synonymous with the term "statistical model" (Krumbein and Graybill, 1965, p. 19). A stochastic-process model may consider random effects such as those contained in the statistical model but in addition has a stochastic process built into it. Generation of a spatially varying permeability field in an aquifer has been considered to be a stochastic process by Bakr and others (1978), Gutjahr and others (1978), and Smith and Freeze (1979a, 1979b). Recently, this type of process has been incorporated into a parameter estimation scheme for a steady-state ground-water flow model

(Kitanidis and Vomvoris, 1983). Stochastic-process models are not considered further here.

3.1.1   Assumed Model Structure

Consider an experiment where two variables, ξ and Y, are measured repeatedly. The independent variable, ξ, is considered to be a precisely defined quantity, whereas the dependent variable, Y, whose values depend upon values of the independent variable, contains some error resulting from the experimental process. A scatter diagram of the data might appear as in figure 3.1-1.

From the scatter diagram or from physical considerations, the experimentalist may decide that an appropriate model equation for the data is

$$Y_i = \beta_1 + \beta_2 \xi_i + \epsilon_i \qquad (3.1-1)$$

where β_1 and β_2 are the intercept and slope of the equation for a straight line, subscript i represents the ith observation of (ξ, Y), and ϵ_i is the true error in Y for observation i. The quantity $\beta_1 + \beta_2 \xi$ is the deterministic part of the equation (the computed value of the dependent variable), and, because ϵ_i is the true error, parameters β_1 and β_2 are the true parameters representing the deterministic part of the model response. True error ϵ_i, often called a disturbance, is a random variable and, thus, represents the stochastic part of the model response. Note that if the model is correct and no other source of bias in ϵ_i exists, $E(\epsilon_i) = 0$.

Equation 3.1-1 is linear in parameters β_1 and β_2. Another example of a model equation that is linear in the parameters is

$$Y_i = \beta_1 + \beta_2 \xi_i + \beta_3 \xi_i^2 + \epsilon_i \qquad (3.1-2)$$

which is the equation for a second degree polynomial. In equation 3.1-2 there is still only one independent variable, ξ, although the equation has two terms containing ξ. An alternate form for equation 3.1-2 is

$$Y_i = X_{i1}\beta_1 + X_{i2}\beta_2 + X_{i3}\beta_3 + \epsilon_i \quad , \qquad (3.1-3)$$

where

$$\left. \begin{array}{l} X_1 = 1 \\ X_2 = \xi \\ X_3 = \xi^2 \end{array} \right\} \qquad (3.1-4)$$
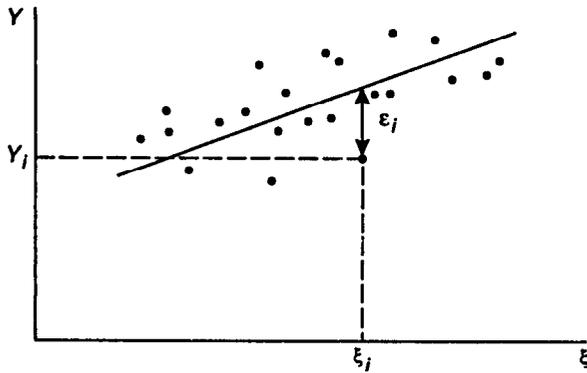
Figure 3.3-1

In general, any equation that is linear in parameters β_1, β_2, . . .,β_p, where there are p parameters for the system, can be written in the form

$$Y_i = X_{i1}\beta_1 + X_{i2}\beta_2 + \ldots + X_{ip}\beta_p + \epsilon_i \quad (3.1\text{-}5)$$

where

$$X_{ij} = X_{ij}(\xi_{i1}, \xi_{i2}, \ldots, \xi_{ik}) \quad (3.1\text{-}6)$$

is a function of k independent variables that multiplies the jth parameter and does not contain the parameters. Because

$$\frac{\partial}{\partial \beta_j}\left(X_{i1}\beta_1 + X_{i2}\beta_2 + \ldots + X_{ip}\beta_p\right) = X_{ij}, \quad (3.1\text{-}7)$$

the X terms are often called <u>sensitivity coefficients</u> or, simply, sensitivities. They indicate the change in the model response (the computed value of the dependent variable) at observation point i for a unit change in parameter β_j. Equation 3.1-5 can be written compactly in matrix form as

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\epsilon} \quad (3.1\text{-}8)$$

where

$$\underline{Y} = \begin{vmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{vmatrix} \quad (3.1\text{-}9)$$

$$\underline{X} = \begin{vmatrix} X_{11} & X_{12} & \ldots & X_{1p} \\ X_{21} & X_{22} & \ldots & X_{2p} \\ \vdots & & & \vdots \\ X_{n1} & \ldots & & X_{np} \end{vmatrix} \quad (3.1\text{-}10)$$

$$\underline{\beta} = \begin{vmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{vmatrix} \quad (3.1\text{-}11)$$

$$\underline{\epsilon} = \begin{vmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{vmatrix} \quad (3.1\text{-}12)$$

and n observations are assumed.

Often we are faced with models where the equation is no longer linear in the parameters, $\underline{\beta}$. For example, suppose that the model equation is

$$Y_i = \sqrt{\beta_1 + \beta_2}\, \tan\left(\frac{\sqrt{\beta_1 + \beta_2}}{\beta_1}\xi_i\right) + \epsilon_i \ . \quad (3.1\text{-}13)$$

Equation 3.1-13 cannot be reduced to the form of equation 3.1-5, and, thus, is not linear in the parameters. Equations of this type are written in the general vector form

$$\underline{Y} = f(\xi_1, \xi_2, \ldots, \xi_k; \beta_1, \beta_2, \ldots, \beta_p) + \underline{\epsilon} \quad (3.1\text{-}14)$$

or, in more compact form,

$$\underline{Y} = f(\underline{\xi}, \underline{\beta}) + \underline{\epsilon} \quad (3.1\text{-}15)$$

where f is an n-vector, each element f_i of which is a general function of the k independent variables, $\xi_{i\ell}$ ($\ell = 1, 2, \ldots, k$), and p parameters, β_j ($j = 1, 2, \ldots, p$). Equation 3.1-15 incorporates equation 3.1-8 because 3.1-8 is simply the special case where $f(\underline{\xi}, \underline{\beta}) = \underline{X}(\underline{\xi})\underline{\beta}$.

Some equations may be nonlinear in $\underline{\beta}$ but linear in some transformation of $\underline{\beta}$. For example, the model

$$Y_i = \beta_1(\beta_2)^{X_i}\epsilon_i \quad (3.1\text{-}16)$$

in which the error ϵ_i is multiplicative, is non-linear in β_1 and β_2. However, equation 3.1-16 may be written as

$$\log Y_i = \log \beta_1 + X_i \log \beta_2 + \log \epsilon_i \quad (3.1-17)$$

which is linear in $\log \beta_1$ and $\log \beta_2$ and has an additive error term. Thus, it is of the standard linear form. Equations such as 3.1-16 are frequently best utilized in their transformed, thus linearized, form. However, all model analyses (to be discussed further on) would probably be made in terms of the transformed variables, and this would have to be remembered when results were interpreted.

Types of models other than the linear and nonlinear ones discussed above also exist. Some types involve a complex model equation that cannot be solved explicitly for the dependent variable. In other cases the function $f(\underline{\xi},\underline{\beta})$, which is assumed to be a known function of $\underline{\xi}$ and $\underline{\beta}$, cannot be obtained and, therefore, must be replaced by a numerical formulation. However, the basic model structure of equation 3.1-15, where the error $\underline{\epsilon}$ in $\underline{Y}$ is assumed to be additive to a deterministic dependent variable vector, is always assumed. Additional complexities of the other types of models are handled by auxiliary equations appended to equation 3.1-15. The other models are introduced at appropriate places further on.

3.1.2 Least-Squares Estimation

Because the true parameter set $\underline{\beta}$ and true error set $\underline{\epsilon}$ are generally unknown, the true model equation 3.1-15 must be regarded as unknown, even though the form of the model is known (or, at least assumed). We do, however, have measurements to make up the independent variable set $\underline{\xi}$ and observation set $\underline{Y}$. We would like to use these measurements and the form of the model to obtain estimates of $\underline{\beta}$ and $\underline{\epsilon}$. The method explained in the following paragraphs is based on the idea that, if estimates of $\underline{\beta}$ and $\underline{\epsilon}$ can be found such that the error structure of the true model is duplicated as closely as possible, then the resulting model should, in some sense, be the best possible approximation of the true model.

Assume that all ϵ_i ($i=1,2,\ldots,n$) as random

variables have finite <u>common variance</u> σ^2 and that ϵ_i and ϵ_j, $i{\neq}j$, are uncorrelated. Then

$$\mathrm{Var}(\underline{\epsilon})=\underline{I}\sigma^2 \ . \quad (3.1-18)$$

The scalar variance σ^2 can be solved for by taking the trace of both sides of equation 3.1-18:

$$\mathrm{tr}[\mathrm{Var}(\underline{\epsilon})]=\mathrm{tr}(\underline{I})\sigma^2$$

or

$$\mathrm{tr}[E((\underline{\epsilon}-E(\underline{\epsilon}))(\underline{\epsilon}-E(\underline{\epsilon}))^T)]=n\sigma^2$$

or

$$E[(\underline{\epsilon}-E(\underline{\epsilon}))^T(\underline{\epsilon}-E(\underline{\epsilon}))]=n\sigma^2$$

from which

$$\sigma^2=\frac{E[(\underline{\epsilon}-E(\underline{\epsilon}))^T(\underline{\epsilon}-E(\underline{\epsilon}))]}{n} \ . \quad (3.1-19)$$

Ordinarily the assumptions would be made that the model being used is the correct one and that no other source of bias in $\underline{\epsilon}$ exists, so that $E(\underline{\epsilon})=\underline{0}$ and

$$\sigma^2=\frac{E(\underline{\epsilon}^T\underline{\epsilon})}{n} \ . \quad (3.1-20)$$

Equation 3.1-20 indicates that the sum of squared disturbances over all observations, averaged over many sets of observations, divided by n yields σ^2.

As indicated previously, the investigator only has available the data and the form of the model, so that $\underline{\epsilon}$, σ^2, and $\underline{\beta}$ must all be considered as unknowns. However, a good approximation of the true model would produce estimates of $\underline{\epsilon}$ that, for many observations, would yield a variance approaching σ^2. Let $\underline{b}$ be an estimator of $\underline{\beta}$. Then a linear model incorporating $\underline{b}$ is

$$\underline{Y}=\underline{X}\underline{b}+\underline{e} \quad (3.1-21)$$

where the vector $\underline{e}$ is an estimate of $\underline{\epsilon}$ called the <u>residual</u> vector. From equations 3.1-20 and 3.1-21, an estimate of σ^2 is

$$\tilde{\sigma}^2=\frac{\underline{e}^T\underline{e}}{n} \quad (3.1-22)$$

Most arbitrary parameter sets are expected to yield values of $\tilde{\sigma}^2$ that are larger than σ^2 because they would yield models that do not fit the data well. Of all possible parameter sets $\underline{b}$, the one that fits the data the best and at the same time minimizes $\tilde{\sigma}^2$ is the set that minimizes the sum of squares function $S(\underline{b})$,

$$S(\underline{b}) = \underline{e}^T \underline{e} \qquad (3.1\text{-}23)$$

with respect to $\underline{b}$. The process of finding estimates of σ^2 and $\underline{\beta}$ by minimizing $S(\underline{b})$ is termed least squares estimation. It is developed fully in sections 3.2 and 3.3.

Recall that equation 3.1-18 and, hence, equations 3.1-22 and 3.1-23 assume that the ϵ_i all come from the same distribution having variance σ^2 and that ϵ_i and ϵ_j, $i \neq j$, are uncorrelated. In some instances ϵ_i and ϵ_j have different variances, say σ_i^2 and σ_j^2, and they may even be correlated so that they have nonzero covariance σ_{ij}. In this case, equation 3.1-18 must be written in the more general form

$$\text{Var}(\underline{\epsilon}) = \underline{V}\sigma^2 \qquad (3.1\text{-}24)$$

where $\underline{V}\sigma^2$ is a symmetric, positive definite variance-covariance matrix defined as

$$\underline{V}\sigma^2 = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ & & \cdots & & \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_n^2 \end{bmatrix} \qquad (3.1\text{-}25)$$

In equation 3.1-24, σ^2 is no longer the common variance of all ϵ_i but is instead another type of common variance. Its exact meaning can be discerned as follows. Define $\underline{V}^{1/2}$ as the nonsingular symmetric matrix such that $\underline{V}^{1/2}\underline{V}^{1/2} = \underline{V}$. Then, from equation 3.1-24,

$$\underline{V}^{-1/2} \text{Var}(\underline{\epsilon})\underline{V}^{-1/2} = \underline{V}^{-1/2}\underline{V}\underline{V}^{-1/2}\sigma^2$$

or

$$\text{Var}(\underline{V}^{-1/2}\underline{\epsilon}) = \underline{I}\sigma^2 \qquad (3.1\text{-}26)$$

from which σ^2 is seen to be the constant or common variance of the transformed disturbances $\underline{V}^{-1/2}\underline{\epsilon}$. Equation 3.1-26 shows that these disturbances are uncorrelated.

The more general conditions represented by equation 3.1-24 may easily be incorporated into the least squares procedure. All developments through equation 3.1-20 are repeated using $\underline{V}^{-1/2}\underline{\epsilon}$ instead of $\underline{\epsilon}$. The result is

$$\sigma^2 = \frac{E[(\underline{V}^{-1/2}\underline{\epsilon})^T(\underline{V}^{-1/2}\underline{\epsilon})]}{n}$$

$$= \frac{E(\underline{\epsilon}^T\underline{V}^{-1}\underline{\epsilon})}{n} \qquad (3.1\text{-}27)$$

which suggests that

$$S(\underline{b}) = \underline{e}^T\underline{V}^{-1}\underline{e} \qquad (3.1\text{-}28)$$

be minimized instead of equation 3.1-23.

If equation 3.1-23 is minimized to find the parameter estimates when the more general error structure given by equation 3.1-24 is correct, then the incorrect error structure will be reflected in parameter estimates that are less precise than if equation 3.1-28 were used. The proper sum of squares function to minimize when equation 3.1-24 represents the correct error structure is equation 3.1-28.

Even more general cases can be postulated to yield $S(\underline{b})$ in the form

$$S(\underline{b}) = \underline{e}^T\underline{\omega}\underline{e} \qquad (3.1\text{-}29)$$

where $\underline{\omega}$ is a general symmetric positive definite weight matrix that subsumes $\underline{V}^{-1}$. To apply equation 3.1-29, $\underline{\omega}$ does not necessarily reflect the error structure of $\underline{\epsilon}$. Instead it may reflect the investigator's desire to emphasize (or de-emphasize) certain components of $S(\underline{b})$. Equation 3.1-29 is used in all developments to follow in which the general form is applicable.

As a final note, least-squares estimation should be viewed as more than simply a parameter estimation procedure. The development given in this section is intended to show that the procedure is an attempt to reproduce the true model structure: the variance, σ^2, the distribution of $\underline{\epsilon}$, and $\underline{\beta}$. Although it is possible to use least-squares estimation as just an algebraic process, making no assumptions

about structure, considerably more information can be gained by taking the more general approach to make model analysis, including analysis of assumptions initially made concerning model structure, an integral part of the regression process.

3.1.3   Inclusion of Prior Information

The model structure given by equation 3.1–15 is general. Nothing is implied about the nature of $\underline{Y}$ except that it is a dependent variable vector, in error by the amount $\underline{\epsilon}$. A very general interpretation of equation 3.1–15 is to assume that $\underline{Y}$, f, and $\underline{\epsilon}$ are each composed of two partitions, one giving sample information and one giving prior information on parameters. This viewpoint amounts to an expansion of the original formulation given in the previous section where only sample information was considered (Theil, 1963).

For example, suppose that an investigator collects data (ξ, Y) on a process for which the model equation is given by equation 3.1–2. However, suppose that he or she also has developed methods to collect some data directly on parameters β_1 and β_3 and suppose that these data can be represented by the equations

$$\left.\begin{aligned} P_1 &= a_{11}\beta_1 + u_1 \\ P_2 &= a_{21}\beta_1 + a_{23}\beta_3 + u_2 \end{aligned}\right\} \qquad (3.1-30)$$

where a_{ij} is a constant and u_i is a random error. If $a_{11} = 1$, then P_1 is a direct observation, subject to error u_1, of β_1. The entire set of equations representing the system, then, can be written

$$\left.\begin{aligned} Y_1 &= \beta_1 + \xi_1\beta_2 + \xi_1^2\beta_3 + \epsilon_1 \\ Y_2 &= \beta_1 + \xi_2\beta_2 + \xi_2^2\beta_3 + \epsilon_2 \\ &\quad\cdots \\ Y_n &= \beta_1 + \xi_n\beta_2 + \xi_n^2\beta_3 + \epsilon_n \\ P_1 &= a_{11}\beta_1 + u_1 \\ P_2 &= a_{21}\beta_1 + a_{23}\beta_3 + u_2 \end{aligned}\right\} \qquad (3.1-31)$$

If $\underline{Y}$, $\underline{X}$, and $\underline{\epsilon}$ are augmented to include the prior information, then equation 3.1–31 is of the form 3.1–8 where

$$\underline{Y} = \begin{vmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \\ P_1 \\ P_2 \end{vmatrix} \qquad \underline{\epsilon} = \begin{vmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdots \\ \epsilon_n \\ u_1 \\ u_2 \end{vmatrix}$$

$$\underline{X} = \begin{vmatrix} 1 & \xi_1 & \xi_1^2 \\ 1 & \xi_2 & \xi_2^2 \\ & \cdots & \\ 1 & \xi_n & \xi_n^2 \\ a_{11} & 0 & 0 \\ a_{21} & 0 & a_{23} \end{vmatrix} \quad .$$

Note that the number of observations is now the number of equations giving sample information (n) plus the number of equations giving prior information (2).

Although the equations giving prior information are often linear, they do not need to be. Hence, a general form of equation 3.1–15 to include the prior information may be assumed:

$$\underline{Y} = f(\underline{\xi}, \underline{\beta}) + \underline{\epsilon} \qquad (3.1-32)$$

where

$$\underline{Y} = \begin{bmatrix} \underline{Y_s} \\ \underline{Y_p} \end{bmatrix} \qquad f = \begin{bmatrix} f_s \\ f_p \end{bmatrix} \qquad \underline{\epsilon} = \begin{bmatrix} \underline{\epsilon_s} \\ \underline{\epsilon_p} \end{bmatrix} \qquad (3.1-33)$$

and subscripts s and p indicate partitions of the respective vectors pertaining to sample and prior information, respectively. Corresponding to these partitions, it is convenient to redefine n as the total number of observations, $n = n_s + n_p$, where n_s is the number of items (or equations) of sample information and n_p is the number of items (or equations) of prior information.

To apply the least squares procedure to equation 3.1–15 as augmented by the prior information, it is assumed for now that

$$\mathrm{Var}(\underline{\epsilon_s}) = \underline{V_s}\sigma^2 \qquad (3.1-34)$$

$$\mathrm{Var}(\underline{\epsilon_p}) = \underline{V_p}\sigma^2 \qquad (3.1-35)$$

$$\text{Cov}(\underline{\epsilon}_s, \ \underline{\epsilon}_p)=\underline{0} \qquad (3.1\text{--}36)$$

where $\underline{V}_s$ is symmetric, positive definite and of order n_s, and $\underline{V}_p$ is symmetric, positive definite and of order n_p. Equation 3.1–36 indicates that sample disturbances $\underline{\epsilon}_s$ and prior information disturbances $\underline{\epsilon}_p$ are not correlated with each other. With use of equations 3.1–34 through 3.1–36, equation 3.1–24 becomes

$$\text{Var}\begin{bmatrix}\underline{\epsilon}_s \\ \underline{\epsilon}_p\end{bmatrix}=\begin{bmatrix}\text{Var}(\underline{\epsilon}_s) & \underline{0} \\ \underline{0} & \text{Var}(\underline{\epsilon}_p)\end{bmatrix}$$

$$=\sigma^2\begin{bmatrix}\underline{V}_s & \underline{0} \\ \underline{0} & \underline{V}_p\end{bmatrix}$$

$$=\underline{V}\sigma^2 \qquad . \qquad (3.1\text{--}37)$$

With use of equation 3.1–37, 3.1–28 becomes

$$S(\underline{b})=\underline{e}^T\underline{\underline{V}}^{-1}\underline{e}$$

$$=\begin{bmatrix}\underline{e}_s^T, & \underline{e}_p^T\end{bmatrix}\begin{bmatrix}\underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{V}_p^{-1}\end{bmatrix}\begin{bmatrix}\underline{e}_s \\ \underline{e}_p\end{bmatrix}$$

$$=\underline{e}_s^T\underline{V}_s^{-1}\underline{e}_s+\underline{e}_p^T\underline{V}_p^{-1}\underline{e}_p \qquad (3.1\text{--}38)$$

where the residual vector is defined as

$$\underline{e}=\begin{bmatrix}\underline{e}_s \\ \underline{e}_p\end{bmatrix} \qquad (3.1\text{--}39)$$

and

$$\underline{V}^{-1}=\begin{bmatrix}\underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{V}_p^{-1}\end{bmatrix} . \qquad (3.1\text{--}40)$$

The least-squares procedure may be generalized even further by using equation 3.1–29 instead of equation 3.1–28 to define $S(\underline{b})$. In this case the weight matrix $\underline{\omega}$ is defined by

$$\underline{\omega}=\begin{bmatrix}\underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p\end{bmatrix} \qquad (3.1\text{--}41)$$

where $\underline{\omega}_s$ is a symmetric positive definite submatrix of order n_s that pertains to the sample information and $\underline{\omega}_p$ is a symmetric positive definite submatrix of order n_p that pertains to the prior information. Because equation 3.1–41 is of block diagonal form like equation 3.1–37, zero correlation of sample and prior information is again assumed. Thus, $S(\underline{b})$ may be written in the same form as equation 3.1–38, or as

$$S(\underline{b})=\underline{e}_s^T\underline{\omega}_s\underline{e}_s+\underline{e}_p^T\underline{\omega}_p\underline{e}_p \qquad . \qquad (3.1\text{--}42)$$

Problem 3.1–1

You are charged with a ground-water study in the vicinity of Lake Ohpupu (figure 1). Estimates of transmissivity and recharge for the confined aquifer surrounding the lake are necessary for the completion of your report. Taking advantage of the unusually colinear equipotential contours on the west side of the lake (constructed from an unbiased source, of course), you decide that estimates based on a uniform stream tube will suffice. Recharge to the aquifer is largely from precipitation and is uniform over the region. Assume that the boundary heads at the range front and the lake are imprecisely known; estimates of these parameters will also be necessary. Your project has limited funds to bore n_s holes along the stream tube and obtain measurements of head at n_s locations of distance, s, from the range front.

The steady-state flow equation for a stream tube is

$$\frac{d}{ds}\left(TD\frac{dh}{ds}\right)+WD=0 \qquad (1)$$

where

$T=$transmissivity (ft^2/d);
$W=$recharge (ft/d);
$D=$width of stream tube (ft);
$h=$hydraulic head (ft); and
$s=$distance along tube from the range front (ft).

The boundary conditions are taken to be

$$\left.\begin{array}{l}h=h_0 \text{ at } s=0 \\ h=h_b \text{ at } s=s_b\end{array}\right\} \qquad . \qquad (2)$$

POTENTIOMETRIC SKETCH MAP FOR THE VICINITY OF LAKE OHPUPU

BALDERDASH MOUNTAINS

$S=0$

$\rightarrow s$

$s=s_b$

LAKE OHPUPU

EXPLANATION

SKETCH EQUIPOTENTIAL LINE
SHOWING DIRECTION
OF EQUAL HEAD.

STREAMTUBE

N
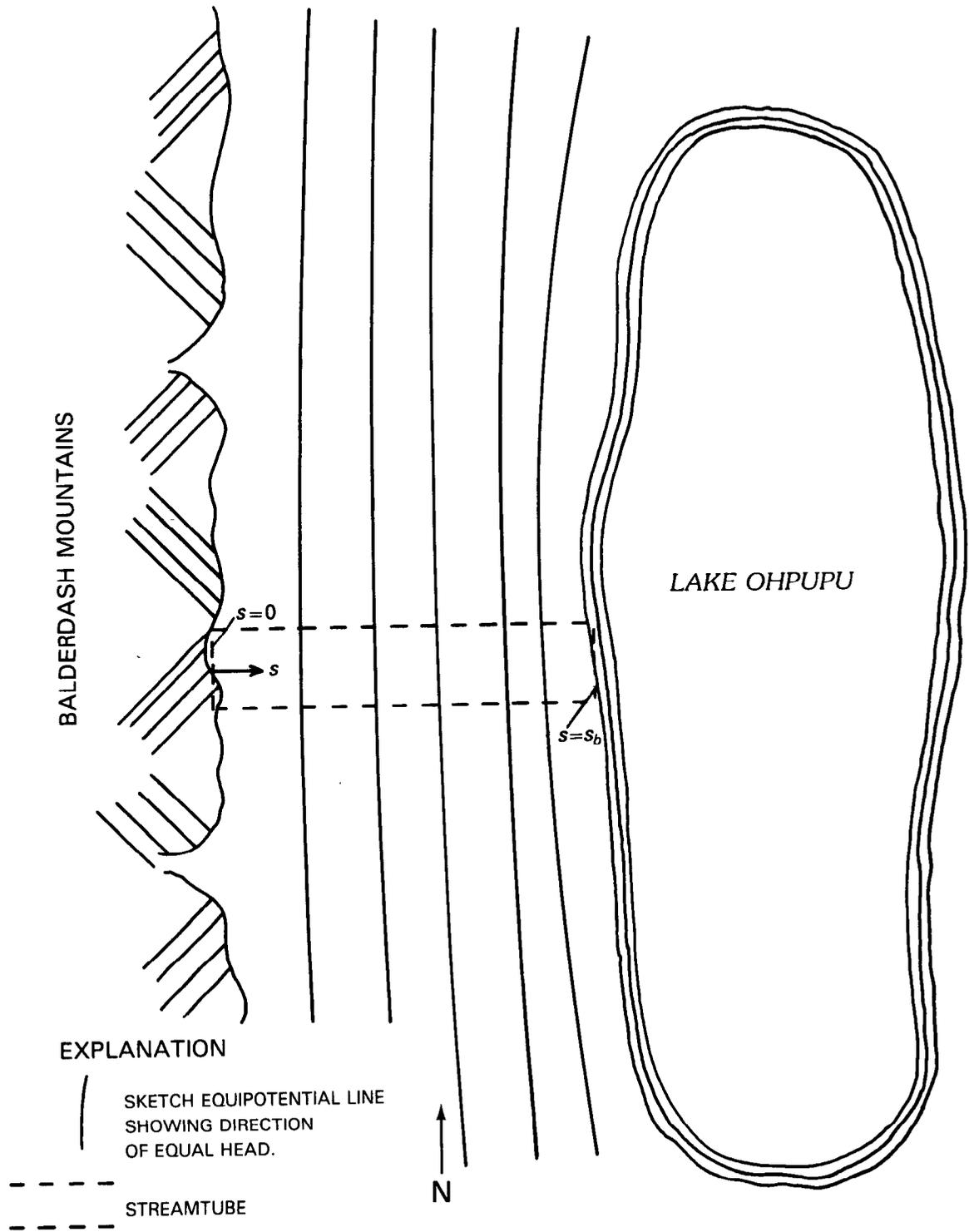
Figure 1

By integrating the flow equation twice with respect to s (assuming T, D, and W to be constant) and using the above boundary conditions, the solution for hydraulic head h along the stream tube is found to be:

$$h = \frac{W}{2T}(s_b - s)s + h_b \frac{s}{s_b} + h_o\left(\frac{s_b - s}{s_b}\right) . \qquad (3)$$

Let

$$\left.\begin{array}{lll} \beta_1 = h_o & \beta_3 = W/T & X_2 = \dfrac{s}{s_b} \\[2mm] \beta_2 = h_b & X_1 = \dfrac{s_b - s}{s_b} & X_3 = \dfrac{(s_b - s)s}{2} \end{array}\right\} \qquad (4)$$

and write the above solution (equation 3) to the flow equation using the definitions of β_i and X_i ($i=1,2,3$). Then write the system of n_s linear regression equations in the three unknown parameters using matrix notation and indicate the contents of each matrix. Identify dependent variable(s), independent variable(s), sensitivities, and parameters.

Let b_i be an estimate of β_i, and:

a. Assume that $\mathrm{Var}(\underline{\epsilon}) = \underline{I}\sigma^2$. Write $S(\underline{b})$ using the matrix form of the model equation 3.1-21 with the estimated parameters b_i. Write a few terms of $S(\underline{b})$ using algebraic notation.

b. Assume that a unique estimate of the variance of the error associated with every head observation is available and that these errors are uncorrelated. Indicate the contents of the resulting weight matrix $\underline{\omega} = \underline{V}^{-1}$. Write $S(\underline{b})$ using the matrix form of the model equation with parameters b_i. Write a few terms of $S(\underline{b})$ using weights ω_{jj} and algebraic notation.

c. Assume case a above, except that there is a prior estimate of h_b having a standard deviation of σ_{h_b}. Indicate the contents of the resulting weight matrix $\omega = V^{-1}$. Write $S(\underline{b})$ using the matrix form of the model equation with parameters b_i. Write a few terms of $S(\underline{b})$ using algebraic notation. Include the term involving the prior information.

3.2 Regression When the Model is Linear

3.2.1 Derivation of Solution

The linear model assumed is

$$Y = X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p + \epsilon \qquad (3.2\text{-}1)$$

where the X_i are not functions of the parameters. If n observations are used, then an equation of the form of equation 3.2-1 is written for each observation, so that the system can be written in matrix form as

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\epsilon} . \qquad (3.2\text{-}2)$$

To find estimates of $\underline{\beta}$ and $\underline{\epsilon}$, the weighted error sum of squares $S(\underline{b})$,

$$\begin{aligned} S(\underline{b}) &= \underline{e}^T\underline{\omega}\underline{e} \\ &= (\underline{Y} - \underline{X}\underline{b})^T\underline{\omega}(\underline{Y} - \underline{X}\underline{b}) \end{aligned} \qquad (3.2\text{-}3)$$

is minimized with respect to $\underline{b}$.

To minimize $S(\underline{b})$ with respect to $\underline{b}$ means to take the derivative of $S(\underline{b})$ with respect to each element of $\underline{b}$, b_j ($j=1,2,\ldots,p$), and set the results to zero, or

$$\frac{\partial}{\partial b_j} S(\underline{b})\big|_{\underline{b}=\hat{\underline{b}}} = \frac{\partial}{\partial b_j}[(\underline{Y}-\underline{X}\underline{b})^T\underline{\omega}(\underline{Y}-\underline{X}\underline{b})]|_{\underline{b}=\hat{\underline{b}}}$$
$$= 0, j=1,2,\ldots,p, \qquad (3.2\text{-}4)$$

where $|_{\underline{b}=\hat{\underline{b}}}$ signifies that $\hat{\underline{b}}$ is the set of parameters that causes the derivatives of $S(\underline{b})$ to be zero. By employing the rule of differentiating a product and noting that $\underline{\omega}$ is independent of $\underline{b}$, it can be seen that

$$\frac{\partial}{\partial b_j}[(\underline{Y}-\underline{X}\underline{b})^T\underline{\omega}(\underline{Y}-\underline{X}\underline{b})] = [\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b})^T]\underline{\omega}(\underline{Y}-\underline{X}\underline{b})$$

$$+ (\underline{Y}-\underline{X}\underline{b})^T\underline{\omega}\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b}). \qquad (3.2\text{-}5)$$

To evaluate $\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b})^T$, note that taking the derivative of a vector, $\underline{Y}-\underline{X}\underline{b}$, with respect to a scalar, b_j, means taking the derivative of each

entry of the vector and, thus, yields a new vector. Therefore, because observations $\underline{Y}$ are independent of $\underline{b}$,

$$\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b})^T = \frac{\partial}{\partial b_j}\left\{\begin{bmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{bmatrix} - [\underline{X}_1,\underline{X}_2,...,\underline{X}_p]\begin{bmatrix}b_1\\b_2\\\vdots\\b_p\end{bmatrix}\right\}^T$$

$$=-\underline{X}_j^T, \qquad (3.2\text{-}6)$$

where $\underline{X}_j$ is the jth column vector of matrix $\underline{X}$. Also, the first and second terms on the right side of equation 3.2-5 are equal because the transpose of a scalar is the scalar, and $\underline{\omega}$ is symmetric so that $\underline{\omega}=\underline{\omega}^T$. Hence, transposing the second term on the right side of equation 3.2-5 gives

$$(\underline{Y}-\underline{X}\underline{b})^T\underline{\omega}\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b})=[(\underline{Y}-\underline{X}\underline{b})^T\underline{\omega}\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b})]^T$$

$$=[\frac{\partial}{\partial b_j}(\underline{Y}-\underline{X}\underline{b})^T]\underline{\omega}(\underline{Y}-\underline{X}\underline{b}). \qquad (3.2\text{-}7)$$

The combination of equations 3.2-4 through 3.2-7 yields

$$-2\underline{X}_j^T\underline{\omega}(\underline{Y}-\underline{X}\hat{\underline{b}})=0 \qquad (3.2\text{-}8)$$

or

$$\underline{X}_j^T\underline{\omega}\underline{X}\hat{\underline{b}}=\underline{X}_j^T\underline{\omega}\underline{Y}, \; j=1,2,...,p. \qquad (3.2\text{-}9)$$

The system of equations implied by equation 3.2-9 can be written as

$$\begin{bmatrix}\underline{X}_1^T\\\underline{X}_2^T\\\vdots\\\underline{X}_p^T\end{bmatrix}\underline{\omega}\underline{X}\hat{\underline{b}}=\begin{bmatrix}\underline{X}_1^T\\\underline{X}_2^T\\\vdots\\\underline{X}_p^T\end{bmatrix}\underline{\omega}\underline{Y}. \qquad (3.2\text{-}10)$$

Each vector $\underline{X}_j^T$ in equation 3.2-10 is a row vector, so that, by definition,

$$\begin{bmatrix}\underline{X}_1^T\\\underline{X}_2^T\\\vdots\\\underline{X}_p^T\end{bmatrix}=\underline{X}^T \qquad (3.2\text{-}11)$$

and equation 3.2-10 becomes

$$\underline{X}^T\underline{\omega}\underline{X}\hat{\underline{b}}=\underline{X}^T\underline{\omega}\underline{Y}. \qquad (3.2\text{-}12)$$

The set of equations symbolized in matrix form by equation 3.2-12 are called the <u>normal equations</u>, and parameters $\hat{\underline{b}}$ are called the <u>estimates</u> of $\underline{\beta}$. The estimates are found from

$$\hat{\underline{b}}=(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}\underline{Y}. \qquad (3.2\text{-}13)$$

Students not comfortable with the preceding development should read Draper and Smith (1981, p. 5-17, 70-80, 85-87). This material covers fitting a straight line without and with matrix nomenclature, and then extends the results to the general linear situation. Weighted least squares (where $\underline{\omega}\neq\underline{I}$) is covered in Draper and Smith (1981, p. 108-116).

Elements of $\underline{X}$ are often of vastly differing magnitudes. Thus, when working with a calculator or computer, round-off error can cause serious errors to develop when solving equation 3.2-13. It is often useful to scale equation 3.2-12 with respect to a matrix $\underline{C}$, which is a diagonal matrix defined as follows: Let $\underline{X}^T\underline{\omega}\underline{X}=\underline{A}$. Then $\underline{C}=\text{diag}\{1/A_{11}^{\frac{1}{2}},1/A_{22}^{\frac{1}{2}},...,1/A_{pp}^{\frac{1}{2}}\}$, where A_{ii} is a diagonal entry of $\underline{A}$. Thus, equation 3.2-12 can be transformed to become

$$\underline{C}^T\underline{X}^T\underline{\omega}\underline{X}\underline{C}\underline{C}^{-1}\hat{\underline{b}}=\underline{C}^T\underline{X}^T\underline{\omega}\underline{Y} \qquad (3.2\text{-}14)$$

or

$$\underline{S}^T\underline{\omega}\underline{S}\hat{\underline{a}}=\underline{S}^T\underline{\omega}\underline{Y} \qquad (3.2\text{-}15)$$

where

$$\underline{S}=\underline{X}\underline{C} \qquad (3.2\text{-}16)$$

$$\hat{\underline{a}}=\underline{C}^{-1}\hat{\underline{b}}. \qquad (3.2\text{-}17)$$

The effect of the scaling is to preserve the symmetry of $\underline{X}^T\underline{\omega}\underline{X}$ while at the same time to produce a matrix having all diagonal entries equal to unity. Thus, variability from entry to entry of the $\underline{S}^T\underline{\omega}\underline{S}$ matrix is usually reduced considerably over that of $\underline{X}^T\underline{\omega}\underline{X}$.

It is also sometimes useful to transform equation 3.2-2 and equation 3.2-12 to incremental form. By definition

$$f(\underline{\xi}, \underline{b}) = \underline{X}\underline{b} \quad . \tag{3.2-18}$$

Then, subtracting equation 3.2-18 from equation 3.2-2 results in

$$\underline{Y} - f(\underline{\xi}, \underline{b}) = \underline{X}(\underline{\beta} - \underline{b}) + \underline{\epsilon} \tag{3.2-19}$$

which is an <u>incremental linear model</u>. To obtain the analog to equation 3.2-12, premultiply equation 3.2-18 by $\underline{X}^T\underline{\omega}$ and subtract the result from equation 3.2-12 to obtain

$$\underline{X}^T\underline{\omega}\underline{X}(\hat{\underline{b}} - \underline{b}) = \underline{X}^T\underline{\omega}(\underline{Y} - f(\underline{\xi}, \underline{b})) \quad . \tag{3.2-20}$$

Equation 3.2-20 can be transformed to obtain a result analogous to equation 3.2-15:

$$\underline{S}^T\underline{\omega}\underline{S}\hat{\underline{\delta}} = \underline{S}^T\underline{\omega}(\underline{Y} - f(\underline{\xi}, \underline{b})) \tag{3.2-21}$$

where

$$\hat{\underline{\delta}} = \underline{C}^{-1}(\hat{\underline{b}} - \underline{b}) \quad . \tag{3.2-22}$$

If $\hat{\underline{b}}$ as calculated initially using equations 3.2-15 and 3.2-17 is in error because of round-off, then $\underline{b}$ can be used to calculate $f(\underline{\xi}, \underline{b})$ which then can be substituted into equation 3.2-21 to calculate $\hat{\underline{\delta}}$. By using equation 3.2-22, a new improved estimate of $\underline{b}$ can be obtained. Writing equation 3.2-2 in incremental form also provides a basis for procedures, involving statistical analysis of the model, that apply for both linear and (with restrictions) nonlinear models. These are discussed later on.

3.2.2 Solution Algorithm

Sequential steps to follow are:
1. Form $\underline{X}^T\underline{\omega}\underline{X}$ and $\underline{X}^T\underline{\omega}\underline{Y}$.
2. Transform equation 3.2-12 to 3.2-15.
3. Solve equation 3.2-15 (or 3.2-21) for $\hat{\underline{a}}$.
4. Solve equation 3.2-17 for $\hat{\underline{b}}$.

Problem 3.2-1

a. By using $\underline{\omega}$ from case c, problem 3.1-1, write out the normal equations used to estimate parameters β explicitly in sum-of-product (algebraic) form. (Use equation 3.2-12.)

b. By using either data set 1 (table 1) or 2 (table 2), generate the least squares coefficient matrix $(\underline{X}^T\underline{\omega}\underline{X})$ and then compute its inverse. <u>Do not</u> round off any intermediate calculations or the final inverse. To aid in the calculations, table 3 gives the sums of products for the sample information from the two data sets. You must add the prior information to complete the sums of products.

c. Find the vector $\hat{\underline{b}}$. Do not round off the results.

Table 1.—Data set 1

s_j (ft)	X_{j1}	X_{j2}	X_{j3}	Observed head, Y_j (ft)
50	0.95	0.05	23,750	48.33
150	.85	.15	63,750	45.76
250	.75	.25	93,750	42.08
350	.65	.35	113,750	38.34
450	.55	.45	123,750	35.30
550	.45	.55	123,750	31.00
650	.35	.65	113,750	25.85
750	.25	.75	93,750	21.76
850	.15	.85	63,750	16.11
950	.05	.95	23,750	12.48

Assume $\sigma^2 = 0.25$ ft^2, $s_b = 1,000$ ft, and prior information as follows: $h_b = 11$ ft and $\sigma_{h_b} = 1.1$ ft.

Table 2.—Data set 2

s_j (ft)	X_{j1}	X_{j2}	X_{j3}	Observed head, Y_j (ft)
100	0.9	0.1	45,000	47.13
200	.8	.2	80,000	44.14
300	.7	.3	105,000	39.89
400	.6	.4	120,000	36.36
500	.5	.5	125,000	32.48
600	.4	.6	120,000	29.70
700	.3	.7	105,000	24.33
800	.2	.8	80,000	19.10
900	.1	.9	45,000	14.96

Assume $\sigma^2 = 0.25$ ft^2, $s_b = 1,000$ ft, and prior information as follows: $h_b = 9.5$ ft and $\sigma_{h_b} = 0.95$ ft.

Table 3.—Sums of products of sample information

	Data Set 1	Data Set 2
$\sum_j X_{j1} X_{j1}$	3.3250	2.8500
$\sum_j X_{j1} X_{j2}$	1.6750	1.6500
$\sum_j X_{j1} X_{j3}$	418,750	412,500
$\sum_j X_{j2} X_{j2}$	3.3250	2.8500
$\sum_j X_{j2} X_{j3}$	418,750	412,500
$\sum_j X_{j3} X_{j3}$	83,340,625,000	83,325,000,000
$\sum_j X_{j1} Y_j$	192.18350	168.2030
$\sum_j X_{j2} Y_j$	124.82650	119.8870
$\sum_j X_{j3} Y_j$	26,879,687.5	26,583,550

3.2.3 Singularity and Conditioning

Singularity of the least-squares coefficient matrix occurs whenever columns of the sensitivity matrix, X, are linearly dependent because this causes rows (or columns) of the coefficient matrix $X^T \omega X$ to be linearly dependent. Linear dependence in X may be stated as

$$Xc = 0 \qquad (3.2\text{-}23)$$

where not all components of the vector c of order p are zero. By premultiplying equation 3.2-23 by $X^T \omega$,

$$X^T \omega X c = 0 \qquad (3.2\text{-}24)$$

which shows that columns of $X^T \omega X$ (or rows since $X^T \omega X$ is symmetric) are linearly dependent. Note that transformation of $X^T \omega X$ to $S^T \omega S$ alters only the form of c.

Near-singularity, also referred to as ill-conditioning, occurs whenever the columns of X (or S) are almost linearly dependent. Often, this condition is indicated by a high degree of correlation among two or more parameter estimates. This correlation reflects the redundancy in the problem. As a result of ill-conditioning, computed parameters can be affected greatly by accumulation of round-off error generated by solving the normal equations. Also, computed variances of the parameters, which are proportional to the diagonal elements of $(X^T \omega X)^{-1}$, will be large.

A common form of ill-conditioning results if a column of X approaches zero so that $c = [0,0, \ldots, 1,0, \ldots, 0]^T$, where the one appears in the row corresponding to the zero column in X. This condition indicates that the model is insensitive to the parameter corresponding to the zero column in X and that the parameter should be eliminated from the model. The problem is readily detected by examining the X matrix. Another readily detected form of ill-conditioning results if two columns of X are nearly proportional, or

$$\alpha X_i \cong X_j \qquad (3.2\text{-}25)$$

so that $c = [0,0, \ldots, \alpha, 0, \ldots, -1, 0, \ldots, 0]^T$, where α appears in row i and -1 appears in row j of c. In this case

$$X_i^T \omega X_j \cong \alpha X_j^T \omega X_j$$

$$\cong \frac{1}{\alpha} X_j^T \omega X_j$$

so that

$$S_i^T \omega S_j = \frac{X_i^T \omega X_j}{\sqrt{(X_i^T \omega X_i)(X_j^T \omega X_j)}}$$

$$\cong \pm 1 \qquad . \qquad (3.2\text{-}26)$$

Thus, to detect this problem, one need only examine the $S^T \omega S$ matrix for an off-diagonal entry nearly equal to ± 1. This type of linear dependence indicates that parameters b_i, and b_j should be combined because the model can be written as

$$Y = X_1 b_1 + \ldots + X_i b_i + \ldots + X_j b_j + \ldots + X_p b_p$$

$$\cong X_1 b_1 + \ldots + X_i(b_i + \alpha b_j) + \ldots + X_p b_p$$

$$= X_1 b_1 + \ldots + X_i b_i^* + \ldots + X_{p-1} b_{p-1} \quad (3.2\text{-}27)$$

where b_i^* replaces $b_i + \alpha b_j$, and all subsequent variables are shifted by one so that the last variable number is $p-1$.

An excellent way to detect general ill-conditioned (or completely linearly dependent) problems is to orthogonalize the columns of the scaled sensitivity matrix, S (Draper and Smith, 1981, p. 275-278). If the columns are all linearly independent, then they can all be transformed

so as to be orthogonal to one another; that is, such that

$$Q^T Q = D \qquad (3.2\text{-}28)$$

where Q is the nonzero transformation of S, and D is a full-rank diagonal matrix. If linear dependence exists in S, then equation 3.2-28 is replaced by a similar diagonal form except that one or more diagonal entries will be zero. The technique is to successively transform columns such that each new column is orthogonal to all of the previously transformed columns. If column dependence exists, then eventually a column will be calculated that exists entirely of very small numbers (theoretically all zeros for a linearly dependent problem). This column, then, is almost (or completely) linearly dependent on one or more of the previous columns.

The transformation procedure is called Gram-Schmidt orthogonalization and takes the following form:

$$\left.\begin{array}{l} Q_1 = S_1 \\[2mm] Q_j = S_j - \displaystyle\sum_{i=1}^{j-1} c_{ij} Q_i, \, j = 2, 3, \ldots, p \end{array}\right\} \qquad (3.2\text{-}29)$$

where

$$c_{ij} = \frac{Q_i^T S_j}{Q_i^T Q_i},$$

Q_i=the transformed vector orthogonal to vectors already in Q,

S_j=the next column vector of S to be transformed.

3.3 Regression When the Model is Nonlinear

3.3.1 Modified Gauss-Newton Method

If the model is nonlinear in the parameters but is linear in the dependent variable, then the model may be written in the standard form for nonlinear regression:

$$Y = f(\xi_1, \xi_2, \ldots, \xi_k; \beta_1, \beta_2, \ldots \beta_p) + \epsilon \, . \qquad (3.3\text{-}1)$$

Because of the nonlinearity, f cannot be written in the form $f = X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p$. The case more complicated than equation 3.3-1, where the model is nonlinear in both the parameters and the dependent variable, is treated in section 6.1. When there are n observations, equation 3.3-1 may be written in matrix form as

$$Y = f(\xi, \beta) + \epsilon \qquad (3.3\text{-}2)$$

or, in terms of general estimate b of β and estimate e of ϵ,

$$Y = f(\xi, b) + e \, . \qquad (3.3\text{-}3)$$

As for the linear case, the regression solution of equation 3.3-3 is obtained by minimizing the weighted error sum of squares:

$$\begin{aligned} S(b) &= e^T \omega e \\ &= (Y - f(\xi, b))^T \omega (Y - f(\xi, b)) \, . \end{aligned} \qquad (3.3\text{-}4)$$

However, because equation 3.3-3 is nonlinear, solution of the problem is not as direct as it was for the linear case.

A convenient and robust solution method is obtained by linearizing equation 3.3-3 around an initial estimate of parameters, then proceeding as if the problem were linear. This yields a new set of parameters that minimizes equation 3.3-4 where f is replaced by the linear approximation. The new parameters are then substituted for the initial set, and the process is repeated to yield a better set of parameters. The iterative process stops whenever the change in calculated parameters from one iteration to the next is small. At that point the minimum of equation 3.3-4 has been found.

To derive the method, first $f(\xi, b)$ is expanded about the initial set of parameters b_0 by using a truncated Taylor series to obtain a linear approximation for $f(\xi, b)$:

$$f(\xi, b) \cong f(\xi, b_0) + X_0(b - b_0) \qquad (3.3\text{-}5)$$

where

$$X_0 = \{X_{ij}^0\} = \left\{\frac{\partial f_i}{\partial b_j}\bigg|_{b=b_0}\right\}(n \times p) \qquad (3.3\text{-}6)$$

and f_i is f calculated at the ith observation point. The components of $\underline{X}_0$ are recognized as sensitivity coefficients, or simply sensitivities.

By using equation 3.3–5, equation 3.3–3 may be written as

$$\underline{Y}-f(\underline{\xi},\underline{b}_0) \cong \underline{X}_0(\underline{b}-\underline{b}_0)+\underline{e} \qquad (3.3–7)$$

which is of the form of the incremental linear model of section 3.2. Note that if the model is linear so that $f(\underline{\xi},\underline{b})=\underline{X}(\underline{\xi})\underline{b}$, then the truncated Taylor series and hence equation 3.3–7 are exact. In this case expansion in a Taylor series is another way of deriving the incremental linear model. If the model is nonlinear, equation 3.3–7 is the approximate (linearized) model for parameters in the vicinity of $\underline{b}_0$, as illustrated by a simple one-parameter example in figure 3.3–1.

An approximate best estimate of β (which is exact for a linear model) can be obtained by minimizing $S(\underline{b})=\underline{e}^T\underline{\omega}\underline{e}\cong(\underline{Y}-f(\underline{\xi},\underline{b}_0)-\underline{X}_0(\underline{b}-\underline{b}_0))^T$ $\cdot\underline{\omega}(\underline{Y}-f(\underline{\xi},\underline{b}_0)-\underline{X}_0(\underline{b}-\underline{b}_0))$ with respect to $\underline{b}$. This process is carried out exactly like it was for the linear model and yields the set of normal equations

$$\underline{X}_0^T\underline{\omega}\underline{X}_0\underline{d}_1=\underline{X}_0^T\underline{\omega}(\underline{Y}-f(\underline{\xi},\underline{b}_0)) \qquad (3.3–8)$$

where subscript 1 indicates the first approximate solution and

$$\underline{d}_1=\underline{b}_1-\underline{b}_0 . \qquad (3.3–9)$$

To reduce round-off error in nonlinear regression, it is generally useful to scale equation

$f(\xi, b_1)$

b_1^0 b_1

$-f(\xi, b_1^0)+X_1^0(b_1-b_1^0)$
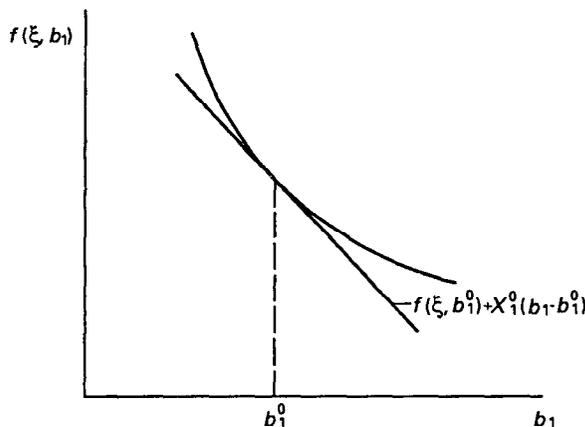
Figure 3.3–1

3.3–8 in the same manner as for linear regression (equations 3.2–14 through 3.2–17), although another type of scaling that is useful for some types of problems is introduced in problem 3.3–1. Scaling equation 3.3–8 produces

$$\underline{S}_0^T\underline{\omega}\underline{S}_0\underline{\delta}_1=\underline{S}_0^T\underline{\omega}(\underline{Y}-f(\underline{\xi},\underline{b}_0)) \qquad (3.3–10)$$

where

$$\underline{S}_0=\underline{X}_0\underline{C}_0 \qquad (3.3–11)$$

$$\underline{\delta}=\underline{C}_0^{-1}\underline{d}_1 \qquad (3.3–12)$$

$$\underline{C}_0=\text{diag}\{(A_{11}^0)^{-\frac{1}{2}},(A_{22}^0)^{-\frac{1}{2}},....,(A_{pp}^0)^{-\frac{1}{2}}\} \qquad (3.3–13)$$

$$\underline{A}_0=\underline{X}_0^T\underline{\omega}\underline{X}_0 \qquad (3.3–14)$$

and A_{ii}^0 is a diagonal component of $\underline{A}_0$.

Because equation 3.3–8 is not exact, equation 3.3–4 is not truly minimized, and $\underline{b}_1$ is not actually the optimal set. Hence, $\underline{b}_1$ is substituted for $\underline{b}_0$, and the entire process is repeated to yield another, hopefully improved, estimate. As a general iteration equation, 3.3–10 may be written in the form

$$\underline{S}_r^T\underline{\omega}\underline{S}_r\underline{\delta}_{r+1}=\underline{S}_r^T\underline{\omega}(\underline{Y}-f(\underline{\xi},\underline{b}_r)) \qquad (3.3–15)$$

where

$$\underline{S}_r=\underline{X}_r\underline{C}_r \qquad (3.3–16)$$

$$\underline{\delta}_{r+1}=\underline{C}_r^{-1}(\underline{b}_{r+1}-\underline{b}_r)=\underline{C}_r^{-1}\underline{d}_{r+1} \qquad (3.3–17)$$

$$\underline{C}_r=\text{diag}\{(A_{11}^r)^{-\frac{1}{2}},(A_{22}^r)^{-\frac{1}{2}},....,(A_{pp}^r)^{-\frac{1}{2}}\} \qquad (3.3–18)$$

and A_{ii}^r is defined analogously to A_{ii}^0. As the process converges, $\underline{\delta}_{r+1}\rightarrow\underline{0}$ and equation 3.3–7 becomes $\underline{Y}-f(\underline{\xi},\hat{\underline{b}})=\hat{\underline{e}}$. At the same point, $\hat{\underline{b}}$ minimizes $S(\underline{b})$ in equation 3.3–4, or $S(\hat{\underline{b}})=(\underline{Y}-f(\underline{\xi},\hat{\underline{b}}))^T\underline{\omega}(\underline{Y}-f(\underline{\xi},\hat{\underline{b}}))$, which is a minimum for the nonlinear equation. This process for finding the minimum of $S(\underline{b})$ is known as the Gauss-Newton method.

A sketch of progression of the iterations to the minimum for a hypothetical two-parameter problem is given in figure 3.3–2. As diagrammed
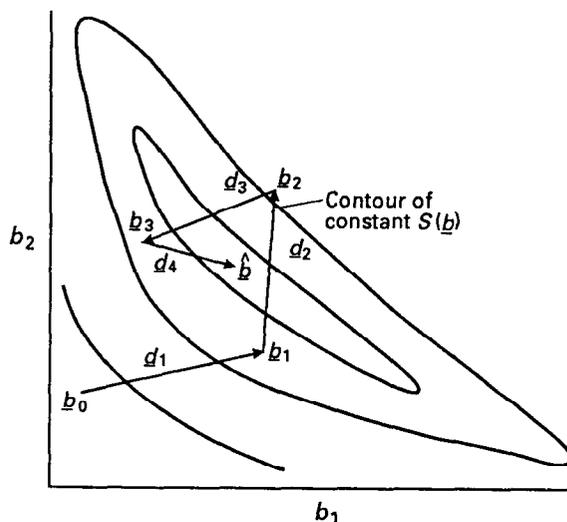
Figure 3.3-2

in figure 3.3-2, the solution often does not progress directly toward the minimum.

Modifications to the Basic Procedure. It is well known that the Gauss-Newton method as defined by equation 3.3-15, does not always converge. To help induce convergence, a damping parameter, ρ, is introduced by modifying equation 3.3-17 to give

$$\underline{b}_{r+1}=\rho\underline{d}_{r+1}+\underline{b}_r \ . \qquad (3.3\text{-}19)$$

where $\underline{d}_{r+1}=\underline{C}_r\underline{\delta}_{r+1}$. If $0<\rho<1$, the changes in computed parameters are less than would result for $\rho=1$; thus the method is an interpolation method. Similarly, if $\rho>1$, the method is an extrapolation method.

Inspection of equation 3.3-19 reveals that ρ changes the magnitude of the displacement from $\underline{b}_r$ to $\underline{b}_{r+1}$. However, because all components of the displacement vector $\underline{d}_{r+1}$ are scaled by the same multiplier ρ, the direction of the displacement vector is not altered. If the displacement vector $\underline{d}_{r+1}$ is oriented in a direction nearly parallel to a contour in the sum of squares surface ($S(\underline{b})$), then little, if any, improvement (in terms of reducing $S(\underline{b})$) can result from solution of equation 3.3-15. In this case it would be desirable to alter the direction of $\underline{d}_{r+1}$ to point closer to a down-gradient direction. For example, in figure 3.3-3 vector $\underline{d}_{r+1}$

yields no improvement in estimates $\underline{b}_r$, but vector $\underline{d}'_{r+1}$ yields a significant improvement in the estimates.

A modification that accomplishes the desired alteration of direction of $\underline{d}_{r+1}$, and reduces its magnitude as well, consists of adding a positive parameter μ, known as the Marquardt parameter (Marquardt, 1963), to the main diagonal of the coefficient matrix $\underline{S}_r^T\omega\underline{S}_r$ of equation 3.3-15. Scaling is needed so that μ can have the same effect on each entry of the main diagonal of the coefficient matrix. The scaling accomplishes this effect because each entry of the main diagonal of the scaled matrix $\underline{S}_r^T\omega\underline{S}_r$ is unity. Mathematically, the Marquardt modification can be stated as follows.

$$\left(\underline{S}_r^T\underline{\omega}\underline{S}_r+\mu\underline{I}\right)\underline{\delta}_{r+1}=\underline{S}_r^T\underline{\omega}\left(\underline{Y}-f(\underline{\xi},\underline{b}_r)\right) . \qquad (3.3\text{-}20)$$

Solution Algorithm. The sequential steps implementing the modified Gauss-Newton procedure are:

1. Calculate $f(\underline{\xi},\underline{b}_r)$ and $\underline{S}_r$ using initial parameters $\underline{b}_r$ and the combination of equation 3.3-6 (with index r replacing 0) and equation 3.3-16.
2. Solve equation 3.3-20 for $\underline{\delta}_{r+1}$.
3. Solve equation 3.3-17 for $\underline{d}_{r+1}$.
4. Solve equation 3.3-19 for $\underline{b}_{r+1}$.
5. Test to determine if $|d_i^{r+1}/c|>\epsilon$, where ϵ is
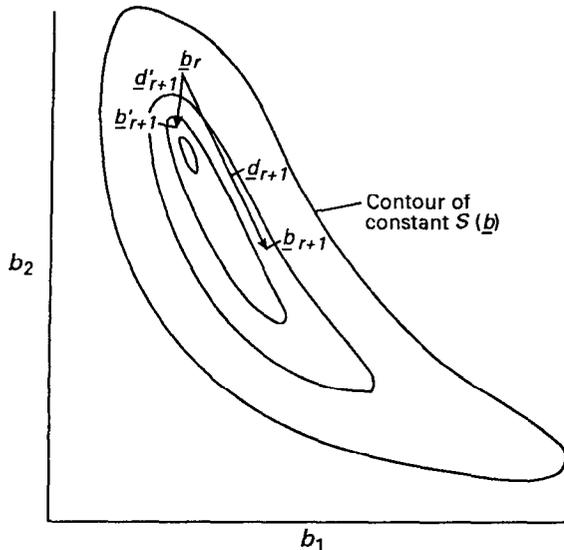
Figure 3.3–3

a small number such as 0.01, $c=b_i^r$ for $b_i^r \neq 0$, and $c=1$ for $b_i^r=0$.

6. If $|d_i^{r+1}/c|>\epsilon$, increment r by one and return to 1. If not, then the process has converged.

Problem 3.3–1

The Theis equation for flow to a well in a confined, nonleaky aquifer is

$$s= \frac{Q}{4\pi T} \int_{r_o^2 S/4Tt}^{\infty} \frac{e^{-z}}{z} dz \qquad (1)$$

where t=time (d);
s=drawdown (ft);
r_o=radial distance to observation well (ft);
Q=discharge (ft³/s);
T=transmissivity (ft²/s); and
S=storage coefficient.

The integral can be evaluated by summing the infinite series

$$\int_u^{\infty} \frac{e^{-z}}{z} dz = -0.577216 - \ln u + u$$

$$-\frac{u^2}{2\cdot2!} + \frac{u^3}{3\cdot3!} - \frac{u^4}{4\cdot4!} + \ldots \qquad (2)$$

where $n! = n\cdot(n-1)\cdot(n-2)\ldots3\cdot2\cdot1$.

The Theis equation is nonlinear in the parameters T and S. Using the information in section 3.3.1, develop an algorithm for solving this equation for T and S, given time and drawdown data.

a. Let

$$f(t, r_o; T, S)= \frac{Q}{4\pi T} W(u) \qquad (3)$$

where $u=r_o^2 S/4Tt$ and W(u) is the integral in equation 2. Find the sensitivities for T and S. (See equation 3.3–6. Hint:

$$\frac{d}{d\alpha} \int_{u(\alpha)}^{\infty} f(x)dx = -f(u(\alpha)) \frac{du(\alpha)}{d\alpha}.)$$

b. Assume that initial estimates of transmissivity and storage coefficient, T_0 and S_0, exist. In equation 3.3–6 let $j=T$ indicate the sensitivity for T, and $j=S$ indicate the sensitivity for S. Then note that X_{iT}^0 can be scaled to become $Z_{iT}^0=X_{iT}^0 T_0$, and X_{iS}^0 can be scaled to become $Z_{iS}^0=X_{iS}^0 S_0$. Modify the functions for X_{iT}^0 and X_{iS}^0 computed in step a to become the scaled functions Z_{iT}^0 and Z_{iS}^0. Do you see any resulting simplifications in arithmetic? Do you think that the scaled sensitivities Z_{iT}^0 and Z_{iS}^0 might be more nearly uniform in value for any fixed i than X_{iT}^0 and X_{iS}^0? What do you think this uniformity in value accomplishes?

c. Construct, for n time observations at a single spatial location, the incremental linear model (equation 3.3–7). Then transform this model so that scaled sensitivities Z_{iT} and Z_{iS} are used instead of X_{iT} and X_{iS}. How does the parameter displacement vector $\underline{d}$ transform? How can you recover $\underline{d}$ from the transformed displacement vector? (Hint: equations 3.3–10 through 3.3–14 are analogous to the present scaling problem.)

Figure 1 is a flow diagram for programming the steps indicated in section 3.3.1. Obtain a coding sheet and proceed to write code according to this flow diagram. Helpful hints, numbered on the flow diagram, follow:

FLOW CHART
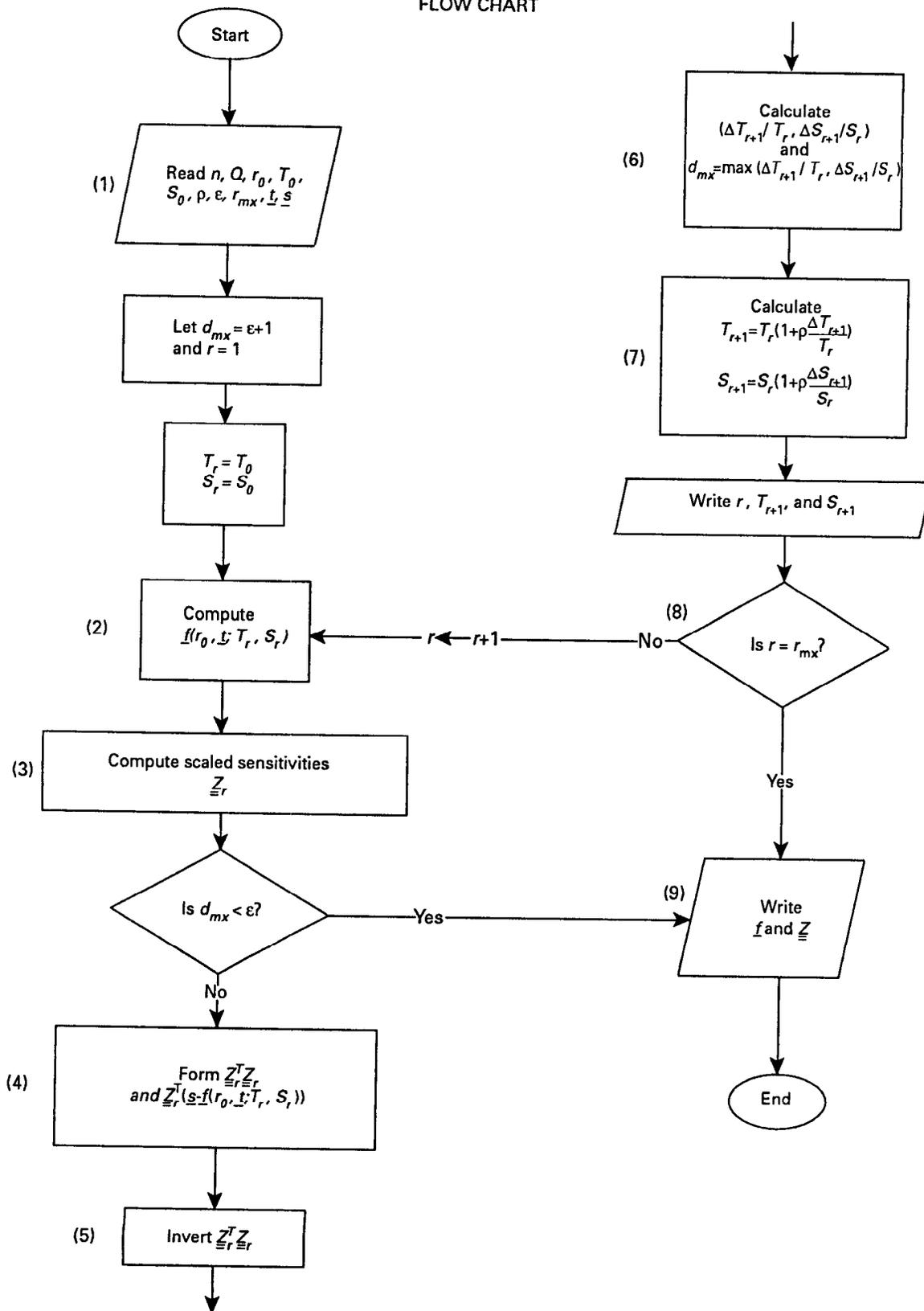
Figure 1

1. n=number of observations;
    r_0=radial distance to observation well;
    t_i=time of each observation;
    s_i=observed drawdown for each time t_i;
    T_0=initial guess, transmissivity;
    S_0=initial guess, storage coefficient;
    Q=discharge;
    ρ=damping parameter;
    ϵ=convergence criterion;
    r_{mx}=maximum number of iterations.

2. $f_i(r_0,t_i;T_r,S_r)= \dfrac{Q}{4\pi T_r}\, W(r_0,t_i;T_r,S_r)$.

You will have to program a finite number of terms of the infinite series to evaluate $W(r_0,t_i;T_r,S_r)$. Note that one computation of f_i will occur for every time observation t_i. These computations form the column vector f.

3. The sensitivities will form an $n\times2$ array (one sensitivity for each parameter T and S; sensitivities are evaluated for each observation). Use the information in parts a and b above.

4. Matrix multiplication is done with the basic algorithm

$$c_{ij}= \sum_{k=1}^{n} a_{ik}b_{kj}$$

where a_{ik}, b_{kj}, and c_{ij} are elements of matrices $\underline{A}$, $\underline{B}$, and $\underline{C}$, respectively, and $\underline{C}=\underline{AB}$. If $\underline{A}=\underline{B}^T$, then

$$c_{ij}= \sum_{k=1}^{n} b_{ki}b_{kj}$$

because $b_{ik}^T=b_{ki}$. Similarly, if $\underline{A}=\underline{D}^T$, then $\underline{C}=\underline{AB}$ becomes

$$c_{ij}= \sum_{k=1}^{n} d_{ki}b_{kj} \ .$$

Note that the model error variance Var($\underline{\epsilon}$) is assumed to be $\underline{I}\sigma^2$.

5. Use the definition of the inverse to construct the inverse.

6. See part c above. Also, note that $\underline{d}_{r+1}$ $=[\Delta T_{r+1},\Delta S_{r+1}]$.

7. Include the damping parameter ρ in your calculation of the new regression parameters. If convergence does not occur, then you may be best advised to set ρ to a value less than one.

8. This is the end of the iteration loop.

9. You may wish to compute the final estimate of residuals $\underline{e}=\underline{f}-\underline{s}$, then print them also.

Because convergence problems may arise, limit the number of passes that can be made through the algorithm. Test your code using the following data.

36-Hour Pumping Test

Test starts: 12 February 1976, 0805 h.
Test ends: 13 February 1976, 2005 h.
Production well was pumped varying between 517 gal/min and 530 gal/min. Time and drawdown data at observation well 175 ft from production well appear in table 1.

Table 1

[Data from S.P. Larson, presently of S.S. Papadopulos and Associates (formerly U.S. Geological Survey), 1978]

Time (s)	Drawdown (ft)
480	1.71
1,020	2.23
1,500	2.54
2,040	2.77
2,700	3.04
3,720	3.25
4,920	3.56

To aid in debugging your computer program, some example calculations follow. Assume that T_0=0.1 ft^2/s, S_0=0.0005, Q=1.16 ft^3/s, and r_0=175 ft. Then

$$u_i^0= \frac{r_0^2 S_0}{4T_0 t_i} = \frac{(175)^2(0.0005)}{4(0.1)t_i} =38.28125/t_i$$

$$f_i^0= \frac{Q}{4\pi T_0}\, W(u_i^0)= \frac{1.16}{4\pi(0.1)}\, W(u_i^0)=0.923099 W(u_i^0)$$

Use of these values for u_i^0 and f_i^0 allows for calculation of the following data.

t_i	u_i^0	$W(u_i^0)$	f_i^0
480	0.0797526	2.02980	1.87371
1,020	.0375306	2.74256	2.53165
1,500	.0255208	3.11640	2.87675
2,040	.0187653	3.41721	3.15442
2,700	.0141782	3.69296	3.40897
3,720	.0102907	4.00957	3.70123
4,920	.00778074	4.28665	3.95700

Data for u_i^0 and f_i^0 are used to compute Z_{iT}^0 and Z_{iS}^0.

$$Z_{iT}^0 = -f_i^0 + \frac{Q}{4\pi T_0} e^{-u_i^0} = -f_i^0 + 0.923099 \ e^{-u_i^0}$$

$$Z_{iS}^0 = -\frac{Q}{4\pi T_0} e^{-u_i^0} = -0.923099 e^{-u_i^0}$$

Note that $Z_{iT}^0 = X_{iT}^0 T_0$ and $Z_{iS}^0 = X_{iS}^0 S_0$ are calculated directly, without first calculating sensitivities X_{iT}^0 and X_{iS}^0, then formally making the multiplication. These scaled sensitivities result from the following scaled linearized model.

$$f_i^1 = f_i^0 + X_{iT}^0 \Delta T_1 + X_{iS}^0 \Delta S_1$$

$$= f_i^0 + X_{iT}^0 T_0 \left(\frac{\Delta T_1}{T_0}\right) + X_{iS}^0 S_0 \left(\frac{\Delta S_1}{S_0}\right)$$

$$= f_i^0 + Z_{iT}^0 \frac{\Delta T_1}{T_0} + Z_{iS}^0 \frac{\Delta S_1}{S_0}$$

The computed values for Z_{iT}^0 and Z_{iS}^0 are:

z_{iT}^0	z_{iS}^0
-1.02137	-0.852339
-1.64255	-.889097
-1.97691	-.899839
-2.24848	-.905938
-2.49887	-.910103
-2.78758	-.913648
-3.04106	-.915944

Use of the computed values for Z_{iT}^0 and Z_{iS}^0 yields the entries in the $\underline{Z}^T\underline{Z}$ matrix:

$$\sum_i \left(Z_{iT}^0\right)^2 = 35.96800209 \quad \sum_i \left(Z_{iS}^0\right)^2 = 5.649402682$$

$$\sum_i Z_{iT}^0 Z_{iS}^0 = 13.75336059$$

Elements of the $\underline{Z}^T(\underline{s}-\underline{f}_0)$ vector are computed in the following manner.

z_{iT}^0	s_i	f_i^0	$z_{iT}^0(s_i-f_i^0)$
-1.02137	1.71	1.87371	0.1672084827
-1.64255	2.23	2.53165	.4954752075
-1.97691	2.54	2.87675	.6657244425
-2.24848	2.77	3.15442	.8643606816
-2.49887	3.04	3.40897	.9220080639
-2.78758	3.25	3.70123	1.257839723
-3.04106	3.56	3.95700	1.20730082

$$\sum_i Z_{iT}^0\left(s_i-f_i^0\right) = 5.579917421$$

z_{iS}^0	s_i	f_i^0	$z_{iS}^0(s_i-f_i^0)$
-0.852339	1.71	1.87371	0.1395364177
-.889097	2.23	2.53165	.2681961101
-.899839	2.54	2.87675	.3030207833
-.905938	2.77	3.15442	.3462606860
-.910103	3.04	3.40897	.3358007039
-.913648	3.25	3.70123	.4122653870
-.915944	3.56	3.95700	.3636297680

$$\sum_i Z_{iS}^0\left(s_i-f_i^0\right) = 2.170709856$$

Finally, the two elements of the scaled displacement vector $(\Delta T_1/T_0, \Delta S_1/S_0)$ are computed and used to compute T_1 and S_1, the new estimates for T and S.

$$\frac{\Delta S_1}{S_0} = \frac{(2.170709856-(13.75336059)}{\cdot (5.579917421)/35.96800209)}$$
$$/(5.649402682-(13.75336059)^2$$
$$/35.96800209)$$

$$= \frac{0.03707407158}{0.3904247981} = 0.09495829097$$

$$\frac{\Delta T_1}{T_0} = (5.579917421$$

$$- (0.09495829097)(13.75336059))$$
$$/35.96800209$$

$$= 0.1188256660$$

$$S_1 = (1+0.0949583)(0.0005) = 0.000547479$$

$$T_1 = (1+0.118826)(0.1) = 0.111883$$

Some similar calculations for the second iteration are:

$$u_i^1 = \frac{(175)^2(0.000547479)}{4(0.111883)t_i} = 37.4645/t_i$$

$$f_i^1 = \frac{1.16}{4\pi(0.111883)} W(u_i^1) = 0.825057 W(u_i^1)$$

t_i	u_i^1	$W(u_i^1)$	f_i^1
480	0.0780510	2.04973	1.69114
1020	.0367299	2.76334	2.27991
1500	.0249763	3.13743	2.58856
2040	.0183650	3.43838	2.83686
2700	.0138757	3.71423	3.06445
3720	.0100711	4.03091	3.32573
4920	.00761474	4.30805	3.55439

$$Z_{iT}^1 = -f_i^1 + 0.825057 e^{-u_i^1}$$

$$Z_{iS}^1 = -0.825057 e^{-u_i^1}$$

z_{iT}^1	z_{iS}^1
-0.928031	-0.763109
-1.48461	-.795303
-1.78385	-.804705
-2.02682	-.810043
-2.25076	-.813688
-2.50894	-.816789
-2.73559	-.818798

3.3.2 Nonlinear Regression When the Model Is Numerical

The basic model equations assumed in all previous developments have been of the closed form or analytical type where the dependent variable f is a known function of ξ and β. In many cases such models may either not exist or be too complicated for practical use. In these cases the basic equation relating the dependent variable to the independent variables and parameters may be a numerical solution that can be stated in the general form

$$\underline{D}(\underline{h},\underline{\xi},\underline{\beta})\underline{h} = \underline{q}(\underline{h},\underline{\xi},\underline{\beta}) \ . \tag{3.3-21}$$

Equation 3.3-21 is a nonlinear matrix equation in which $\underline{h}$ is the solution (dependent variable) vector of order m; $\underline{D}$ is a nonsingular coefficient matrix of order m that is a function of $\underline{h}$, $\underline{\xi}$, and $\underline{\beta}$; and $\underline{q}$ is a vector of order m that is a function

of $\underline{h}$, $\underline{\xi}$, and $\underline{\beta}$. Order m is not related to the number of observations n, but instead is simply the order required to give a good numerical approximation to the solution of the problem.

If equation 3.3-21 is linear in $\underline{h}$ so that $\underline{D}$ and $\underline{q}$ are not functions of $\underline{h}$, then equation 3.3-21 may be solved directly for $\underline{h}$. In this case the Gauss-Newton method may be used to obtain the regression solution. (The nonlinear case is considered in section 6.1.2.) The procedure is as follows. First write equation 3.3-21 in the form

$$\underline{h} = \underline{D}^{-1}\underline{q} \tag{3.3-22}$$

which is explicit in the dependent variable $\underline{h}$. Next, note that $\underline{h}$ in equation 3.3-22 and f in equation 3.3-2 (or, as an estimate, equation 3.3-3) are expressions of the same quantity, the only difference between them being that elements of $\underline{h}$ are values of the dependent variable computed at points defined by the numerical solution, and elements of f are values of the dependent variable computed at observation points. If all n_s observation points are contained in the set of points required for the numerical solution, which implies that $m \geq n_s$, then f is obtained from $\underline{h}$ simply by eliminating those entries in $\underline{h}$ not corresponding to observation points. In other instances the points in m may not correspond to those in n_s. For these instances an interpolation scheme would be used to obtain f from $\underline{h}$. In either case, the vector $f(\underline{\xi}, \underline{b}_r)$ is obtained by using $\underline{h}$ computed from equation 3.3-22 in which $\underline{b}_r$ was used to evaluate $\underline{D}$ and $\underline{q}$.

The final step in forming the Gauss-Newton solution is to derive the sensitivity matrix $\underline{\underline{X}}$. To accomplish this step, write equation 3.3-21 in terms of a general parameter set $\underline{b}$, then differentiate it with respect to $\underline{b}$ to yield

$$\underline{D}\frac{\partial \underline{h}}{\partial b_j} + \frac{\partial \underline{D}}{\partial b_j}\underline{h} = \frac{\partial \underline{q}}{\partial b_j}, j=1,2,\ldots,p \tag{3.3-23}$$

or

$$\frac{\partial \underline{h}}{\partial b_j} = \underline{D}^{-1}\left(\frac{\partial \underline{q}}{\partial b_j} - \frac{\partial \underline{D}}{\partial b_j}\underline{h}\right), j=1,2,\ldots,p. \tag{3.3-24}$$

The quantity $\partial \underline{h}/\partial b_j$ forms a column of the sensitivity matrix for points in m. Sensitivity

matrix $\underline{X}_r$ would be found by first computing $\partial h/\partial b_j$ using equation 3.3-24 written in terms of $\underline{h}_r$ and $\underline{b}_r$, then following one of the two procedures described above for obtaining f from $\underline{h}$. By incorporating the procedures to compute numerical estimates of $f(\xi,\underline{b}_r)$ and $\underline{X}_r$, the algorithm given for the Gauss-Newton method may be followed exactly to find the regression solution of equation 3.3-22.

By studying the sequence of calculations in the solution algorithm carefully, it will be noted that to calculate $\underline{h}_r$ and $\underline{X}_r$, $\underline{D}_r$ ($\underline{D}$ computed using $\underline{b}_r$) must already have been computed. This requires one of two possible calculation schemes. Either $\underline{D}_r$ and p matrices of the form $(\partial\underline{D}/\partial b_j)_r$ must be formed at the same time and stored before $\underline{h}_r$ and $\underline{X}_r$ are calculated, or $\underline{D}_r$ must be formed before $\underline{h}_r$ is calculated and then each matrix $(\partial\underline{D}/\partial b_j)_r$ formed as needed to calculate each column (that is, $\underline{X}_j^r$) of $\underline{X}_r$. The first alternative could require a considerable amount of computer memory or the use of scratch files, whereas the second alternative could require repetitive calculation because many arithmetic operations could be the same for forming both $\underline{D}_r$ and $(\partial\underline{D}/\partial b_j)_r$. Often, however, matrices $(\partial\underline{D}/\partial b_j)_r$ can be written in a condensed or decomposed form to conserve computer memory. In this way $\underline{D}_r$ and the decomposed form of $(\partial\underline{D}/\partial b_j)_r$ can be computed together without using a significant amount of extra memory. Then each matrix $(\partial\underline{D}/\partial b_j)_r$ may be assembled as needed without performing numerous repetitive calculations.

A significant amount of computer memory can also be wasted unless care is taken when forming $\underline{X}_r$. The general procedure is to form the column vector $(\partial q/\partial b_j)_r - (\partial\underline{D}/\partial b_j)_r\underline{h}_r$, then use equation 3.3-24 to form the vector $(\partial\underline{h}/\partial b_j)_r$, which replaces the first vector in central computer memory. From this, vector $\underline{X}_j^r$ is immediately formed and stored. The matrix composed of vectors $(\partial\underline{h}/\partial b_j)_r$ should not be stored in central computer memory because it is often large. If desired, it may be stored column by column on a scratch file for later retrieval and printing.

Problem 3.3-2

Assume the finite difference representation of a flow problem shown in figure 1.
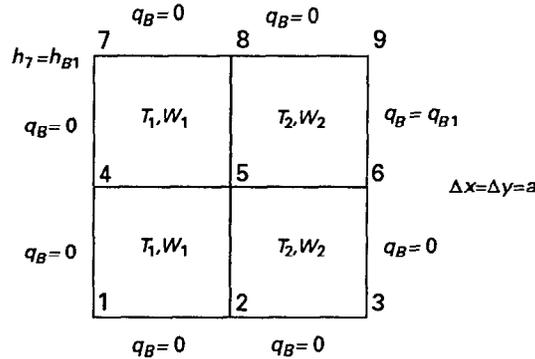
Figure 1

Finite difference equations for this problem can be written as follows.

1. $\frac{1}{2}\Delta y T_1(h_2-h_1)/\Delta x + \frac{1}{2}\Delta x T_1(h_4-h_1)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_1.$

2. $\frac{1}{2}\Delta y T_2(h_3-h_2)/\Delta x - \frac{1}{2}\Delta y T_1(h_2-h_1)/\Delta x$
$+ \frac{1}{2}\Delta x T_1(h_5-h_2)/\Delta y + \frac{1}{2}\Delta x T_2(h_5-h_2)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_1 - \frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2.$

3. $-\frac{1}{2}\Delta y T_2(h_3-h_2)/\Delta x + \frac{1}{2}\Delta x T_2(h_6-h_3)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2.$

4. $\Delta y T_1(h_5-h_4)/\Delta x + \frac{1}{2}\Delta x T_1(h_{B1}-h_4)/\Delta y$
$-\frac{1}{2}\Delta x T_1(h_4-h_1)/\Delta y = -\frac{1}{2}\Delta x\Delta y W_1.$

5. $\Delta y T_2(h_6-h_5)/\Delta x - \Delta y T_1(h_5-h_4)/\Delta x$
$+ \frac{1}{2}\Delta x T_1(h_8-h_5)/\Delta y + \frac{1}{2}\Delta x T_2(h_8-h_5)/\Delta y$
$-\frac{1}{2}\Delta x T_1(h_5-h_2)/\Delta y - \frac{1}{2}\Delta x T_2(h_5-h_2)/\Delta y$
$= -\frac{1}{2}\Delta x\Delta y W_1 - \frac{1}{2}\Delta x\Delta y W_2.$

6. $-\Delta y T_2(h_6-h_5)/\Delta x + \frac{1}{2}\Delta x T_2(h_9-h_6)/\Delta y$
$-\frac{1}{2}\Delta x T_2(h_6-h_3)/\Delta y = -\frac{1}{2}\Delta x\Delta y W_2 - \frac{1}{2}\Delta y q_{B1}.$

7. $h_7 = h_{B1}.$

8. $\frac{1}{2}\Delta y T_2(h_9-h_8)/\Delta x - \frac{1}{2}\Delta y T_1(h_8-h_{B1})/\Delta x$
$-\frac{1}{2}\Delta x T_1(h_8-h_5)/\Delta y - \frac{1}{2}\Delta x T_2(h_8-h_5)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_1 - \frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2.$

9. $-\frac{1}{2}\Delta y T_2(h_9-h_8)/\Delta x - \frac{1}{2}\Delta x T_2(h_9-h_6)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2 - \frac{1}{2}\Delta y q_{B1}.$

Or, by assuming that $\Delta x = \Delta y = a$,

1. $T_1(h_2-h_1) + T_1(h_4-h_1) = -\frac{1}{2}a^2 W_1.$
2. $T_2(h_3-h_2) - T_1(h_2-h_1) + (T_1+T_2)(h_5-h_2)$
$= -\frac{1}{2}a^2(W_1+W_2).$
3. $-T_2(h_3-h_2) + T_2(h_6-h_3) = -\frac{1}{2}a^2 W_2.$
4. $2T_1(h_5-h_4) + T_1(h_{B1}-h_4) - T_1(h_4-h_1)$
$= -a^2 W_1.$
5. $2T_2(h_6-h_5) - 2T_1(h_5-h_4) + (T_1+T_2)(h_8-h_5)$
$-(T_1+T_2)(h_5-h_2) = -a^2(W_1+W_2).$
6. $-2T_2(h_6-h_5) + T_2(h_9-h_6) - T_2(h_6-h_3)$
$= -a^2 W_2 - aq_{B1}.$
7. $h_7 = h_{B1}.$
8. $T_2(h_9-h_8) - T_1(h_8-h_{B1}) - (T_1+T_2)(h_8-h_5)$
$= -\frac{1}{2}a^2(W_1+W_2).$
9. $-T_2(h_9-h_8) - T_2(h_9-h_6) = -\frac{1}{2}a^2 W_2 - aq_{B1}.$

a. Verify the finite difference equations so that you understand their physical basis. (Hint: read appendix, section 4.3.1.)

b. Write the equations in matrix form:

$$\underline{\underline{D}}\underline{h}=\underline{q}$$

by explicitly writing out $\underline{\underline{D}}$, $\underline{h}$, and $\underline{q}$.

c. Let vectors $(\partial \underline{q}/\partial b_j)_r - (\partial \underline{\underline{D}}/\partial b_j)_r \underline{h}_r = \underline{J}_j^r$, and develop $\underline{J}_j^r$ $(j=1,2,3,4)$ for the parameters $\beta_1 = T_1$, $\beta_2 = T_2$, $\beta_3 = W_1$, and $\beta_4 = q_{B1}$.

d. Using the modified Gauss-Newton method, develop the algorithm (not computer program) to solve for the parameters. Assume that all nodes except number 7 are observation points.

3.3.3 Convergence and Conditioning

A value of ρ, $0 < \rho \leq 1$, can be shown to exist for which the Gauss-Newton procedure, as modified using equation 3.3-19, will converge to the global minimum value of $S(\underline{b})$ provided that:

1. An initial estimate of the parameters can be found such that they lie within a parameter region R bounded by sets of parameters $\underline{b}^*$ defined by

$$S(\underline{b}_0) \leq \min_{\underline{b}^*} S(\underline{b}^*) \qquad (3.3-25)$$

and the global minimum point lies within this region.

2. For all $\underline{b}$ belonging to R, $\underline{\underline{X}}$ is a continuous and unique matrix function.

3. The matrix $\underline{\underline{S}}^T \underline{\underline{\omega}} \underline{\underline{S}}$ is nonsingular and is a continuous function of $\underline{b}$.

Condition 2 is almost always met. Condition 1 requires that the system be well enough understood that intelligent initial estimates of parameters can be made. Difficulties frequently arise in connection with condition 3. Unless the problem is correctly specified, the least squares coefficient matrix (for example, $\underline{\underline{S}}^T \underline{\underline{\omega}} \underline{\underline{S}}$) can be singular. Moreover, problems often arise because of ill-conditioning (that is, near-singularity) of the matrix. Although the addition of the Marquardt parameter, μ, is intended to help these cases, convergence can be difficult to obtain. In the following paragraphs, the general question of convergence is considered first. This is followed by discussions of singularity and ill-conditioning.

In general, the rate of convergence has been found to be related to the number of parameters being estimated, as predicted by theory. That is, the greater the number of parameters, the slower the rate, all other things being equal. It is also related to the conditioning of the problem and to the nearness of the initial set of parameters to the optimum set, in that the rate of convergence is usually much faster near a minimum of $S(\underline{b})$. As a rule of thumb, one may often expect convergence within a number of iterations equal to either 5 or twice the number of parameters, whichever is greater. Fewer iterations are required for well-conditioned problems.

A problem that frequently retards the convergence rate, or even causes divergence, is overshoot. This happens when the parameter correction vector $\rho \underline{\delta}_{r+1}$ has a favorable orientation but is much longer than an ideal value. The result is that the new set of parameters $\underline{b}_{r+1}$ is almost as far as (or even further than) the old set $\underline{b}_r$ from the optimum value. A two-parameter example is illustrated in figure 3.3-4.

Overshoot is detected as large oscillations with accompanying changes in sign of components of $\rho \underline{\delta}_{r+1}$ from one iteration to the next. The remedy is to decrease the value of ρ such that $0 < \rho < 1$. In figure 3.3-4, a good value would be $\rho = 0.5$, which would give $\underline{b}_{r+1}$ at point 1.

Care must be taken not to make ρ too small so that undershoot becomes a problem. Undershoot occurs when $\rho \underline{\delta}_{r+1}$ is too small, and it manifests itself as small steps $\rho \underline{\delta}_{r+1}$, the components of which usually do not change sign. The remedy is to increase the value of ρ, in rare occasions such that $\rho > 1$. As a practical guide, the best value of ρ is one that causes some
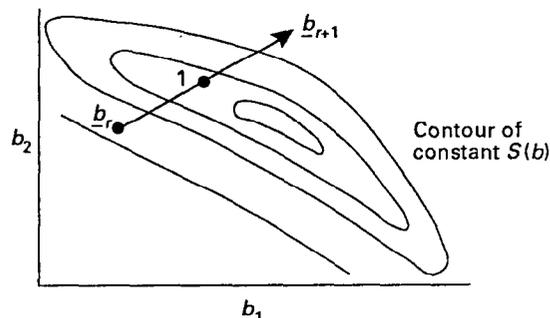
Figure 3.3-4

oscillations in sign of a few components of $\rho\underline{\delta}_{r+1}$ from one iteration to the next.

Sometimes the solution may converge to a local minimum instead of the global minimum. Most commonly this can occur (1) if ρ is too large and the overshoot causes the search vector $\rho\underline{\delta}_{r+1}$ to escape from the region defined by equation 3.3-25, or (2) if the initial estimate $\underline{b}_0$ is not in the region defined by equation 3.3-25. Detection of this problem is accomplished through adequate knowledge of the system so that the wrong solution can be recognized. If the computed parameter vector $\hat{\underline{b}}$ is physically illogical or the model analysis discussed in section 5 suggests that the model is not correct, then one might suspect that a local minimum has been reached. The remedy to the problem is to decrease ρ and (or) choose another initial estimate $\underline{b}_0$. If $\underline{b}_0$ is changed, usually the distance between $\underline{b}_0$ and the vector $\hat{\underline{b}}$ computed for the local minimum should be increased. Thus, if some components of $\hat{\underline{b}}$ are far too large to be realistic, then the same components of $\underline{b}_0$ should be reduced in value. If several attempts at changing ρ and (or) $\underline{b}_0$ do not produce a change in $\hat{\underline{b}}$, then the cause of the poor results is probably not a local minimum.

As for the linear case, singularity of the least-squares coefficient matrix occurs whenever columns of the sensitivity matrix are linearly dependent. Near-singularity, caused by near-linear dependence, is a more frequent occurrence. As a result of ill-conditioning, step sizes $\underline{\delta}_{r+1}$ can be highly erratic, appearing to head toward no well-defined point and can be dominated by overshoot. In addition, some problems may start fairly well conditioned for the initial parameters but may become progressively more poorly conditioned during the iterative solution process.

The same techniques for analysis of poorly conditioned problems as are used for linear problems may be used for nonlinear problems as well. Whenever a problem is poorly conditioned, the sensitivity matrix $\underline{X}$ may be examined for a near-zero column, and $\underline{S}^T\omega\underline{S}$ may be examined for off-diagonal components near unity. Also, the orthogonal transformation may be used to indicate that ill-conditioning exists and to point out possible columns where linear dependency occurs.

Use of the Marquardt parameter, μ, is intended to improve conditioning by adding a small quantity to the main diagonal of the least-squares coefficient matrix. Although conditioning is always artificially improved by employing μ, the parameters resulting from applying the least squares process to a very poorly conditioned problem may be considerably in error unless the actual causes of the poor conditioning are discovered and the conditioning improved without using μ.

3.3.4 Computation of μ and ρ

For best efficiency, both μ and ρ should be recomputed at each iteration, r. A number of schemes exist in the literature for making these computations, but virtually all schemes involve assuming several trial values of μ and ρ, then performing all of the calculations for iteration r for each of the trial values. The best values to use are then computed so as to minimize or substantially decrease $S(\underline{b})$. The problem with these schemes is that they require so much time that one is often much better off settling for approximate values of μ and ρ computed by using a much simpler scheme.

The scheme adopted here is derived from the considerations discussed in section 3.3.3. Parameter μ is used only when the problem is so poorly conditioned that the search direction $\underline{\delta}$ must be altered. Overshoot and undershoot are controlled primarily through use of ρ.

Because the Marquardt parameter is used with a scaled problem formulation, computations must be made using scaled quantities (see equation 3.3-16). By direct computation, it can be verified that the scaled gradient $(\partial S(\underline{b})/\partial b_j)$ $\cdot C_{jj}$ $(j=1,2,...,p)$ of the sum of squares $S(\underline{b})$ is given by

$$\underline{C}^T\frac{\partial S(\underline{b})}{\partial\underline{b}}=-\underline{g}=-\underline{S}^T\omega(\underline{Y}-f(\underline{\xi},\underline{b}))\ . \qquad (3.3\text{-}26)$$

Thus, by definition (Spiegel, 1959, p. 16), the angle between $\underline{g}$, which points directly down the sum of squares surface, and displacement vector $\underline{\delta}$ is given for the rth iteration by

$$\cos\Theta=\frac{\underline{\delta}_{r+1}^T\underline{g}_r}{\sqrt{(\underline{\delta}_{r+1}^T\underline{\delta}_{r+1})(\underline{g}_r^T\underline{g}_r)}}\ . \qquad (3.3\text{-}27)$$

For two parameters, b_1 and b_2, the relationships given by equation 3.3-27 are illustrated in figure 3.3-5. (Note the use of scaled parameters; see equation 3.3-17.)

If $\Theta=90°$, then, as discussed in section 3.3.1, no improvement in parameters is likely to result from application of the Gauss-Newton procedure. However, application of the Marquardt parameter, $\mu>0$, will result in $\Theta<90°$ (Marquardt, 1963) because vector $\underline{\delta}$ is shifted progressively toward g as μ increases. Thus, a viable scheme for choosing μ is to define a maximum value of Θ, $\Theta_{mx}<90°$, and compute μ so that Θ never exceeds Θ_{mx}. This can be accomplished rather simply. At the beginning of the regression set $\mu_0=0$. Then at each iteration r, check and recompute μ as necessary:

$$\mu_r=\mu_\ell$$
$$\text{if } \underline{\delta}_{r+1}^T\underline{g}_r\geq\cos\Theta_{mx}\sqrt{(\underline{\delta}_{r+1}^T\underline{\delta}_{r+1})(\underline{g}_r^T\underline{g}_r)} \text{ or}$$
$$\mu_{\ell+1}=\frac{3}{2}\mu_\ell+0.001 \qquad (3.3\text{-}28)$$
$$\text{if } \underline{\delta}_{r+1}^T\underline{g}_r<\cos\Theta_{mx}\sqrt{(\underline{\delta}_{r+1}^T\underline{\delta}_{r+1})(\underline{g}_r^T\underline{g}_r)}$$

At the beginning of iteration r, $\ell=1$ and $\mu_\ell=\mu_{r-1}$. Then equation 3.3-20 is solved and equation 3.3-28 is applied. If the second part of equation 3.3-28 is employed, equation 3.3-20 is resolved using $\mu_{\ell+1}$, ℓ is incremented by one, and equation 3.3-28 is used again. This process is continued until the first part of equation 3.3-28 is used, at which point the appropriate value of μ for iteration r has been found. The formula for computing $\mu_{\ell+1}$ from μ_ℓ is empirical but gives what experience has shown to be a good range in values of μ. For each resolution of equation 3.3-20, $\underline{S}_{r\omega}^T\underline{S}_r$ and g_r are not recomputed. Thus, the calculations are not extensive.

Computation of ρ is designed to prevent disastrous overshoot and to keep $\rho\underline{\delta}$ within the region R defined by equation 3.3-25. A simple but usually effective scheme is to estimate the maximum fraction that any of the parameters could change and still remain within R and then to prevent any parameter from changing any more than this amount over any iteration. Let t_{mx} be this maximum fractional change. Then at iteration r, ρ is calculated as follows:

$$t=\max_i|d_i^{r+1}/c| \qquad (3.3\text{-}29)$$
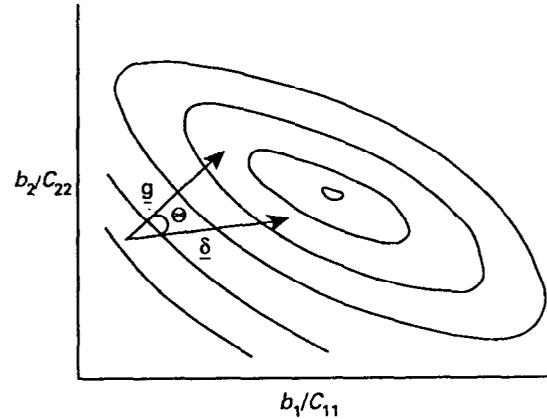
Figure 3.3-5

$$\left.\begin{array}{l}\rho=1 \text{ if } t\leq t_{mx} \text{ or} \\[2mm] \rho=t_{mx}/t \text{ if } t>t_{mx}\end{array}\right\} \qquad (3.3\text{-}30)$$

where $c=b_i^r$ if $b_i^r\neq 0$ and $c=1$ if $b_i^r=0$.

3.4 Regression Including Prior Information

3.4.1 Model Structure

Recall that the standard nonlinear regression model including prior information on the parameters may be written in the form (equation 3.1-32)

$$\underline{Y}=\underline{f}(\underline{\xi},\underline{\beta})+\underline{\epsilon} \qquad (3.4\text{-}1)$$

where

$$\underline{Y}=\begin{bmatrix}\underline{Y}_s\\[2mm]\underline{Y}_p\end{bmatrix} \qquad (3.4\text{-}2)$$

$$\underline{f}(\underline{\xi},\underline{\beta})=\begin{bmatrix}\underline{f}_s(\underline{\xi},\underline{\beta})\\[2mm]\underline{f}_p(\underline{\xi},\underline{\beta})\end{bmatrix} \qquad (3.4\text{-}3)$$

$$\underline{\epsilon}=\begin{bmatrix}\underline{\epsilon}_s\\[2mm]\underline{\epsilon}_p\end{bmatrix} \qquad (3.4\text{-}4)$$

and subscripts s and p indicate sample and prior information, respectively. To correspond with the above partitions into sample and prior information, the sensitivity matrix should be written in the form

$$\underline{X}=\begin{bmatrix} \underline{X}_s \\ \underline{X}_p \end{bmatrix} \qquad (3.4\text{-}5)$$

where $\underline{X}$ is a function of β for a nonlinear model. The transformation of $\underline{X}$ to $\underline{S}$ used to control round-off error is defined analogously. Finally, recall that the weight matrix is partitioned as

$$\underline{\omega}=\begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix} \qquad (3.4\text{-}6)$$

where $\underline{\omega}_s$ and $\underline{\omega}_p$ correspond to sample and prior information, respectively.

Often partition $\underline{X}_p$ will be obtained in a different manner than $\underline{X}_s$. For example, the model for the sample information may be numerical of the type in equation 3.3-22, whereas the model for the prior information may be of the analytical linear or nonlinear form. Thus, $\underline{X}_s$ would be obtained as described in section 3.3.2, and $\underline{X}_p$ would be obtained as described in section 3.3.1. Other types of differences are handled in a similar fashion. Obviously, if sample and prior models are of the same type, then $\underline{X}_s$ and $\underline{X}_p$ are obtained in the same manner.

Despite the possibilities for combinations of linear, nonlinear, analytical, and numerical models, remember that all models have the general form of the incremental linear model when expanded in the Taylor series. Because all models resolve to the incremental linear form, for simplicity subsequent discussions in this section are based on this model only.

3.4.2 Solution Procedures

Whenever $\underline{\omega}_s$ and $\underline{\omega}_p$ are both known, solution for both linear and nonlinear models is unaltered from that given in the previous sections. However, recall that, because of the block diagonal form of equation 3.4-6, $S(\underline{b})$ and, hence, the normal equations can be written in a special form.

By applying the standard minimization technique to $S(\underline{b})$ as given by equation 3.1-42, which can be written in the form

$$S(\underline{b})=[\underline{e}_s^T,\underline{e}_p^T]\begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix}\begin{bmatrix} \underline{e}_s \\ \underline{e}_p \end{bmatrix}, \qquad (3.4\text{-}7)$$

the normal equations for the incremental linear model become

$$[\underline{X}_s^T,\underline{X}_p^T]\begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix}\begin{bmatrix} \underline{X}_s \\ \underline{X}_p \end{bmatrix}(\underline{\hat{b}}-\underline{b})$$

$$=[\underline{X}_s^T,\underline{X}_p^T]\begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix}\begin{bmatrix} \underline{Y}_s-\underline{f}_s(\underline{\xi},\underline{b}) \\ \underline{Y}_p-\underline{f}_p(\underline{\xi},\underline{b}) \end{bmatrix} \qquad (3.4\text{-}8)$$

or

$$(\underline{X}_s^T\underline{\omega}_s\underline{X}_s+\underline{X}_p^T\underline{\omega}_p\underline{X}_p)(\underline{\hat{b}}-\underline{b})$$

$$=\underline{X}_s^T\underline{\omega}_s(\underline{Y}_s-\underline{f}_s(\underline{\xi},\underline{b}))+\underline{X}_p^T\underline{\omega}_p(\underline{Y}_p-\underline{f}_p(\underline{\xi},\underline{b})) \quad (3.4\text{-}9)$$

where $\underline{f}_s(\underline{\xi},\underline{b})$ and $\underline{f}_p(\underline{\xi},\underline{b})$ are for sample and prior information, respectively. Equation 3.4-9 is of the same form, and thus applies, as the equation for each iteration of solution of a nonlinear regression problem.

Frequently, the weight matrix is constructed from variance-covariance matrices for ϵ_s and ϵ_p that are given in the form

$$\mathrm{Var}(\underline{\epsilon}_s)=\underline{V}_s\sigma^2 \qquad (3.4\text{-}10)$$

$$\mathrm{Var}(\underline{\epsilon}_p)=\underline{U} \qquad (3.4\text{-}11)$$

where the usual form $\underline{V}_p\sigma^2$ for equation 3.4-11 cannot be used because $\mathrm{Var}(\underline{\epsilon}_p)$ is not known as a function of σ^2. Thus, with $\underline{\omega}$ defined as

$$\underline{\omega}=\begin{bmatrix} [\mathrm{Var}(\underline{\epsilon}_s)]^{-1} & \underline{0} \\ \underline{0} & [\mathrm{Var}(\underline{\epsilon}_p)]^{-1} \end{bmatrix}\sigma^2$$

$$=\begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{U}^{-1}\sigma^2 \end{bmatrix} \qquad (3.4\text{-}12)$$

then equation 3.4–9 becomes

$$(\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s + \underline{X}_p^T \underline{U}^{-1} \underline{X}_p \sigma^2)(\hat{\underline{b}} - \underline{b})$$

$$= \underline{X}_s^T \underline{V}_s^{-1}(\underline{Y}_s - \underline{f}_s(\underline{\xi}, \underline{b}))$$

$$+ \underline{X}_p^T \underline{U}^{-1} \sigma^2 (\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b})). \tag{3.4-13}$$

Hence, σ^2 apparently would have to be known to form the regression solution, whereas σ^2 is considered to be an unknown.

Theil (1963) showed that, for a linear model, σ^2 may be estimated for use in equation 3.4–13 by its ordinary least squares estimate (that is, the estimate obtained when prior information is not used). Bias produced by this estimate was shown by Theil (1963) to be of the order of $n_s^{-\frac{1}{2}}$. The procedure to be followed is to first solve the ordinary least squares problem by omitting all prior information; then find the estimate of σ^2 (to be given further on); finally use this estimate of σ^2 in the normal equations to solve the complete problem, including prior information.

References Cited

Bakr, A.A., Gelhar, L.W., Gutjahr A.L., and MacMillan, J.W., 1978, Stochastic analysis of spatial variability in subsurface flows, 1—Comparison of one- and three-dimensional flows: Water Resources Research, v. 14, no. 2, p. 263–271.

Draper, N. R., and Smith, H., 1981, Applied regression analysis [2d ed]: New York, John Wiley, 709 p.

Gutjahr, A.L., Gelhar, L.W., Bakr, A.A., and MacMillan, J.W., 1978, Stochastic analysis of spatial variability in subsurface flows, 2— Evaluation and application: Water Resources Research, v. 14, no. 5, p. 953–959.

Kitanidis, P.K., and Vomvoris, E.G., 1983, A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations: Water Resources Research, v. 19, no. 3, p. 677–690.

Krumbein, W.C., and Graybill, F.A., 1965, An introduction to statistical models in geology: New York, McGraw-Hill, 475 p.

Marquardt, D.W., 1963, An algorithm for least squares estimation of nonlinear parameters: Journal of the Society for Industrial and Applied Mathematics, v. 11, no. 2, p. 431–441.

Smith, L., and Freeze, R.A., 1979a, Stochastic analysis of steady-state groundwater flow in a bounded domain, 1—One-dimensional simulations: Water Resources Research, v. 15, no. 3, p. 521–528.

———1979b, Stochastic analysis of steady-state groundwater flow in a bounded domain, 2—Two-dimensional simulations: Water Resources Research, v. 15, no. 6, p. 1543–1559.

Spiegel, M.R., 1959, Theory and problems of vector analysis: New York, Schaum, 225 p.

Theil, H., 1963, On the use of incomplete prior information in regression analysis: American Statistical Association Journal, v. 58, no. 302, p. 401–414.

Additional Reading

Beck, J.V., and Arnold, K.J., 1977, Parameter estimation in engineering and science: New York, John Wiley, 501 p.

Draper, N. R., and Smith, H., 1981, Applied regression analysis [2d ed]: New York, John Wiley, 709 p.

4   Numerical Nonlinear Regression Solution of General Steady-State Ground-Water Flow Problems

4.1   Assumed Model and Solution Procedure

A model to solve fairly general steady-state ground-water flow problems by using the regression procedures presented in section 3 is given in this section. A complete description of the method is given first. Documentation and listing of the computer program are given in the appendix, section 4.3.4.

4.1.1   Problem Specification

The equation assumed to govern ground-water flow for the class of problems to be analyzed is derived from equation 1.1-1 by letting $\partial h/\partial t \rightarrow 0$, which results in

$$\frac{\partial}{\partial x}(T_{xx}\frac{\partial h}{\partial x})+\frac{\partial}{\partial y}(T_{yy}\frac{\partial h}{\partial y})+R(H-h)+W$$

$$+\sum_{\ell=1}^{N}\delta(x-a_\ell)\delta(y-b_\ell)Q_\ell=0 \ , \qquad (4.1\text{-}1)$$

where the symbols are defined the same as for equation 1.1-1.

Functions $T_{\xi\xi}$ (that is, T_{xx} and T_{yy}), R, and W are each formulated within the region being modeled as the product of a parameter and a given (or known) function. To provide for spatial variability of parameters, the region is subdivided into a number of discrete zones within each of which the parameters are assumed

constant. Hence, known spatial variability (often, but not necessarily, smooth or continuous) is superimposed upon the discontinuous spatial variability dictated by the parameter zonation. As an example, hydraulic conductivity $K_{\xi\xi}$ may often be considered to be constant within particular rock types, each of which may be considered to be a discrete zone. Thus $K_{\xi\xi}$ may be considered to be a parameter. Thickness b may be known from measurements and may vary continuously. The function $T_{\xi\xi}$ is, of course, given as $K_{\xi\xi}b$. Finally, because the controls that dictate a particular zonation may vary from parameter to parameter, zones for one type of parameter (for example, the parameter contained in $T_{\xi\xi}$) do not necessarily correspond to zones for another type (for example, the parameter contained in W). An example of zonation is given in figure 4.1-1 where the given function is unity so that the parameters are $T_{\xi\xi}$ and W.

Internal boundary conditions applying at discontinuities in $T_{\xi\xi}$ are that the specific discharge normal to the boundary and the hydraulic head both remain unchanged as the boundary is crossed. External boundary conditions applying on the periphery of the region being modeled include specified specific discharge normal to the boundary, specified hydraulic head at the boundary, or a mixture of the two types along the boundary.

Specific discharge q_B normal to the boundary is assumed to vary along the boundary in a manner similar to that of $T_{\xi\xi}$. It may have discontinuities and may vary smoothly between discontinuities. Discontinuities in q_B might often be expected to correspond to discontinuities in $T_{\xi\xi}$.

Hydraulic head variation along a specified head boundary is a continuous function, h_B, although the boundary may be subdivided into segments within each of which head h_B may vary linearly or curvilinearly with distance.

Dashed lines separate W zones, and solid lines separate T zones
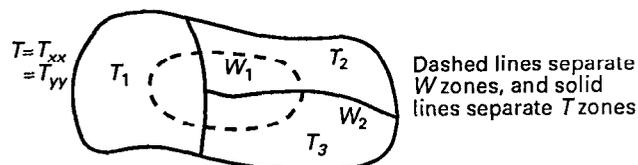
Figure 4.1-1

Subdivision into segments is often based on the causes of the known head conditions along various segments of the boundary.

Unknown quantities to be determined are $T_{\xi\xi}$, R, W, q_B, and h_B. Single values of T_{xx} or T_{yy}, R, and W (or multipliers such as K_{xx} or K_{yy}) are assumed to be parameters in each zone, although any of these parameters may be held constant (assumed known as <u>exact prior information</u>) in each zone. Separate zones (or segments) are specified for the values of q_B, and single multipliers for the fluxes in each flux zone are assumed to be parameters. Within each of the separate segments (zones) of a specified head boundary the heads are adjusted as a linear function of distance by the regression procedure so that the parameters are the values of head at each end of each segment. Even though the adjustment is linear, the actual shape of the head profile along the boundary may be curvilinear.

4.1.2 Matrix Form of Regression Model

For most field problems, equation 4.1–1 with its attendant boundary conditions cannot be solved analytically. Thus, the regression solution must be based on a numerical solution of equation 4.1–1, which is expressed as a matrix equation. The particular numerical solution method is given in the appendix, section 4.3.1.

The matrix equation comprising the numerical solution is given as

$$\underline{D}\underline{h}=\underline{q} \qquad (4.1\text{-}2)$$

where

$\underline{D}=$ the square coefficient matrix of order m, the number of nodes used to discretize the modeled region;

$\underline{h}=$ the hydraulic head vector of order m; and

$\underline{q}=$ the known vector of order m.

Matrix $\underline{D}$ contains parameters for $T_{\xi\xi}$ and R, whereas vector $\underline{q}$ can contain all parameter types. To express any specified head value, say $h_j=h_{Bj}$,

$$D_{jj}=1 \qquad D_{ji}=D_{ij}=0,\, i\neq j \qquad (4.1\text{-}3)$$

$$q_j=h_{Bj} \qquad (4.1\text{-}4)$$

where D_{jj}, D_{ij}, and D_{ji} are components of $\underline{D}$, and q_j is a component of $\underline{q}$. To accomplish the condition that $D_{ij}=0$ in equation i, $i\neq j$, the term $D_{ij}h_{Bj}$ is transferred to the right-hand side of equation i so that q_i contains the term $-D_{ij}h_{Bj}$. Then D_{ij} in $\underline{D}$ is set to zero.

The known head value is computed from

$$h_{Bj}=\frac{h^0_{Bj}[L_j H_t+(1-L_j)H_s]}{L_j H^0_t+(1-L_j)H^0_s} \qquad (4.1\text{-}5)$$

where

$s=$ node at one end of the boundary segment within which node j lies;

$t=$ node at the other end;

$H_s=$ head (parameter) at node s;

$H_t=$ head (parameter) at node t;

$L_j=$ ratio of distance along the boundary from node s to node j, to distance along the boundary from node s to node t; and

superscript $0=$ an initial or reference value.

Indices s and t can be equal so that $j=s=t$ for the case where only one specified head is present. Also, H_s and H_t can represent the same parameter, so that the entire specified head boundary behaves as a unit.

Because equation 4.1–2 is a linear matrix equation, the modified Gauss-Newton procedure is employed to solve the regression problem. Equation 4.1–2 is the same as equation 3.3–21, except that in equation 4.1–2 the coefficient matrix $\underline{D}$ and right-side vector $\underline{q}$ are not functions of dependent variable vector $\underline{h}$. Hence, sensitivities may be calculated using equation 3.3–24.

Prior information is assumed to be given (if available) on each parameter individually so that

$$\underline{X}_p=[\underline{I}_{n_p},\underline{0}]_{(n_p\times p)} \qquad (4.1\text{-}6)$$

where $\underline{I}_{n_p}$ is the identity matrix of order n_p. Thus, in equation 4.1–6, direct prior information is assumed to be given on the first n_p parameters. Placement of these parameters first in the vector $\underline{\beta}$ simplifies theoretical statement of equation 4.1–6 but is not necessary in practice.

The linearized regression model assumed, then, is of the form of equation 3.3-7 partitioned as suggested by equations 3.4-2 through 3.4-5:

$$\underline{Y}_s = \underline{f}_s(\underline{\xi},\underline{b}_r) + \underline{X}_s^r(\underline{b}_{r+1} - \underline{b}_r) + \underline{e}_s$$

$$\underline{Y}_p = \underline{f}_p(\underline{\xi},\underline{b}_r) + \underline{X}_p^r(\underline{b}_{r+1} - \underline{b}_r) + \underline{e}_p, \qquad \Big\} \quad (4.1-7)$$

where $\underline{Y}_s$ is the vector of heads at observation points, $\underline{f}_s(\underline{\xi},\underline{b}_r)$ is the vector of computed heads at observation points for iteration r, $\underline{Y}_p$ is the vector of prior estimates of the first n_p parameters, $\underline{f}_p(\underline{\xi},\underline{b}_r)$ is the vector of the first n_p elements of the computed parameter vector (on which there is prior information) for iteration r, and, from equation 4.1-6, $\underline{X}_p^r = [\underline{I}_{n_p}, \underline{0}]$.

No correlation or other coupling is assumed to exist among components of either $\underline{\epsilon}_s$ or $\underline{\epsilon}_p$. Matrix $\underline{\omega}$ is assumed to be of the form

$$\underline{\omega} = \begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{U}^{-1}\sigma^2 \end{bmatrix} \qquad (4.1-8)$$

where

$$\underline{V}_s^{-1} = [\mathrm{Var}(\underline{\epsilon}_s)]^{-1}\sigma^2 \qquad (4.1-9)$$

$$\underline{U}^{-1}\sigma^2 = [\mathrm{Var}(\underline{\epsilon}_p)]^{-1}\sigma^2 \qquad (4.1-10)$$

and both $\underline{V}_s^{-1}$ and $\underline{U}^{-1}$ are diagonal.

4.1.3 Nonlinear Regression Solution

Nonlinear-regression solution for the model given in section 4.1.2 is accomplished by using the algorithm at the end of section 3.3.1. Nodal sensitivities are calculated as illustrated in the appendix, section 4.3.2. We assume that observation points may be located anywhere within the flow region, so that computed heads, $\underline{f}_s(\underline{\xi},\underline{b}_r)$ and sensitivities, $\underline{X}_s^r$, at observation points in general must be obtained by interpolation from surrounding nodal values. Standard bilinear interpolation using the four adjacent nodes surrounding an observation is used as the interpolation method (Wang and Anderson, 1982, p. 153-155).

The normal equations used are equation 3.3-20, as modified to include prior information also (see equation 3.4-13):

$$[(\underline{S}_s^r)^T \underline{V}_s^{-1} \underline{S}_s^r + (\underline{S}_p^r)^T \underline{U}^{-1} \underline{S}_p^r s^2$$

$$+ \mu \underline{I}]\underline{\delta}_{r+1} = (\underline{S}_s^r)^T \underline{V}_s^{-1}(\underline{Y}_s - \underline{f}_s(\underline{\xi},\underline{b}_r))$$

$$+ (\underline{S}_p^r)^T \underline{U}^{-1} s^2(\underline{Y}_p - \underline{f}_p(\underline{\xi},\underline{b}_r)) \qquad (4.1-11)$$

where subscripts s and p refer to sample and prior information, respectively;

$\underline{S}_p^r =$ the matrix $[\underline{I}_{n_p}, \underline{0}]$ from equation 4.1-6, transformed using equation 3.3-16;

$\underline{f}_p(\underline{\xi},\underline{b}_r) =$ a vector composed of the rth estimate of those parameters on which there is prior information; and

$s^2 =$ the ordinary least-squares estimate of σ^2 (to be developed later).

4.2 Singularity and Conditioning

Singularity of the least-squares coefficient matrix can occur whenever (1) no measured flow rates (such as well or spring discharges) are in the model and (2) an attempt is made to compute all parameters. To understand how this occurs, consider first the case where there are no specified head parameters and no prior information, but all other parameters are to be computed. Also, assume for simplicity that all observation points correspond to node points. In this case it can be shown (appendix, section 4.3.3) that

$$\underline{J}\underline{b} = \underline{0} \qquad (4.2-1)$$

where subscripts r were omitted to simplify nomenclature, and $\underline{J} = \{J_j\} = \{\partial q/\partial b_j - (\partial \underline{D}/\partial b_j)\underline{h}\}$. By using equation 3.3-23,

$$J_j = \frac{\partial q}{\partial b_j} - \frac{\partial \underline{D}}{\partial b_j}\underline{h} = \underline{D}\frac{\partial \underline{h}}{\partial b_j} \qquad (4.2-2)$$

so that

$$\underline{J}\underline{b} = \sum_{j=1}^{p} \underline{J}_j b_j = \underline{D}\sum_{j=1}^{p} \frac{\partial \underline{h}}{\partial b_j} b_j \quad . \qquad (4.2-3)$$

Because $\underline{D}$ is nonsingular,

$$\underline{D}^{-1}\underline{D} \sum_{j=1}^{p} \frac{\partial \underline{h}}{\partial b_j}\, b_j = \underline{0}$$

or, by eliminating those nodes not corresponding to observation points from $\partial \underline{h}/\partial b_j$,

$$\sum_{j=1}^{p} \underline{X}_{sj} b_j = \underline{0} \ . \qquad (4.2\text{-}4)$$

Equation 4.2–4 can also be written in the form

$$\underline{X}_s \underline{c} = \underline{0} \qquad (4.2\text{-}5)$$

where

$$\underline{c} = \underline{b} = [b_1, b_2, ..., b_p]^T \ . \qquad (4.2\text{-}6)$$

Recall that equation 4.2–5 implies that the least-squares coefficient matrix is singular.

If a known flow rate Q_i is at node i, then equation 4.2–1 becomes

$$\underline{J}\underline{b} = \underline{Q} \qquad (4.2\text{-}7)$$

where

$$\underline{Q} = [0, 0, ..., Q_i, ..., 0]^T_{(m)} \qquad (4.2\text{-}8)$$

so that, all other things being equal,

$$\underline{X}_s \underline{c} \neq \underline{0} \ . \qquad (4.2\text{-}9)$$

In the case where at least one parameter j is fixed, then $\underline{J}\underline{b}$ has column j of $\underline{J}$ and element j of $\underline{b}$ deleted. Thus, equation 4.2–1 no longer holds, so that equation 4.2–9 will hold, if no other source of singularity exists.

Whenever there are specified head parameters, $\underline{J}$ and $\underline{X}_s$ both contain columns resulting from these parameters, and $\underline{J}\underline{b} \neq \underline{0}$. However, for those columns not involving the specified head parameters, $\sum_j \underline{J}_j b_j = \underline{0}$, where j denotes all parameters except specified head parameters. If

$c_j = 0$ for those columns resulting from specified head parameters and $c_j = b_j$ for the remaining columns, then equation 4.2–5 holds for the entire sensitivity matrix, which indicates that the problem is again singular. Because addition of specified head parameters has no influence on this type of singularity, it is assumed for simplicity for the remainder of this section that there are no specified head parameters.

Singularity caused by attempting to find all parameters in the absence of known flow rates can be rectified by using prior information. For the case of prior information, equation 4.2–5 can be written

$$\begin{bmatrix} \underline{X}_s \\ \underline{X}_p \end{bmatrix} \underline{c} = \underline{0} \ . \qquad (4.2\text{-}10)$$

The only way for equation 4.2–10 to hold is if $\underline{X}_s \underline{c} = \underline{0}$ and $\underline{X}_p \underline{c} = \underline{0}$. If the only cause of the singularity is given by equation 4.2–1, then $\underline{c} = \underline{b}$ is the only linearly independent solution of equation 4.2–10, and, if $\underline{X}_p$ is derived from equation 4.1–6, $X_{p_{ii}} = 0$ ($i = 1, 2, ..., p$) is the only way that equation 4.2–10 can hold. Hence, prior information on any parameter can theoretically condition the problem so that all parameters can be found.

The maximum number of parameters that can be found for any problem can also be obtained through nondimensionalization (or partial nondimensionalization) to find the smallest number of independent groups. In addition, nondimensionalization also illustrates the idea that solution is actually often best expressed in terms of ratios of the parameters. As an example, consider the case where a region is composed of two zones where $T_1, T_2, W_1,$ and W_2 are parameters. Then the flow equations for each zone are

$$\left. \begin{aligned} T_1\left|\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right| + W_1 = 0 \\[2em] T_2\left|\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right| + W_2 = 0 \end{aligned} \right\} \qquad (4.2\text{-}11)$$

and the boundary conditions between zones are

$$T_1\left|\frac{\partial h}{\partial n}\right|_1 = T_2\left|\frac{\partial h}{\partial n}\right|_2$$

$$(h)_1 = (h)_2$$

(4.2-12)

where the notation $(\cdot)_1$ indicates that the quantity in parentheses is evaluated just within the 1 side of the boundary and similarly for $(\cdot)_2$. If equations 4.2-11 and 4.2-12 are written in the alternative forms

$$\frac{T_1}{T_2}\left|\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right| + \frac{W_1}{T_2} = 0$$

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{W_2}{T_2} = 0$$

(4.2-13)

$$\frac{T_1}{T_2}\left|\frac{\partial h}{\partial n}\right|_1 = \left|\frac{\partial h}{\partial n}\right|_2$$

$$(h)_1 = (h)_2 \,,$$

(4.2-14)

instead of four independent parameters, there are only three written as ratios: $W_1/T_2, W_2/T_2, T_1/T_2$. A known flow rate in zone 1, Q_1, would add the term Q_1/T_2 to the first part of equation 4.1-13. In this case W_1/T_2, W_2/T_2, T_1/T_2, and Q_1/T_2 could all be considered parameters. Knowing Q_1 (either exactly or with uncertainty) would provide unique estimates of the four original parameters. Another approach would be to find T_1, T_2, W_1, W_2 as parameters, knowing that the problem is not singular because there are four independent ratios for the problem.

Another common way for singularity to occur is if a column of $\underline{X}_s$ is zero: $\underline{X}_{sj}=\underline{0}$. This results if measurements are taken at points where there is no sensitivity to the parameter, b_j, corresponding to the column. If this is the only source of singularity, then $\underline{c}=[0,0,....,b_j, 0,....,0]^T$ is the only linearly independent solution of equation 4.2-5. In this case if $\underline{X}_p$ is derived from equation 4.1-6 and there is prior information on b_j, then $\underline{X}_p\underline{c}\neq\underline{0}$, so that the prior information solves the singularity problem. Addition

of prior information on any other parameter alone obviously will not help.

Two sources of singularity result if $\underline{X}_{sj}=\underline{0}$ and $\underline{Jb}=\underline{0}$. In this case one solution of equation 4.2-5 is, as before, given by equation 4.2-6. However, because $\underline{X}_{sj}=\underline{0}$, c_j can be any arbitrary value less than infinity and so can be set to zero. Hence, $\underline{c}=[b_1,b_2,....,0,b_{j+1},....,b_p]^T$, where the zero appears in row j of $\underline{c}$, is another solution to equation 4.2-5. Addition of prior information on parameter j alone does not solve the singularity problem because, even though X_{pjj} does not equal zero, c_j and X_{pii}, $i\neq j$, do equal zero so that $\underline{X}_p\underline{c}=\underline{0}$. A third solution of equation 4.2-5 is $\underline{c}=[0,0,....,b_j,0,....,0]^T$. In this instance addition of prior information on any or all parameters except parameter j yields $\underline{X}_p\underline{c}=\underline{0}$. Thus, if $\underline{X}_{sj}=\underline{0}$ and $\underline{Jb}=\underline{0}$, then the problem is singular unless prior information is added on parameter j and at least one other parameter.

If the columns of $\underline{X}_s$ are almost linearly dependent, then the problem is ill-conditioned. Thus, if either Q_i in equation 4.2-8 is almost zero or $\underline{X}_{sj}\cong\underline{0}$, then an ill-conditioned problem can result. However, ill-conditioning can occur in a number of ways. The techniques given in section 3.2.3 can be used to detect conditioning problems.

Problem 4.2-1

Solve problem 3.2-1 with the regression computer program (appendix 4.3.4). Assume that the stream tube is one foot wide and that transmissivities are unity. Place a row of nodes along each side of the stream tube, but specify observed heads only along one row or down the center of the tube (number of observed heads should be the same as in problem 3.2-1). Allow two iterations. What would happen if you were to attempt to estimate both W and T?

Problem 4.2-2

Figure 1 gives the zone map for a steady-state ground-water flow system in a hypothetical region. The finite difference mesh and types of boundary conditions also are shown on the map. Use the regression program (appendix 4.3.4) to construct a regression flow model for the region.
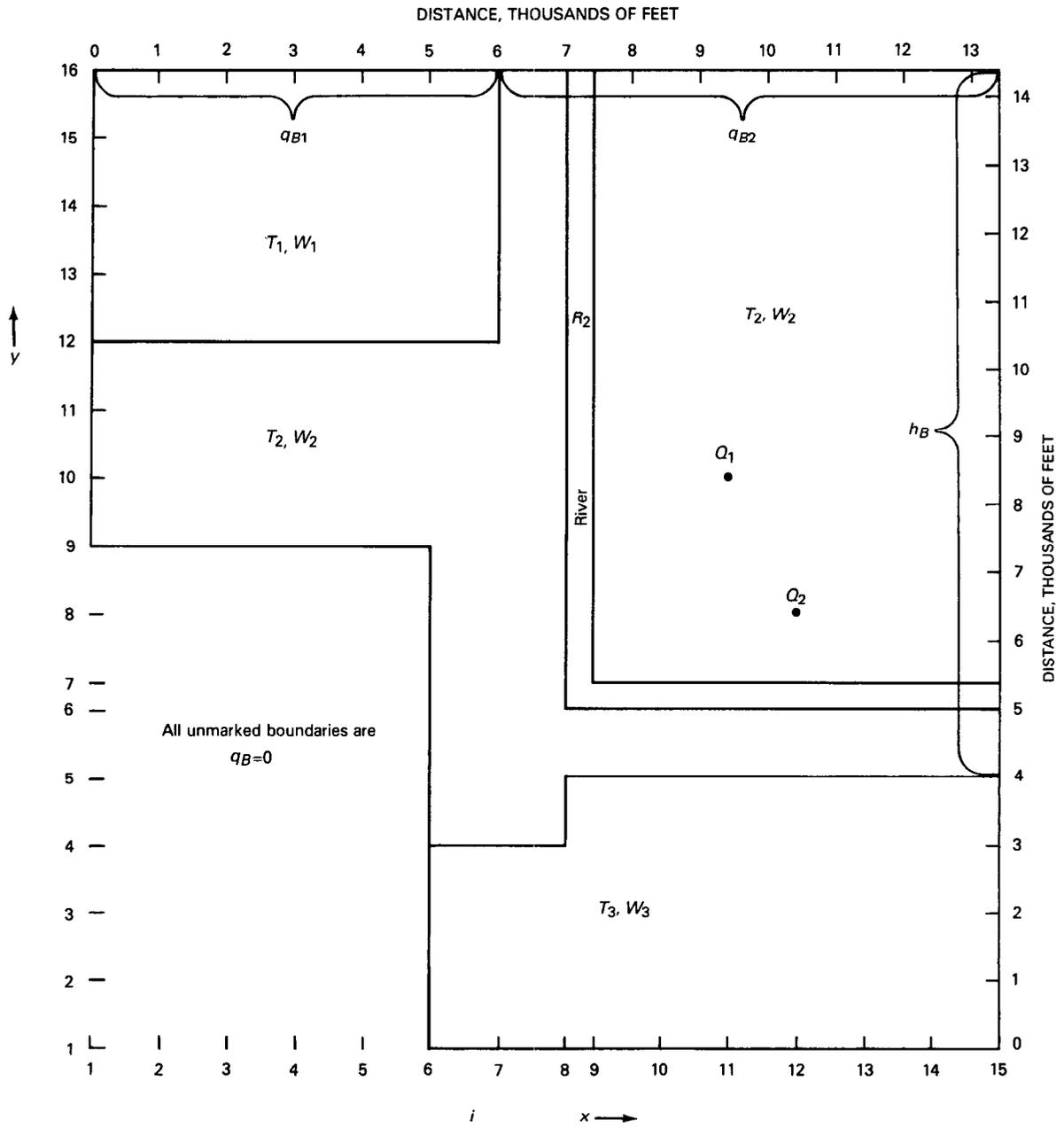
Figure 1

Prior information exists on the following parameters:

	standard deviation=84
$T_2=420$ ft^2/d	
$W_1=0.0004$ ft/d	0.00012
$W_3=0.00017$ ft/d	0.000051
$R_2=0.08$ d^{-1}	0.008
$Q_1=-97,000$ ft^3/d	1940
$Q_2=-51,000$ ft^3/d	1020
h_B at (15,16)=10.4 ft	1.04
h_B at (15,7)=4.8 ft ⎱ one parameter	0.48
h_B at (15,6)=4.8 ft ⎰	0.48
h_B at (15,5)=5.4 ft	0.54

Assume that h_B varies linearly between the estimated values. For one reason or another, the estimates of h_B at the four nodes are not observations. (They may have been interpolated from a contour map, for example.)

There is no prior information on the remaining parameters, but probable limits of variation for these parameters are

$$30 \leq T_1 \leq 80$$
$$10 \leq T_3 \leq 40$$
$$-0.0003 < W_2 \leq -0.00005$$
$$0.2 \leq q_{B1} \leq 0.8$$
$$0.15 \leq q_{B2} \leq 0.4$$

From these ranges, initial estimates of the parameters may be determined.

The observed head data in table 1 were collected. They are of uniform reliability.

Table 1.

Node	Value (ft)	Node	Value (ft)
(8,2)	60.70	(7,11)	6.68
(14,2)	75.64	(13,11)	-15.32
(12,3)	60.27	(3,12)	16.88
(10,4)	29.67	(5,12)	15.87
(7,5)	4.22	(9,12)	4.48
(11,5)	4.37	(11,12)	-18.34
(13,5)	6.07	(13,13)	-2.47
(15,5)	5.81	(15,13)	8.10
(10,7)	4.57	(3,14)	54.12
(8,8)	5.21	(5,14)	38.27
(12,8)	-44.89	(10,14)	0.053
(15,8)	7.01	(12,14)	-2.92
(7,9)	6.95	(7,15)	8.30
(4,10)	12.21	(14,15)	4.54
(9,10)	4.04	(2,16)	85.82
(11,10)	-89.36	(11,16)	2.26

The river stage is about 4.5 ft everywhere.

Assuming that $\sigma^2=1$, find all possible parameters for the model. First, however, determine how many parameters you can find!

Examine the sensitivity maps. Are there data in relatively high sensitivity areas for all parameters? Do you think that there are places where new data points would improve the results?

4.3 Appendices

4.3.1 Integrated Finite Difference Model

The numerical solution of equation 4.1-1 is obtained by using integrated finite difference methods. A rectangular grid of nodes is assumed as indicated in figure 4.3-1. Each node point is enclosed by a subdomain, which is a rectangular region bounded by sides located half way between adjacent node points.

The coordinates of a typical node (i,j) are given as (x_i, y_j). With the nomenclature shown in figure 4.3-1, equation 4.1-1 can be integrated over a subdomain enclosing node (i,j) to produce

$$\int_{\Delta y_j} (T_{xx} \frac{\partial h}{\partial x})_{i+1/2} dy - \int_{\Delta y_j} (T_{xx} \frac{\partial h}{\partial x})_{i-1/2} dy$$

$$+ \int_{\Delta x_i} (T_{yy} \frac{\partial h}{\partial y})_{j+1/2} dx - \int_{\Delta x_i} (T_{yy} \frac{\partial h}{\partial y})_{j-1/2} dx$$

$$+ \int_{\Delta x_i} \int_{\Delta y_j} R(H-h)dxdy$$

$$+ \int_{\Delta x_i} \int_{\Delta y_j} Wdxdy + \sum_{p=1}^{N_{i,j}} Q_p = 0 \qquad (4.3-1)$$

where

$$\left. \begin{array}{l} \Delta x_i = \frac{1}{2}(\Delta x_{i+1/2} + \Delta x_{i-1/2}) \\ \\ \Delta y_j = \frac{1}{2}(\Delta y_{j+1/2} + \Delta y_{j-1/2}) \end{array} \right\} \qquad (4.3-2)$$

and $N_{i,j}$ is the number of pumping wells in subdomain $\Delta x_i \Delta y_j$.

If $T_{\xi\xi}$, R, and W are assumed to be constant in each cell (A,B,C,D) adjacent to node (i,j), then a valid numerical approximation of equation 4.3-1 is
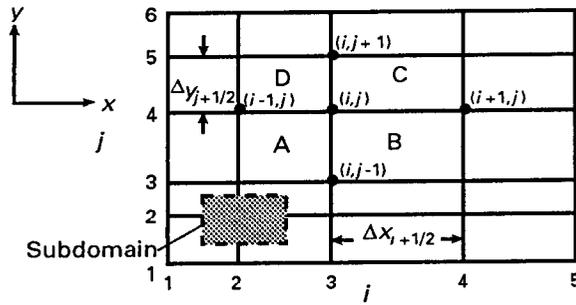
Figure 4.3-1

$$T_{xxi+\frac{1}{2}j}\Delta y_j \frac{h_{i+1,j}-h_{i,j}}{\Delta x_{i+\frac{1}{2}}}$$

$$-T_{xxi-\frac{1}{2}j}\Delta y_j \frac{h_{i,j}-h_{i-1,j}}{\Delta x_{i-\frac{1}{2}}}$$

$$+T_{yyi,j+\frac{1}{2}}\Delta x_i \frac{h_{i,j+1}-h_{i,j}}{\Delta y_{j+\frac{1}{2}}}$$

$$-T_{yyi,j-\frac{1}{2}}\Delta x_i \frac{h_{i,j}-h_{i,j-1}}{\Delta y_{j-\frac{1}{2}}}$$

$$+R_{i,j}\Delta x_i \Delta y_j(H_{i,j}-h_{i,j})$$

$$+W_{i,j}\Delta x_i \Delta y_j+Q_{i,j}=0 \qquad (4.3-3)$$

where

$$T_{xxi+\frac{1}{2}j}=\frac{\Delta y_{j-\frac{1}{2}}T_{xxB}+\Delta y_{j+\frac{1}{2}}T_{xxC}}{2\Delta y_j} \qquad (4.3-4)$$

$$T_{xxi-\frac{1}{2}j}=\frac{\Delta y_{j-\frac{1}{2}}T_{xxA}+\Delta y_{j+\frac{1}{2}}T_{xxD}}{2\Delta y_j} \qquad (4.3-5)$$

$$T_{yyi,j+\frac{1}{2}}=\frac{\Delta x_{i-\frac{1}{2}}T_{yyD}+\Delta x_{i+\frac{1}{2}}T_{yyC}}{2\Delta x_i} \qquad (4.3-6)$$

$$T_{yyi,j-\frac{1}{2}}=\frac{\Delta x_{i-\frac{1}{2}}T_{yyA}+\Delta x_{i+\frac{1}{2}}T_{yyB}}{2\Delta x_i} \qquad (4.3-7)$$

$$R_{i,j}=\frac{\Delta x_{i-\frac{1}{2}}\Delta y_{j-\frac{1}{2}}R_A+\Delta x_{i+\frac{1}{2}}\Delta y_{j-\frac{1}{2}}R_B+\Delta x_{i+\frac{1}{2}}\Delta y_{j+\frac{1}{2}}R_C+\Delta x_{i-\frac{1}{2}}\Delta y_{j+\frac{1}{2}}R_D}{4\Delta x_i\Delta y_j} \qquad (4.3-8)$$

$$W_{i,j}=\frac{\Delta x_{i-\frac{1}{2}}\Delta y_{j-\frac{1}{2}}W_A+\Delta x_{i+\frac{1}{2}}\Delta y_{j-\frac{1}{2}}W_B+\Delta x_{i+\frac{1}{2}}\Delta y_{j+\frac{1}{2}}W_C+\Delta x_{i-\frac{1}{2}}\Delta y_{j+\frac{1}{2}}W_D}{4\Delta x_i\Delta y_j} \qquad (4.3-9)$$

$$Q_{i,j}=\sum_{p=1}^{N_{i,j}}Q_p \ . \qquad (4.3-10)$$

Because of the way that the cells are designated, all zone boundaries are assumed to pass through node points; for example, see figure 4.3-2.

If the node points in equation 4.3-3 are designated as

$$k=i+NC\cdot(j-1) \qquad (4.3-11)$$

where NC is the number of columns (in the i direction), then the grid is renumbered as in figure 4.3-3. Equation 4.3-3 then becomes

$$T_{xxk,3}\Delta y_j \frac{h_{k+1}-h_k}{\Delta x_{i+\frac{1}{2}}}-T_{xxk,1}\Delta y_j \frac{h_k-h_{k-1}}{\Delta x_{i-\frac{1}{2}}}$$

$$+T_{yyk,4}\Delta x_i \frac{h_{k+NC}-h_k}{\Delta y_{j+\frac{1}{2}}}-T_{yyk,2}\Delta x_i \frac{h_k-h_{k-NC}}{\Delta y_{j-\frac{1}{2}}}$$

$$+R_k\Delta x_i\Delta y_j(H_k-h_k)+W_k\Delta x_i\Delta y_j+Q_k=0 \ (4.3-12)$$

where

$$T_{xxk,1}=T_{xxi-\frac{1}{2}j}, T_{yyk,2}=T_{yyi,j-\frac{1}{2}},$$

$$T_{xxk,3}=T_{xxi+\frac{1}{2}j} \ , \quad \text{and}$$

$$T_{yyk,4}=T_{yyi,j+\frac{1}{2}} \ .$$

In matrix form the numerical solution is

$$\underline{D}h=q \qquad (4.3-13)$$
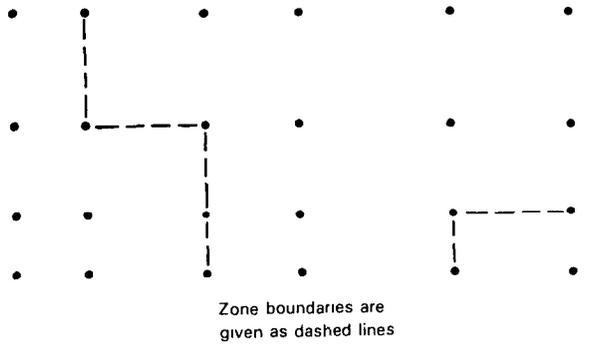
Zone boundaries are
given as dashed lines

Figure 4.3-2

Figure 4.3-3

$$D_{k,k+1} = -T_{xxk,3} \frac{\Delta y_j}{\Delta x_{i+\frac{1}{2}}} \quad , \qquad (4.3\text{-}17)$$

$$D_{k,k+NC} = -T_{yyk,4} \frac{\Delta x_i}{\Delta y_{j+\frac{1}{2}}} \quad , \qquad (4.3\text{-}18)$$

$$q_k = R_k \Delta x_i \Delta y_j H_k + W_k \Delta x_i \Delta y_j + Q_k \quad . \quad (4.3\text{-}19)$$

For node k on a specified head boundary, $D_{k,k-NC} = D_{k,k-1} = D_{k,k+1} = D_{k,k+NC} = 0, D_{k,k} = 1,$ and $q_k = h_{Bk}$, the specified head. All remaining $D_{k,\ell} = 0$ for equation k in both cases. To preserve symmetry of $\underline{D}$, equations ℓ, $\ell \neq k$, are modified as indicated just after equation 4.1-4.

The flow across specified flow boundaries is incorporated by using the Q_k term, so that the total flow crossing the specified flow boundary of the subdomain around node k is added into Q_k. If $Q_k = 0$ on a boundary node and the head at the node is not specified, then the boundary for the node is automatically a no-flow type. When computing the total flow to add into a specified flow node, remember that nodes are on boundaries so that subdomains for boundary nodes are only fractions of the full subdomains. For example, see figure 4.3-4.

4.3.2 Computation of Nodal Sensitivities for the Integrated Finite Difference Model

Partial derivatives J_ℓ, defined by

$$J_\ell = \frac{\partial q}{\partial b_\ell} - \frac{\partial \underline{D}}{\partial b_\ell} h, \quad \ell = 1, 2, \dots, p \qquad (4.3\text{-}20)$$

where, from equation 4.3-12 for node k not on a specified head boundary,

$$D_{k,k-NC} = -T_{yyk,2} \frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}} \quad , \qquad (4.3\text{-}14)$$

$$D_{k,k-1} = -T_{xxk,1} \frac{\Delta y_j}{\Delta x_{i-\frac{1}{2}}} \quad , \qquad (4.3\text{-}15)$$

$$D_{k,k} = T_{xxk,3} \frac{\Delta y_j}{\Delta x_{i+\frac{1}{2}}} + T_{yyk,4} \frac{\Delta x_i}{\Delta y_{j+\frac{1}{2}}}$$

$$+ T_{xxk,1} \frac{\Delta y_j}{\Delta x_{i-\frac{1}{2}}} + T_{yyk,2} \frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}}$$

$$+ R_k \Delta x_i \Delta y_j \quad , \qquad (4.3\text{-}16)$$

Dashed lines
enclose
subdomains
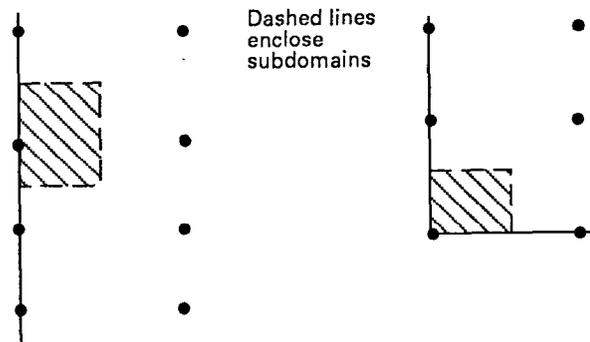
Figure 4.3-4

are employed in equation 3.3-24 to compute sensitivities for the nonlinear regression solution of the numerical model given in section 4.3.1. Using the definitions of the elements of $\underline{D}$ and $\underline{q}$ given in section 4.3.1, partial derivatives for the various parameters are computed as follows:

1. $b_\ell = T_{xxA}$ at node k

$$\frac{\partial D_{k,k-NC}}{\partial b_\ell} = 0 \qquad (4.3-21)$$

$$\frac{\partial D_{k,k-1}}{\partial b_\ell} = -\frac{\Delta y_j}{\Delta x_{i-\frac{1}{2}}} \frac{\partial}{\partial T_{xxA}}$$

$$\cdot \left(\frac{\Delta y_{j-\frac{1}{2}} T_{xxA} + \Delta y_{j+\frac{1}{2}} T_{xxD}}{2 \Delta y_j} \right) = -\frac{\Delta y_{j-\frac{1}{2}}}{2 \Delta x_{i-\frac{1}{2}}} (4.3-22)$$

$$\frac{\partial D_{k,k}}{\partial b_\ell} = \frac{\Delta y_{j-\frac{1}{2}}}{2 \Delta x_{i-\frac{1}{2}}} \qquad (4.3-23)$$

$$\frac{\partial D_{k,k+1}}{\partial b_\ell} = 0 \qquad (4.3-24)$$

$$\frac{\partial D_{k,k+NC}}{\partial b_\ell} = 0 . \qquad (4.3-25)$$

Similar expressions result for $b_\ell = T_{xxB}, T_{xxC},$ and T_{xxD}. If more than one of the cells $A, ...,D$ lie in the same zone, then derivatives for the individual expressions are summed to form the final value. For example, if $b_\ell = T_{yyA} = T_{yyB} = T_{yyC} = T_{yyD}$, then

$$\frac{\partial D_{k,k-NC}}{\partial b_\ell} = \frac{\partial D_{k,k-NC}}{\partial T_{yyA}} + \frac{\partial D_{k,k-NC}}{\partial T_{yyB}}$$

$$+ \frac{\partial D_{k,k-NC}}{\partial T_{yyC}} + \frac{\partial D_{k,k-NC}}{\partial T_{yyD}}$$

$$= -\frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}} \left(\frac{\Delta x_{i-\frac{1}{2}}}{2 \Delta x_i} + \frac{\Delta x_{i+\frac{1}{2}}}{2 \Delta x_i} + 0 + 0 \right)$$

$$= -\frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}} . \qquad (4.3-26)$$

Derivatives for any configuration of zone boundary are handled by combinations of the form of equations 4.3-21 through 4.3-26.

An example of application of equations 4.3-21 through 4.3-26 for an irregular zone boundary is given in figure 4.3-5. Let $T_{xx} = T_{yy} = T$ for simplicity. Then, for a two-zone problem, the two transmissivities are T_1 and T_2.

For node 7:

$$\frac{\partial D_{7,2}}{\partial T_1} = -\frac{\Delta x_2}{\Delta y_{1\frac{1}{2}}} , \frac{\partial D_{7,6}}{\partial T_1} = -\frac{\Delta y_2}{\Delta x_{1\frac{1}{2}}} , \frac{\partial D_{7,7}}{\partial T_1}$$

$$= \frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}} + \frac{\Delta x_2}{\Delta y_{2\frac{1}{2}}} + \frac{\Delta y_2}{\Delta x_{1\frac{1}{2}}} + \frac{\Delta x_2}{\Delta y_{1\frac{1}{2}}} ,$$

$$\frac{\partial D_{7,8}}{\partial T_1} = -\frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}} , \frac{\partial D_{7,12}}{\partial T_1} = -\frac{\Delta x_2}{\Delta y_{2\frac{1}{2}}} .$$

For node 7 the derivatives of the D_{ij}'s with respect to T_2 are all zero because T_2 does not appear in any of the D_{ij}'s. For node 8:

$$\frac{\partial D_{8,3}}{\partial T_1} = -\frac{\Delta x_3}{\Delta y_{1\frac{1}{2}}} , \frac{\partial D_{8,7}}{\partial T_1} = -\frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}} , \frac{\partial D_{8,8}}{\partial T_1}$$

$$= \frac{\Delta y_{1\frac{1}{2}}}{2 \Delta x_{3\frac{1}{2}}} + \frac{\Delta x_{2\frac{1}{2}}}{2 \Delta y_{2\frac{1}{2}}} + \frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}} + \frac{\Delta x_3}{\Delta y_{1\frac{1}{2}}} ,$$

$$\frac{\partial D_{8,9}}{\partial T_1} = -\frac{\Delta y_{1\frac{1}{2}}}{2 \Delta x_{3\frac{1}{2}}} , \frac{\partial D_{8,13}}{\partial T_1} = -\frac{\Delta x_{2\frac{1}{2}}}{2 \Delta y_{2\frac{1}{2}}}$$

$$\frac{\partial D_{8,3}}{\partial T_2} = \frac{\partial D_{8,7}}{\partial T_2} = 0, \frac{\partial D_{8,8}}{\partial T_2} = \frac{\Delta y_{2\frac{1}{2}}}{2 \Delta x_{3\frac{1}{2}}} + \frac{\Delta x_{3\frac{1}{2}}}{2 \Delta y_{2\frac{1}{2}}}$$

$$\frac{\partial D_{8,9}}{\partial T_2} = -\frac{\Delta y_{2\frac{1}{2}}}{2 \Delta x_{3\frac{1}{2}}} , \frac{\partial D_{8,13}}{\partial T_2} = -\frac{\Delta x_{3\frac{1}{2}}}{2 \Delta y_{2\frac{1}{2}}}$$

Figure 4.3-5

2. $b_\ell = R_A$ at node k

$$\frac{\partial D_{k,k-NC}}{\partial b_\ell} = \frac{\partial D_{k,k-1}}{\partial b_\ell} = \frac{\partial D_{k,k+1}}{\partial b_\ell}$$

$$= \frac{\partial D_{k,k+NC}}{\partial b_\lambda} = 0 \qquad (4.3\text{-}27)$$

$$\frac{\partial D_{k,k}}{\partial b_\ell} = \Delta x_i \Delta y_j \left| \frac{\Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}}}{4\Delta x_i \Delta y_j} \right| = \frac{1}{4} \Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}}$$

$$(4.3\text{-}28)$$

$$\frac{\partial q_k}{\partial b_\ell} = \frac{1}{4} \Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} H_k . \qquad (4.3\text{-}29)$$

Derivatives for R_B, R_C, and R_D are similar.

3. $b_\ell = W_A$ at node k

$$\frac{\partial q_k}{\partial b_\ell} = \frac{1}{4} \Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} . \qquad (4.3\text{-}30)$$

All derivatives of D_{ij} are zero.

4. $b_\ell = q_{B1}$, boundary flux 1 in Q_k, where, for example, $Q_k = q_{B2} \cdot \frac{1}{2} \Delta y_{j-\frac{1}{2}} + q_{B1} \cdot \frac{1}{2} \Delta y_{j+\frac{1}{2}}$ (see figure 4.3-6).

$$\frac{\partial q_k}{\partial b_\ell} = \frac{1}{2} \Delta y_{j+\frac{1}{2}} . \qquad (4.3\text{-}31)$$

A similar expression results for $b_\ell = q_{B2}$, and if $q_{B1} = q_{B2}$, the derivative is the sum of two parts, of equation 4.3-31 and its analog for q_{B2}.

5. $b_\ell = H_s$ or H_t

The specified head at any node k along a specified head boundary is given as

$$h_{Bk} = A_k[L_k H_t + (1-L_k)H_s] \qquad (4.3\text{-}32)$$

where

$$A_k = \frac{h_{Bk}^0}{L_k H_t^0 + (1-L_k)H_s^0} \qquad (4.3\text{-}33)$$
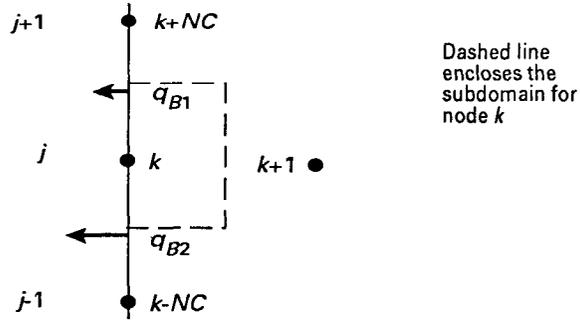
Figure 4.3-6

and the meanings of the symbols are defined after equation 4.1-5.

If node m (here only m indicates an arbitrary node number) is adjacent to a boundary segment bounded by nodes s and t, and node k lies in the segment so that it appears in equation m, then for $b_\ell = H_s$,

$$\frac{\partial q_m}{\partial b_\ell} = -D_{m,k} \frac{\partial h_{Bk}}{\partial H_s} = -D_{m,k} A_k (1-L_k) . \quad (4.3\text{-}34)$$

Similarly, for $b_\ell = H_t$,

$$\frac{\partial q_m}{\partial b_\ell} = -D_{m,k} \frac{\partial h_{Bk}}{\partial H_t} = -D_{m,k} A_k L_k . \qquad (4.3\text{-}35)$$

If node k lies on the boundary, then the equation for node k in equation 4.3-13 becomes

$$h_k = h_{Bk}$$

$$= A_k[L_k H_t + (1-L_k)H_s] \qquad (4.3\text{-}36)$$

and, for $b_\ell = H_s$,

$$\frac{\partial q_k}{\partial b_\ell} = A_k (1-L_k) \qquad (4.3\text{-}37)$$

and similarly for $b_\ell = H_t$.

4.3.3 Derivation of Equation 4.2-1

By careful examination of equations 4.3-12 through 4.3-19 it can be seen that, if there are

no specified head parameters, $q-\underline{Dh}=\underline{0}$ can be written in the form

$$a_{i1}b_1+a_{i2}b_2+...+a_{ip}b_p-Q_i=0,$$
$$i=1, 2,...,m \qquad (4.3\text{-}38)$$

where

> a_{ij}=coefficient containing Δx, Δy, and head differences;
>
> b_j=any parameter except a specified head parameter; and
>
> Q_i=term not containing parameters in $\underline{b}$.

Define

$$J_j=\frac{\partial q}{\partial b_j} - \frac{\partial \underline{D}}{\partial b_j}\,\underline{h} \ . \qquad (4.3\text{-}39)$$

Then, by carrying out the differentiations indicated in equation 4.3–39 and comparing the result with equation 4.3–38 it can be seen that

$$J_{ij}b_j=a_{ij}b_j \qquad (4.3\text{-}40)$$

so that

$$\sum_{j=1}^{P}J_{ij}b_j=\sum_{j=1}^{P}a_{ij}b_j=Q_i \ . \qquad (4.3\text{-}41)$$

If $\underline{b}$ contains all possible parameters (except specified head parameters) and there are no known fluxes, then $Q_i=0$ and

$$\sum_{j=1}^{p}J_{ij}b_j=0 \qquad (4.3\text{-}42)$$

or

$$\underline{Jb}=\underline{0} \ . \qquad (4.3\text{-}43)$$

4.3.4 Documentation of Program for Nonlinear Regression Solution of Steady-State Ground-Water Flow Problems

Introduction.—This program is designed to obtain a nonlinear regression solution to the finite-difference model of steady-state ground-water flow given in section 4.3.1. Basic calculation methods are given in sections 4.1 and 4.3.2.

The program was developed using the Microsoft[1] Fortran Compiler, Version 3.3, with the DOS[1] 2.0 operating system on an IBM[1] PC/XT computer with the IBM[1] 8088 Math Coprocessor and 256 KB memory. Except for the OPEN statements near the beginning of the code, Fortran 66 was used throughout to make the code as machine independent as possible. The source code is contained in files INVFD.FOR and INVSUB.FOR in the 5¼ in. diskette accompanying this report. These two files must be linked or compiled together.

The computer program is composed of a main program and eight subroutines. The main program controls input-output and performs all computations that cannot be accomplished more effectively with subroutines. The eight subroutines (D4SOLV, COEF, LSTSQ, PRTOT, ORDER, ARRAY, ARRAYI, HOBS) perform the following specialized tasks:

D4SOLV   Obtains an LDU factorization solution of the set of linear algebraic equations resulting from application of the finite difference methods, assuming the equations are ordered in an alternating diagonal fashion (Price and Coats, 1974).

COEF   Computes coefficients necessary for the determination of sensitivities and heads.

LSTSQ   Computes the coefficients of the normal equations and solves the system of equations to determine the vectors of parameter changes and parameters.

PRTOT   Prints matrices or vectors in a column configuration.

ORDER   Computes equation numbers at grid points corresponding to the alternating diagonal ordering scheme.

ARRAY   Reads and (or) prints 1- and 2-dimensional real array variables.

ARRAYI   Reads and (or) prints 1- and 2-dimensional integer array variables.

---

[1]Use of the trade names in this report is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.

HOBS Reads and prints observed heads and weighting values; computes coefficients for bilinear interpolation of computed heads and sensitivities at observation points.

The basic flow of the program can be described as follows:

A. Data are input and variables are initialized.

B. Using coefficients generated in COEF, an initial solution corresponding to the initial parameter estimates is computed by D4SOLV.

C. In an iterative fashion, the following four steps are taken until the regression technique converges or until the number of iterations exceeds the maximum allowed.

(1) Sensitivities are calculated using coefficients computed in COEF and in the main program.

(2) LSTSQ is employed to form and solve the normal equations.

(3) Parameters are updated in LSTSQ using the parameter change vector generated.

(4) Various coefficients involving the updated parameters are computed in COEF, and current estimates of head are computed using D4SOLV.

D. Various statistics associated with the regression analysis are computed.

Aquifer Property Zonation and Variable Definition.—Basic model geometry is defined by the finite difference grid that is constructed over the region to be modeled. Nodes, consisting of grid intersections, are numbered from the lower left-hand corner of the grid (columns from left to right and rows from bottom to top). Cells, consisting of intragrid areas bounded by four adjacent nodes, are numbered similarly (figure 4.3–7).

The finite difference grid is divided into aquifer property zones, which define zonal values for the aquifer properties T_{xx}, T_{yy}, R, and W. Each zonal value is constant within the zone. Variation of a property within a zone is accomplished by assigning cell values. The aquifer property at any particular cell is computed as the product of the zonal value and the cell value of the property. Thus, if all cell values for a property within a zone are given a value of unity, the zonal value becomes the value of the property for each cell within that zone.
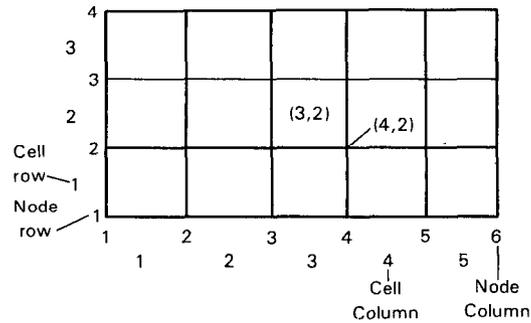
Figure 4.3–7

Zones are created by subdividing the grid into groups of cells having distinct combinations of zonal values. Cells belonging to these zones are accordingly assigned distinct zone numbers (IZN), reserving IZN=0 to indicate groups of cells outside of the model area.

All zonal values are either regarded as regression parameters to be determined by the procedure or are held constant and, thus, are not regarded as regression parameters to be determined, as specified in the input. From a conceptual viewpoint there is no difference between these two designations because zonal values that are held constant can be regarded as regression parameters having exact prior information. However, from a computational viewpoint it is most efficient to eliminate parameters being held constant from the calculations. Hence, these types of parameters do not appear in the normal equations or any of the vectors and matrices derived from them. To simplify nomenclature, in subsequent discussions the term regression parameter refers only to those regression parameters to be determined.

Cell values for an aquifer property are input using rectangular blocks of cells, which are defined for each property for convenience of input only. These blocks need not bear any relationship to the zones. Cell values may be constant or variable within each block. Cell values of a property are unaffected by the regression procedure.

Boundary Conditions and Boundary Parameters.—Two types of boundary conditions may be used: specified flow and (or) specified head. Nonzero specified-flow boundaries where the flow rate is known can be imposed

by assigning the appropriate value of the specified flow rate to the nodal value of variable WELL. Because the no-flow boundary is the default condition, zeros do not have to be assigned to WELL to simulate this condition. Specified-head boundaries where the head is known can be imposed by assigning the value of the known head at the boundary nodes to variable HC, which also describes nodal values of computed head. A negative one must be assigned to variable IN corresponding to each node that is to be considered a specified head node. Segments of the boundary that will be considered as either flow or head regression parameters must not be entered into WELL or HC arrays.

Different variables are used if specified-flow or specified-head boundary conditions are to be considered as regression parameters. Variable QB is used to indicate the zonal value for discharge across groups of nodes that form a specified-flow segment (or zone). The specified volumetric flow per unit width for each cell boundary within a zone is the product of QB for that zone and a multiplier for the cell boundary. By restricting a flow-boundary zone to a single node, point recharge and discharge can be simulated. QB can be a regression parameter and, therefore, can be modified by the regression procedure; the multiplier is unaffected. If the variable IP that specifies the regression parameter number for the boundary zone is set to zero, then QB is held constant, and the segment is treated as a known-flow boundary, thus giving two possible ways (via WELL and via QB) to designate known-flow boundaries.

Specified-head boundary nodes that are to be considered regression parameters are defined by segments composed of a sequence of nodes (variables ILOC and JLOC) along portions of the boundary. The specified heads at the first and last nodes in the sequence can be either different regression parameters or a single regression parameter or held constant, depending on the nature of the problem. Adjustments to heads at these nodes computed by the regression procedure are apportioned to other nodes in the sequence. The proportion is the ratio of the distance (along the sequence of nodes) between the end node and the node of interest, to the distance between the two end nodes. These

factors (PLA and PLB) are computed by the program for a given segment. As in the case with flow-boundary parameters, if the regression parameter number is set to zero, then that parameter is held constant. If the parameter numbers at both ends of the segment are set to zero, then the segment is treated as a known-head boundary, thus giving two possible ways to designate known-head boundaries. A negative one must be entered into IN for all nodes on the specified-head boundary, whether or not the boundary involves regression parameters.

The definition of some of the more important variables related to aquifer properties and boundary conditions in the computer program are given below.

Variable name	Definition
PAR(1), PAR(2), PAR(3), PAR(4)	Zonal value for transmissivity $(T_{xx}$ and $T_{yy})$, hydraulic conductance (R), and distributed recharge (W), respectively.
IZN	An integer array that indicates the zone number of each cell.
CX, CY, VL, QR ..	Cell values for x-direction transmissivity, y-direction transmissivity, hydraulic conductance, and distributed recharge, respectively.
HR	Nodal values of head on the boundary of the confining bed opposite the aquifer.
WELL	Nodal (or point) values of known volumetric discharge (or recharge) from a well or other known-rate source-sink phenomena.
QBF	Cell-boundary multipliers for specified flow.
PLA, PLB	Arrays, which give the proportional distances from any point to either end of a segment where the specified heads at one (or both) end(s) of the segment is a (or are) regression parameter(s).

Numbering of Regression Parameters.—The three types of regression parameters (aquifer parameters, specified-flow parameters, and specified-head parameters) are numbered consecutively and in any order. Also, any of the zonal values for aquifer properties or specified-flows and segment-end values for specified head zones can be given a single regression parameter number. In this case the zonal or segment-end values sharing the same number must have

identical regression parameter values. Any differences in property values must be specified by differences in multipliers. A common example is to use only one regression parameter for both the x- and y-direction zonal transmissivity values and to fix the degree of anisotropy for the zone by using the CX and CY arrays. As another example, if good estimates of the hydraulic gradient across cell boundaries forming a specified-flow boundary for an aquifer property zone were available, and if transmissivity were isotropic, then these gradient values could be used as multipliers so that the regression parameter QB for the specified-flow zone comprising the specified-flow boundary would be transmissivity. Thus, the x- and y-direction transmissivities for the aquifer property zone and QB for the specified-flow zone could all be the same regression parameter. A high degree of flexibility for distributing aquifer properties and boundary conditions while, at the same time, keeping the number of regression parameters to a minimum is achieved with these types of schemes.

Prior Information on Regression Parameters.—If estimates of the regression parameters and their (less than infinite) reliability are available from other sources (for example, aquifer tests), it is desirable to introduce this information into the regression analysis. For this case, initial values for the parameters are taken to be the prior information. The reliability of each estimate is represented by a standard deviation. Array WP is used to store these values for the regression parameters. However, only the standard deviations for the aquifer regression parameters are read directly into WP. The standard deviations for boundary regression parameters are read in through temporary variables (SDQB for specified-flow regression parameters, and SDHA and SDHB for specified-head regression parameters) and are only subsequently placed into the WP array.

The use of prior information of known reliability requires an estimate of the error variance of the heads (variable, EV) computed using ordinary least squares. If the estimate differs substantially from the value computed by the analysis using prior information, the problem should be resolved using the latter computed value as the estimate of error variance.

In some instances prior information of unknown reliability may be available. Use of this type of information is an advanced topic and is not covered in this report. The papers by Cooley (1982, 1983) cover the method in detail. Variables RP and BP are used to input the additional information needed for this method.

Solution-Only Mode.—To facilitate the calculation of certain statistical measures, the program is capable of bypassing the regression analysis and computing only head distributions for various combinations of parameter values. This is accomplished by specifying the solution-only option (variable ISO) and providing the various combinations of parameter values for which solutions are desired.

Using the Program.—The computer code has been designed to be as machine independent as possible. Also, to minimize confusion, all arrays have been dimensioned explicitly. The following list summarizes the minimum dimensions required for the program to operate properly for a specific problem. If

N_g is the number of node points $(N_x \times N_y)$,

N_x is the number of node columns,

N_y is the number of node rows,

N_e is the number of active (nonspecified head) nodes in the grid,

N_o is the number of observed heads,

N_{oT} is the total number of observed heads plus the number of parameters on which there is prior information,

N_z is the number of aquifer property zones in the model grid,

N_p is the number of regression parameters associated with aquifer property zones,

N_n is the number of zonal aquifer properties that are not regression parameters,

N_{hs} is the number of specified-head zones (or segments),

N_{qp} is the total number of nodes on boundaries where flow is a regression parameter,

N_{hp} is the total number of nodes along boundaries where head is composed of one or more specified-head regression parameters,

N_{mh} is the maximum number of nodes in any specified-head zone, and

N_R is the total number of regression parameters;

then the array variables should be dimensioned as in table 4.3-1.

Table 4.3-1

Variable name	Dimension
WELL,HR,HC,XV,ILOC,JLOC,IN	N_g
CX,CY,VL,QR,CXS,CYS,VLS,QRS,IZN	$(N_x-1)(N_y-1)$
DX	N_x
DY,JPOS	N_y
HCI,BK,BL,BM,BN,HO,W,KOBS	N_o
PAR,XS,KN	4
IPRM	$4,N_z$
QBF,IBNA,IBNB	N_{qp}
PLA,PLB,IBHN	N_{hp} or N_{mh}, whichever is larger.
CXHR,CXHL,CYHT,CYHB	N_{hp}
IHSN	$2N_{hp}$
IBPA, IBPB	N_{hs}
IBZN	$N_{qp}+N_{hs}$
P,WP,NCBA,NCEA,NCBF,NCEF,NCBH,NCEH	N_R
X	N_R,N_o
S	N_R,N_{oT}
A	N_R,N_R
B	N_R+N_n
V	N_e or $3N_R$, whichever is larger.
AU,IC	$5,N_e/2$
AL	$^1N_m,N_e/2$

[1]These dimensions are approximate. The exact sizes required are calculated and printed in subroutine ORDER. N_m is N_x+3 or N_y+3, whichever is smaller.

Note that array variables that have a single dimension (CX, CY, QBF, HO, etc.) and are passed to subroutines are dimensioned as unity within the subroutines (only the initial address of an array is actually passed to a subroutine). This unit dimension should not be changed in subroutines when the dimension of the variable is changed in the main program. A similar system is used for multidimensional arrays, and their dimensions within subroutines should not be changed either. To accompany any change in program dimensions, variable NVD must be set equal to the dimensions of A and the first dimension of X, and variable NAD must be set equal to the first dimension of AL. These variables are defined near the beginning of the main program.

Input Data.—The input is arranged into data sets, each data set being composed of one or more lines of logically related input data, such as cell-by-cell multipliers for x-direction transmissivity or zone-by-zone initial parameter values. Each input line in a data set is a maximum of 80 columns, or characters, long. The formats for the data applying for each line are given with the discussions of the data sets.

Data Set A.
Three title lines of user's choice (format, 20A4).
Data Set B.
Problem size information; one line (format, 16I5).

Line columns	Variable	Definition
1-5	ID	Number of node columns.
6-10	JD	Number of node rows.
11-15	NZNS	Number of aquifer property zones.
16-20	NOBS	Number of observations of head.
21-25	NPAR	Number of regression parameters associated with aquifer property zones.
26-30	NVAR	Total number of regression parameters.
31-35	NWELS	Number of known point flows.
36-40	NQBZ	Number of specified-flow boundary zones.
41-45	NHBZ	Number of specified-head boundary zones.
46-50	NUM	Maximum number of iterations allowed for the regression analysis.
51-55	IPRX	Additional print sensitivities and orthogonalize-sensitivities option. Code 1 to select the option.
56-60	IPO	Additional printout option. Code 1 to select the option.
61-65	ISO	Head-solutions only option. Code 1 to select the option.

Data Set C.
Special input parameters; one line (format, 8F10.0).

Line columns	Variable	Definition
1-10	DMX	Maximum fractional change, t_{mx}, allowed any regression parameter over any iteration.
11-20	CSA	Cosine, $\cos\Theta_{mx}$, of the maximum angle allowed between the gradient direction and the search direction (normally set to 0.08).
21-30	RP	Ridge parameter for regression analysis using prior information of unknown reliability. Code 0.0 if not used.
31-40	BP	Bias parameter for regression using prior information of unknown reliability. Code 0.0 if not used.
41-50	EV	Estimated error variance for problems using prior information of known reliability. Code 0.0 if not used.

Data Sets D through K.

A number of variables are input into the code by first subdividing the grid into rectangular regions (blocks) and then reading the variables for each block. Blocking can be applied to either cells or nodes, depending upon the variable being input. Blocking allows considerable flexibility in the input of certain variables and, once understood, can speed the construction of a model. Block and zone boundaries do not necessarily have to coincide; blocking is basically a convenient way of assigning variable values to every node or cell in the grid.

Data sets D through K represent the real (floating-point) variables subject to blocking. These variables, in the order they must appear, are listed next.

Data Set	Variable	Type of Variable	Definition
D	DX	Cell array	Distance between node points in x or I direction.
E	DY	Cell array	Distance between node points in y or J direction.
F	CX	Cell array	Multiplier (cell value) for x-direction transmissivity.
G	CY	Cell array	Multiplier (cell value) for y-direction transmissivity.
H	VL	Cell array	Multiplier (cell value) for hydraulic conductance of confining bed.
I	HR	Nodal array ...	Head on boundary of confining bed opposite the aquifer.
J	QR	Cell array	Multiplier (cell value) for recharge rate per unit area.
K	HC	Nodal array ...	Initial head at active node or fixed head at specified-head node.

Each data set D through K consists of an initial line defining the number of blocks (NOBL) into which the grid has been subdivided, and then a subsequent line or set of lines that define the blocks and the value or values of the variable to be input. The initial line, read with a A4, IX, 2I5 format, has the following form:

Line columns	Variable	Definition
1-4	NME	Array name for the variable.
6-10	NOBL	Number of rectangular input blocks into which the variable has been subdivided.
11-15	IPRN	Print option for full array. Set to 0 for print. Set to 1 for no output.

The initial line directs the program to seek NOBL blocks of information for a particular variable. If the variable is uniform over the block, then a single line suffices to define the block location and the uniform value to be assigned to every node or cell. If the variable is nonuniform over the block, then by specifying a value of IVAR equal unity on this line the program can be directed to seek additional lines specifying values of the variable for each node or cell in the block. This information is input through the format 4I5, F10.0,I5 as follows:

Line columns	Variable	Definition
1-5	IB	Beginning column of the rectangular input block.
6-10	IE	Final column of the rectangular input block.
11-15	JB	Beginning row of the rectangular input block.
16-20	JE	Final row of the rectangular input block.
21-30	FACT	If the array set is uniform for the entire block, FACT is the cell or nodal value that is assigned to each element. If the array set is not uniform, each cell or nodal value on the subsequent data lines will be multiplied by FACT.
31-35	IVAR	Code 0 if the array set is uniform. Code 1 if it is not uniform.

If a value of IVAR equal to unity is specified, then the program will seek subsequent node or cell data sufficient to define the variable at every node or cell in the block. This information is input through the temporary variable A(I,J) with an 8F10.0 format in the following manner:

Line columns	Variable	Definition
1–10	A(IB,JB)	Temporary variable specifying nodal or cell values for the
11–20	A(IB+1,JB)	arrays in data sets D through K. Note that the information
⋮	⋮	is read in row by row for the grid, each new row beginning
	A(IE,JB)	a new line.
1–10	A(IB,JB+1)	
⋮	A(IB+1,JB+1)	
	⋮	
	A(IE,JB+1)	
	⋮	
	A(IE,JE)	

Any variable that is not defined by blocking over a particular part of the grid will be automatically set to zero on that part.

Data sets D and E, representing the internal spatial dimensions of the grid, are read in by blocking for convenience to the programmer. Because both the horizontal spacing DX and vertical spacing DY (as measured from the lower left corner of the grid) are, in reality, singly dimensioned arrays, it is necessary to set JB and JE equal to unity for both variables. The variable IE then equals ID–1 in the case of DX, and JD–1 in the case of DY (IB equals one, of course, in both cases). Variable grid spacing can be input by specifying a value of IVAR equal to unity and following this with the necessary array information in an 8F10.0 format.

Note that only cell values of x- and y-direction transmissivities, hydraulic conductance, and recharge (data sets F, G, H, and J) are read in through the blocking scheme. Data set P contains the zonal values by which these cell values are multiplied. A typical example of usage would be to form transmissivity as the product of hydraulic conductivity and thickness. Data sets F and G would contain the variable thickness of the aquifer, and the variable in data set P would represent the hydraulic conductivity zone by zone. Their product would be the transmissivity.

Data Set L.

Two integer variables also are input by blocking. Both are defined below, as read in by a 16I5 format, although only that variable associated with data set L is input at this location.

Data Set	Variable	Type of Variable	Definition
L	IZN	Cell array	Zone number of each cell. Each cell having a nonzero zone number must have CX or CY>0.
S	IN	Nodal array ...	Denotes specified head. Set to –1 at nodes where head is specified, including nodes in segments involving specified-head regression parameters, and leave as zero at all remaining nodes.

The initial line for these data is identical to that of the real variable case. The line defining blocks into which integer variables are divided is similar to that of the real variable case with the exception of the variable IFACT, as noted subsequently (format, 6I5):

Line columns	Variable	Definition
1-5	IB	Beginning column of the rectangular input block.
6-10	IE	Final column of the rectangular input block.
11-15	JB	Beginning row of the rectangular input block.
16-20	JE	Final row of the rectangular input block.
21-25	IFACT	*If the array set is uniform for the entire block, IFACT is the cell or nodal value that is assigned to each grid point. If the array set is not uniform, each cell or nodal value on the subsequent data lines will <u>not</u> be multiplied by IFACT.*
26-30	IVAR	Code 0 if the array set is uniform. Code 1 if it is not uniform.

The nonuniform integer input is identical to the real variable case, except that the temporary variable $INT(I,J)$ input with format $16I5$ is used in place of $A(I,J)$.

Data set M.
    Observed head data; set contains NOBS lines (format 3I5,4F10.0).

Line columns	Variable	Definitions
1-5	N	Observation number.
6-10	IL	Cell column in which observation lies.
11-15	JL	Cell row in which observation lies.
16-25	XL	x location of observation.
26-35	YL	y location of observation.
36-45	HO(N)	Observed value of head.
46-55	W(N)	Reliability weight, ω.

Observations are numbered from 1 through NOBS but may be read in any order. Observation N is assumed to lie in cell (IL,JL) at x and y location (XL,YL). If the observation lies on a node point, it may be assigned to any adjacent cell bounded by the node. The origin for x and y is assumed to be node (1,1). <u>Omit</u> this data set if NOBS=0.

Data Set N.
    Aquifer regression parameter numbers; set contains NZNS lines (format, 16I5).

Line columns	Variable	Definition
1-5	I	Zone number.
6-10	IPRM(1,I)	Parameter number of x-transmissivity in zone I. Code 0 if it is not a regression parameter.
11-15	IPRM(2,I)	Parameter number of y-transmissivity in zone I. Code 0 if it is not a regression parameter.
16-20	IPRM(3,I)	Parameter number for hydraulic conductance in zone I. Code 0 if it is not a regression parameter.
21-25	IPRM(4,I)	Parameter number for distributed recharge in zone I. Code 0 if it is not a regression parameter.

Lines may appear in any order with respect to zone number, but there must be NZNS lines. Parameters may have any number from 1 through NVAR. Note that parameters of the same (or even different) property in different zones may have identical parameter numbers. <u>Omit</u> this data set if NPAR equals zero.

Data Set O.

Standard deviations for aquifer regression parameters; one value per line for a total of NPAR lines (format, I5,F10.0).

Line columns	Variable	Definition
1-5	K	Aquifer parameter number.
6-15	WP(K)	Standard deviation of each aquifer regression parameter. Code 0.0 if no prior information exists for the parameter.

Omit data set if NPAR equals zero.

Data Set P.

Zonal aquifer property values; set contains NZNS lines (format, I5,4F10.0).

Line columns	Variable	Definitions
1-5	I	Zone number.
6-15	PAR(1)	Zonal x-transmissivity value for zone I.
16-25	PAR(2)	Zonal y-transmissivity value for zone I.
26-35	PAR(3)	Zonal hydraulic conductance value for zone I.
36-45	PAR(4)	Zonal distributed recharge value for zone I. (PAR(4)*QR has units of volumetric rate per unit area.)

Lines may appear in any order with respect to zone number, but there must be NZNS lines.

Data Set Q.

Known point flow rates; set contains NWELS lines (format, 2I5,F10.0).

Line columns	Variable	Definition
1-5	I	Column location of point flow.
6-10	J	Row location of point flow.
11-20	WELL(I,J)	Total volumetric flow to or from node, negative for withdrawal.

If a point flow is located between node points, the total rate can be apportioned among the four adjacent nodes using bilinear or similar interpolation. The apportioned flow rates must be supplied by the user. Omit data set if NWELS equals zero.

Data Set R.

Specified boundary-flow zones and flow-zone parameters; set contains NQBZ lines, one line for each zone (format, 5I5,3F10.0).

Line columns	Variable	Definition
1-5	IA	Column location of the A end of the segment (zone).
6-10	JA	Row location of the A end of the segment (zone).
11-15	IB	Column location of the B end of segment (zone).
16-20	JB	Row location of the B end of segment (zone).
21-25	IP	Regression parameter number. Set equal to zero if QB is not a regression parameter.
26-35	QB	Zonal flow value.
36-45	SDQB ...	Standard deviation for regression parameter. Code as 0.0 if no prior information exists on the parameter.
46-55	QBM	Multiplier for zonal flow value.

Note that IA≤IB and JA≤JB, which define the A and B ends of the segment. If IA equals IB and JA equals JB, the flow is restricted to a single node. In this case, the product QB*QBM equals total volumetric flow into or out of the node. Otherwise, the product is a volumetric rate per unit cell width. When data set is used to model flow boundary conditions, zone must follow either a row or column. Regression parameters can have any numbers from 1 to NVAR. Note that by setting IP equal to zero, a fixed specified-flow condition is simulated. <u>Omit</u> data set when NQBZ equals zero.

Data Set S.
Specified-head boundary designation; see data set L. This data set, when used in conjunction with data set K or T (following), can be used to construct specified-head boundaries. In particular, nodes designated in this data set by –1 are forced to take on values specified in data set K or T. In addition to peripheral boundary conditions, this data set, in conjunction with data set K or T, can be used to model other constant head conditions such as bodies of open water.

Data Set T.
This data set defines specified-head boundary parameter zones. Each zone is associated with a subset of the T data and the number of subsets will equal NHBZ. The initial line in each subset contains size and descriptive information about the zone and appears as follows (format, 4I5,2F10.0).

Line columns	Variable	Definition
1–5	IZ	Segment or zone number.
6–10	NN	Number of nodes in the segment.
11–15	IBPA	Regression parameter number for A end of segment. Set equal to zero if head at the A end is not a regression parameter.
16–20	IBPB	Regression parameter number for B end of segment. Set equal to zero if head at the B end is not a regression parameter.
21–30	SDHA	Standard deviation of head at A end of segment. Code as 0.0 if no prior information exists on the parameter.
31–40	SDHB	Standard deviation of head at B end of segment. Code as 0.0 if no prior information exists on the parameter.

The A and B ends are arbitrary. Segments are numbered from 1 through NHBZ, and regression parameters may have any number from 1 through NVAR. A single head change can be found for all the intermediate specified head nodes by allowing IBPA to equal IBPB.

After reading the initial line, the program then seeks NN subsequent lines in each subset that define the heads along the boundary segment (format, 2I5,F10.0).

Line columns	Variable	Definition
1–5	ILOC	Column location of node.
6–10	JLOC	Row location of node.
11–20	V	Estimated head at node ILOC, JLOC.

Note that any shape of head surface can be input along a boundary segment. Because heads at the segment ends are the only regression parameters in each segment, their influence is distributed to the intermediate nodes by linear interpolation. The linear interpolation is based upon distance from the

parameter in question, with a weight of one assigned at the end node occupied by the parameter and zero at the node of the opposite end of the segment. If zeros are assigned to the head variables V at intermediate nodes in a boundary segment, then the program automatically assumes that the head surface along the segment is simply a straight line between the heads specified at the end nodes. Omit the entire data set if NHBZ equals zero.

The following data sets are required only if the solution-only option (ISO, data set B) is specified. In the following description input variables are loosely termed parameters for convenience, but it must be realized that they are not actually regression parameters because no regression is performed. Also, data set V requires that the parameters being varied have nonzero numbers corresponding to regression parameter numbers defined for the initial solution. Hence, the input for the initial solution must be coded as if it were to be a regression for these parameters although no regression will actually be performed.

Data Set U.
Additional solution specification, one line (format, I5).

Line columns	Variable	Definition
1–5	N	Number of solutions required using alternative parameter sets. Code 0 if a solution is desired only for the initial set of parameters.

Data Set V.
New set of parameters; the set contains NVAR lines (format I5,F10.0).

Line columns	Variable	Definition
1–5	I	Parameter number.
6–15	B(I)	Parameter value.

Parameters are numbered in the same order as used for the initial solution. This data set is repeated N times, and the program will compute and print the solution corresponding to each set of parameters.

Output.—The following discussion gives the content and order of the output obtained from the program. It should be noted that some of the output is only obtained under certain specified conditions. All output is clearly labeled. However, order numbers in the following discussion do not appear in the output; they are for convenience in listing the order of output only. The statistical measures cited below are described in section 5.

1. Three title lines (data set A).
2. Problem size information (data set B).
3. Special input parameters (data set C).
4. Array sets (data sets D through L). For each variable, input information (variable name, block number, IB, IE, JB, JE, and value of FACT or IFACT) for each block is listed block by block. If specified, this information is then followed by the values of all entries in the array.
5. Observed head data (data set M).
6. Aquifer regression parameter numbers (data set N). This input is printed only if NPAR is greater than zero.
7. Coefficients of variation for aquifer regression parameters (data set O). This input is printed only if NPAR is greater than zero.
8. Zonal aquifer properties (data set P).
9. Known point-flow rates (data set Q). This input is printed only if NWELS is greater than zero.
10. Specified boundary-flow information (data set R). This input is printed only if NQBZ is greater than zero.
11. Specified boundary-head distribution (data set S). This input is printed in the same manner as indicated under 4.
12. Specified boundary-head information (data set T). For each segment, input information (segment number, number of nodes in the segment, IBPA, IBPB, SDHA, and SDHB) is printed followed by a listing of node locations and input values of specified heads. This input is printed only if NHBZ is greater than zero.
13. Error message. If IZN >0, CX≤0, and CY≤0 occur at the same cell, then the following error message is printed: "AT

CELL (i,j), IZN >0, CX=0, AND CY=0."
14. Error message. If the error in 13 happens at one or more cells, then the following message is printed: "PROGRAM ABORTED BECAUSE OF CONFLICT BETWEEN IZN, CX, AND CY." Execution then terminates.
15. Error message. If an active node is isolated from other active nodes, then the following message is printed: ACTIVE NODE (i,j) CANNOT BE ISOLATED." Execution then terminates.
16. Information on matrix solution procedure used to compute heads. Message that solution is by LDU factorization is followed by the computed minimum required dimensions of arrays used in the solution.
17. Initial solution for heads.
18. Program branch. If the solution-only option is specified, then output skips to 35.
19. Number of parameters having prior information.
20. Sensitivity matrix $\underline{X}$. This is printed only if optional print-out was selected.
21. Error message. If any diagonal term of the coefficient matrix $\underline{S}_s^T\underline{V}_s^{-1}\underline{S}_s + \underline{S}_p^T\underline{U}^{-1} \cdot \underline{S}_p s^2$ of the normal equations is smaller than 10^{-10}, then the following message is printed: "SENSITIVITIES FOR PARAMETER i EFFECTIVELY ZERO." If this error occurs, the current weighted residuals $\hat{\underline{u}}$ (section 5.5.1) are then printed. If the sensitivity print and orthogonalization option was selected, the sensitivities and orthogonalized sensitivities are also printed in the forms given in 33 and 34. Execution then terminates.
22. Coefficient matrix $\underline{S}_s^T\underline{V}_s^{-1}\underline{S}_s + \underline{S}_p^T\underline{U}^{-1}\underline{S}_p s^2 + \mu\underline{I}$ and the gradient vector $\underline{S}_s^T\underline{V}_s^{-1}(\underline{Y}_s -f_s(\underline{\xi},\underline{b})) + \underline{S}_p^T\underline{U}^{-1}s^2(\underline{Y}_p - f_p(\underline{\xi},\underline{b}))$. This output is printed only if the optional print-out was selected. (Note: If prior information of unknown reliability was used, then the matrix and vector will be modified to include this information.)
23. Error message. If, during solution of the normal equations, it becomes evident

that the problem is singular, then the following message is printed: "LEAST SQUARES COEFFICIENT MATRIX SINGULAR; SOLUTION FOR PARAMETERS NOT UNIQUE." The course of action is then the same as that given under 21.

24. Iteration number, current sum of weighted, squared deviations of computed from observed heads, determinant of the current least squares coefficient matrix defined in 22 above, current value of μ, current value of ρ, followed by the current parameter vector $\underline{b}_{r+1}$ defined by equation 3.3–19.

25. Error message. If a parameter is more than one thousand times smaller in magnitude than initially specified, then, the message "PARAMETER i EFFECTIVELY ZERO" is printed. If this problem occurs, then further iterations are aborted, the current solution is taken as the final one, and the course of action given in 21 is taken.

26. Solution converged message and final number of iterations. If the solution did not converge in the allotted number of iterations, then a message to this effect is written instead, and output skips to 32.

27. Error message. If the coefficient matrix of the normal equations (see 22) is singular when $\mu=0$, and this status has not been detected because $\mu>0$ has been computed and used by the program, then the message given in 23 is printed. In this case the subsequent course of action is the same as in 21.

28. Error variance (s^2) (section 5.4.1), final total sum of squares (sum of weighted, squared deviations of computed from observed heads plus sum of weighted, squared deviations of computed from prior estimates of parameters), and correlation coefficient (R_y) (section 5.4.2).

29. Final parameter estimates and their estimated standard errors (section 5.4.3).

30. Estimated variance-covariance matrix, $(\underline{X}_s^T\underline{V}_s^{-1}\underline{X}_s+\underline{X}_p^T\underline{U}^{-1}\underline{X}_p s^2)^{-1}s^2$ (section

5.4.3). The ordering of rows and columns matches that for $\underline{b}$ in 24.

31. Correlation matrix for parameters, $\{r_{ij}\}$ $=\{\mathrm{Cov}(b_i,b_j)/(\mathrm{Var}(b_i)\cdot\mathrm{Var}(b_j))^{\frac12}\}$ (section 5.4.4). Again, the ordering matches that for $\underline{b}$.

32. Computed and observed heads, and weighted residuals, $\hat{\underline{u}}$ (section 5.5.1).

33. Nodal sensitivities printed parameter by parameter. These are printed only if the sensitivity print and orthogonalization option was selected or if the solution did not converge.

34. Orthogonalized, scaled sensitivities, $\underline{Q}$, with the sensitivities for the prior information forming the last n_p rows. These are printed only if the sensitivities in 33 above are printed. This completes the output for regression solutions.

35. New parameters (data set V), all in sequential order; the solution number, and the solution for heads. These are printed for all solutions when the solution-only option is invoked.

Example Problem.—The following example problem illustrates use of most of the program options. As illustrated in figure 4.3–8, the modeled area consists of three aquifer zones bounded by three specified boundary-flow zones containing three boundary-flow regression parameters, a no-flow boundary, and two specified boundary-head segments containing three boundary-head regression parameters. Initial values for the regression parameters are:

$$q_{B1}=8 \qquad T_3=1,000$$
$$q_{B2}=0.8 \qquad W_1=0.0001$$
$$q_{B3}=1 \qquad W_2=0.0005$$
$$h_{B1}=40 \qquad W_3=-0.0001$$
$$h_{B2}=10 \qquad R_1=0.001$$
$$h_{B3}=16 \qquad R_2=0.0007$$
$$T_1=4,000 \qquad R_3=0.0015.$$
$$T_2=400$$

Parameter q_{B1} has prior information with a standard deviation of 0.8 on it. Because all aquifer properties and boundary flows are constant within their respective zones, multipliers for these parameters may be assigned values of unity. Assume that the estimated error variance

Figure 4.3-8

Table 4.3-2

Node	Value	Node	Value
(1,1)	29.51	(10,4)	22.46
(3,1)	26.38	(12,4)	17.16
(5,1)	25.16	(14,4)	15.74
(7,1)	26.81	(16,4)	14.69
(9,1)	23.82	(1,5)	36.05
(11,1)	23.59	(3,5)	31.64
(13,1)	16.68	(5,5)	32.91
(15,1)	15.31	(7,5)	25.05
(2,2)	28.27	(9,5)	25.47
(4,2)	25.17	(11,5)	20.39
(6,2)	26.94	(13,5)	14.40
(8,2)	23.59	(15,5)	14.37
(10,2)	22.53	(2,6)	38.66
(12,2)	17.89	(4,6)	34.51
(14,2)	15.87	(6,6)	27.81
(16,2)	15.98	(8,6)	25.98
(1,3)	31.25	(10,6)	22.16
(3,3)	27.50	(12,6)	15.73
(5,3)	26.22	(14,6)	14.79
(7,3)	25.65	(16,6)	10.48
(9,3)	24.35	(1,7)	41.28
(11,3)	22.48	(3,7)	34.82
(13,3)	15.85	(5,7)	32.00
(15,3)	15.71	(7,7)	26.97
(2,4)	29.36	(9,7)	25.67
(4,4)	27.76	(11,7)	18.12
(6,4)	25.51	(13,7)	18.40
(8,4)	24.43	(15,7)	11.90

Heads H on the distal side of the aquitard are computed by first assigning constant values to all cells in an aquifer property zone, then computing nodal values as the average of all adjacent cell values. Cell values are 25 for zone 1, 35 for zone 2, and 15 for zone 3.

Input data for the example problem are coded on figure 4.3-9 and are contained in file EXPROB.DAT in the diskette accompanying this report. These data should be compared with the data input instructions given above.

Output is given in figure 4.3-10. Statistical measures listed are described in section 5.

for use with the prior information has a value of unity, $t_{mx} = 1.5$, and $\cos\Theta_{mx} = 0.08$.

Locations of observed heads are indicated by the small open circles on figure 4.3-8. Values corresponding to these locations are given in table 4.3-2. All observations have a weight ω_{ii} of unity.

EXAMPLE PROBLEM FOR DATA INPUT

```
      16      7      3     56     9     15     0     3     2    20     1     1     0
              1.5          .08           0           0           1
DX            1      0
      1      15      1      1           1000
DY            1      0
      1       6      1      1           1000
CX            1      0
      1      15      1      6              1
CY            1      0
      1      15      1      6              1
VL            1      0
      1      15      1      6              1
HR            8      0
      1       6      1      3             25
      7      10      1      7             25
      1       5      4      4             30
      6       6      4      4           27.5
      6       6      5      7             30
      1       5      5      7             35
     11      11      1      7             20
     12      16      1      7             15
QR            1      0
      1      15      1      6              1
HC            1      0
      1      16      1      7             10
IZN           4      0
      1       5      1      3      1
      6      10      1      6      1
      1       5      4      6      2
     11      15      1      6      3
      1       1      1             0           0     29.51           1
      2       3      1          2000           0     26.38           1
      3       5      1          4000           0     25.16           1
      4       7      1          6000           0     26.81           1
      5       9      1          8000           0     23.82           1
      6      11      1         10000           0     23.59           1
      7      13      1         12000           0     16.68           1
      8      15      1         14000           0     15.31           1
      9       2      2          1000        1000     28.27           1
     10       4      2          3000        1000     25.17           1
     11       6      2          5000        1000     26.94           1
     12       8      2          7000        1000   · 23.59           1
     13      10      2          9000        1000     22.53           1
     14      12      2         11000        1000     17.89           1
     15      14      2         13000        1000     15.87           1
     16      15      2         15000        1000     15.98           1
     17       1      3             0        2000     31.25           1
     18       3 ⌐    3          2000        2000     27.50           1
     19       5      3          4000        2000     26.22           1
     20       7      3          6000        2000     25.65           1
```

Figure 4.3-9

21	9	3	8000	2000	24.35	1
22	11	3	10000	2000	22.48	1
23	13	3	12000	2000	15.85	1
24	15	3	14000	2000	15.71	1
25	2	4	1000	3000	29.36	1
26	4	4	3000	3000	27.76	1
27	6	4	5000	3000	25.51	1
28	8	4	7000	3000	24.43	1
29	10	4	9000	3000	22.46	1
30	12	4	11000	3000	17.16	1
31	14	4	13000	3000	15.74	1
32	15	4	15000	3000	14.69	1
33	1	5	0	4000	36.05	1
34	3	5	2000	4000	31.64	1
35	5	5	4000	4000	32.91	1
36	7	5	6000	4000	25.05	1
37	9	5	8000	4000	25.47	1
38	11	5	10000	4000	20.39	1
39	13	5	12000	4000	14.40	1
40	15	5	14000	4000	14.37	1
41	2	6	1000	5000	38.66	1
42	4	6	3000	5000	34.51	1
43	6	6	5000	5000	27.81	1
44	8	6	7000	5000	25.98	1
45	10	6	9000	5000	22.16	1
46	12	6	11000	5000	15.73	1
47	14	6	13000	5000	14.79	1
48	15	6	15000	5000	10.48	1
49	1	6	0	6000	41.28	1
50	3	6	2000	6000	34.82	1
51	5	6	4000	6000	32.00	1
52	7	6	6000	6000	26.97	1
53	9	6	8000	6000	25.67	1
54	11	6	10000	6000	18.12	1
55	13	6	12000	6000	18.40	1
56	15	6	14000	6000	11.90	1
1	1	1	4	7		
2	2	2	5	8		
3	3	3	6	9		
1		0				
2		0				
3		0				
4		0				
5		0				
6		0				
7		0				
8		0				
9		0				
1		4000	4000	.001	.0001	
2		400	400	.0007	.0005	
3		1000	1000	.0015	-.0001	
1	1	1	4	10	8	.8	1
1	4	1	7	11	.8	0	1

Figure 4.3-9—Continued

```
        8    1   16    1   12         1         0         1
IN           2    0
       16   16    1    7   -1
        1   15    7    7   -1
        1   16   13   14         0         0
        1    7        40
        2    7
        3    7
        4    7
        5    7
        6    7
        7    7
        8    7
        9    7
       10    7
       11    7
       12    7
       13    7
       14    7
       15    7
       16    7        10
        2    7   14   15         0         0
       16    7        10
       16    6
       16    5
       16    4
       16    3
       16    2
       16    1        16
```

Figure 4.3–9—Continued

EXAMPLE PROBLEM FOR DATA INPUT

```
NUMBER OF COLUMNS (ID) ------------------------------------------ =    16
NUMBER OF ROWS (JD) --------------------------------------------- =     7
NUMBER OF AQUIFER ZONES (NZNS) ---------------------------------- =     3
NUMBER OF OBSERVATIONS (NOBS) ----------------------------------- =    56
NUMBER OF AQUIFER PARAMETERS (NPAR) ----------------------------- =     9
TOTAL NUMBER OF PARAMETERS (NVAR) ------------------------------- =    15
NUMBER OF KNOWN POINT FLOWS (NWELS) ----------------------------- =     0
NUMBER OF SPECIFIED FLOW ZONES (NQBZ) --------------------------- =     3
NUMBER OF SPECIFIED HEAD ZONES (NHBZ) --------------------------- =     2
MAXIMUM NUMBER OF ITERATIONS (NUM) ------------------------------ =    20
SENSITIVITY PRINT AND ORTHOGONALIZATION OPTION (IPRX) ----------- =     1
ADDITIONAL PRINTOUT OPTION (IPO) -------------------------------- =     1
SOLUTION ONLY OPTION (ISO) -------------------------------------- =     0
MAXIMUM ALLOWABLE PARAMETER CORRECTION (DMX) -- =  1.5000
SEARCH DIRECTION ADJUSTMENT PARAMETER (CSA) -- =  .80000E-01
RIDGE PARAMETER FOR REGRESSION (RP) ------------ =  .00000
BIAS PARAMETER FOR REGRESSION (BP) ------------- =  .00000
ESTIMATED ERROR VARIANCE (EV) ------------------ =  1.0000
```

```
DX
  1  IB =   1  IE =  15  JB =   1  JE =   1  FACT =  1000.0

DX  ARRAY

        1        2        3        4        5        6        7        8        9       10
  1  1000.0   1000.0   1000.0   1000.0   1000.0   1000.0   1000.0   1000.0   1000.0   1000.0

       11       12       13       14       15
     1000.0   1000.0   1000.0   1000.0   1000.0

DY
  1  IB =   1  IE =   6  JB =   1  JE =   1  FACT =  1000.0

DY  ARRAY

        1        2        3        4        5        6
  1  1000.0   1000.0   1000.0   1000.0   1000.0   1000.0

CX
  1  IB =   1  IE =  15  JB =   1  JE =   6  FACT =  1.0000
```

CX ARRAY

	1	2	3	4	5	6	7	8	9	10
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	11	12	13	14	15
6	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000

CY   IB = 1   IE = 15   JB = 1   JE = 6   FACT = 1.0000

CY ARRAY

	1	2	3	4	5	6	7	8	9	10
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	11	12	13	14	15
6	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000

VL   IB = 1   IE = 15   JB = 1   JE = 6   FACT = 1.0000

Figure 4.3-10

VL ARRAY

	1	2	3	4	5	6	7	8	9	10
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	11	12	13	14	15
6	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000

HR

	IB		IE		JB		JE		FACT	
1	IB =	1	IE =	1	JB =	6	JE =	1	FACT =	25.000
2	IB =	7	IE =	7	JB =	10	JE =	1	FACT =	25.000
3	IB =	1	IE =	1	JB =	5	JE =	4	FACT =	30.000
4	IB =	6	IE =	6	JB =	6	JE =	4	FACT =	27.500
5	IB =	6	IE =	6	JB =	6	JE =	7	FACT =	30.000
6	IB =	1	IE =	1	JB =	5	JE =	5	FACT =	35.000
7	IB =	11	IE =	11	JB =	11	JE =	1	FACT =	20.000
8	IB =	12	IE =	12	JB =	16	JE =	1	FACT =	15.000

HR	ARRAY 1	2	3	4	5	6	7	8	9	10
7	35.000	35.000	35.000	35.000	35.000	30.000	25.000	25.000	25.000	25.000
6	35.000	35.000	35.000	35.000	35.000	30.000	25.000	25.000	25.000	25.000
5	35.000	35.000	35.000	35.000	35.000	30.000	25.000	25.000	25.000	25.000
4	30.000	30.000	30.000	30.000	30.000	27.500	25.000	25.000	25.000	25.000
3	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000
2	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000
1	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000	25.000

HR	11	12	13	14	15	16
7	20.000	15.000	15.000	15.000	15.000	15.000
6	20.000	15.000	15.000	15.000	15.000	15.000
5	20.000	15.000	15.000	15.000	15.000	15.000
4	20.000	15.000	15.000	15.000	15.000	15.000
3	20.000	15.000	15.000	15.000	15.000	15.000
2	20.000	15.000	15.000	15.000	15.000	15.000
1	20.000	15.000	15.000	15.000	15.000	15.000

Figure 4.3-10—Continued

QR   IB =   1   IE =   15   JB =   1   JE =   6   FACT =   1.0000

QR ARRAY

	1	2	3	4	5	6	7	8	9	10
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	11	12	13	14	15
6	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000

HC   IB =   1   IE =   16   JB =   1   JE =   7   FACT =   10.000

HC ARRAY

	1	2	3	4	5	6	7	8	9	10
7	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
6	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
5	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
4	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
3	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
2	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000
1	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000

	11	12	13	14	15	16
7	10.000	10.000	10.000	10.000	10.000	10.000
6	10.000	10.000	10.000	10.000	10.000	10.000
5	10.000	10.000	10.000	10.000	10.000	10.000
4	10.000	10.000	10.000	10.000	10.000	10.000
3	10.000	10.000	10.000	10.000	10.000	10.000
2	10.000	10.000	10.000	10.000	10.000	10.000
1	10.000	10.000	10.000	10.000	10.000	10.000

IZN

IZN	IB =	IE =	JB =	JE =	IFACT =
1	1	5	1	3	1
2	6	10	1	6	1
3	1	5	4	6	2
4	11	15	1	6	3

IZN ARRAY

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
6	2	2	2	2	2	1	1	1	1	1	3	3	3	3	3
5	2	2	2	2	2	1	1	1	1	1	3	3	3	3	3
4	2	2	2	1	1	1	1	1	1	1	3	3	3	3	3
3	1	1	1	1	1	1	1	1	1	1	3	3	3	3	3
2	1	1	1	1	1	1	1	1	1	1	3	3	3	3	3
1	1	1	1	1	1	1	4	1	1	1	3	3	3	3	3

OBSERVED HEAD DATA

OBS. NO.	CELL LOC.		X LOC.	Y LOC.	OBS. HEAD	WEIGHT
1	1	1	.00000	.00000	29.510	1.0000
2	3	1	2000.0	.00000	26.380	1.0000
3	5	1	4000.0	.00000	25.160	1.0000
4	7	1	6000.0	.00000	26.810	1.0000
5	9	1	8000.0	.00000	23.820	1.0000
6	11	1	10000.	.00000	23.590	1.0000
7	13	1	12000.	.00000	16.680	1.0000
8	15	1	14000.	.00000	15.310	1.0000
9	2	2	1000.0	1000.0	28.270	1.0000
10	4	2	3000.0	1000.0	25.170	1.0000
11	6	2	5000.0	1000.0	26.940	1.0000
12	8	2	7000.0	1000.0	23.590	1.0000
13	10	2	9000.0	1000.0	22.530	1.0000
14	12	2	11000.	1000.0	17.890	1.0000
15	14	2	13000.	1000.0	15.870	1.0000
16	15	2	15000.	1000.0	15.980	1.0000
17	1	3	.00000	2000.0	31.250	1.0000
18	3	3	2000.0	2000.0	27.500	1.0000
19	5	3	4000.0	2000.0	26.220	1.0000
20	7	3	6000.0	2000.0	25.650	1.0000
21	9	3	8000.0	2000.0	24.350	1.0000
22	11	3	10000.	2000.0	22.480	1.0000
23	13	3	12000.	2000.0	15.850	1.0000
24	15	3	14000.	2000.0	15.710	1.0000
25	2	4	1000.0	3000.0	29.360	1.0000
26	4	4	3000.0	3000.0	27.760	1.0000

Figure 4.3-10—Continued

27	6	4	5000.0	3000.0	25.510	1.0000
28	8	4	7000.0	3000.0	24.430	1.0000
29	10	4	9000.0	3000.0	22.460	1.0000
30	12	4	11000.	3000.0	17.160	1.0000
31	14	4	13000.	3000.0	15.740	1.0000
32	15	4	15000.	3000.0	14.690	1.0000
33	1	5	.00000	4000.0	36.050	1.0000
34	3	5	2000.0	4000.0	31.640	1.0000
35	5	5	4000.0	4000.0	32.910	1.0000
36	7	5	6000.0	4000.0	25.050	1.0000
37	9	5	8000.0	4000.0	25.470	1.0000
38	11	5	10000.	4000.0	20.390	1.0000
39	13	5	12000.	4000.0	14.400	1.0000
40	15	5	14000.	4000.0	14.370	1.0000
41	2	6	1000.0	5000.0	38.660	1.0000
42	4	6	3000.0	5000.0	34.510	1.0000
43	6	6	5000.0	5000.0	27.810	1.0000
44	8	6	7000.0	5000.0	25.980	1.0000
45	10	6	9000.0	5000.0	22.160	1.0000
46	12	6	11000.	5000.0	15.730	1.0000
47	14	6	13000.	5000.0	14.790	1.0000
48	15	6	15000.	5000.0	10.480	1.0000
49	1	6	.00000	6000.0	41.280	1.0000
50	3	6	2000.0	6000.0	34.820	1.0000
51	5	6	4000.0	6000.0	32.000	1.0000
52	7	6	6000.0	6000.0	26.970	1.0000
53	9	6	8000.0	6000.0	25.670	1.0000
54	11	6	10000.	6000.0	18.120	1.0000
55	13	6	12000.	6000.0	18.400	1.0000
56	15	6	14000.	6000.0	11.900	1.0000

AQUIFER PARAMETER NUMBERS

ZONE	TRANX	TRANY	VLEAK	QDIST
1	1	2	4	7
2		3	5	8
3			6	9

STANDARD DEVIATIONS
FOR AQUIFER PARAMETERS

PAR. NO.	STD. DEV.
1	.00000
2	.00000
3	.00000
4	.00000
5	.00000
6	.00000
7	.00000
8	.00000
9	.00000

INITIAL AQUIFER PARAMETERS BY ZONE

ZONE	TRANX	TRANY	VLEAK	QDIST
1	4000.0	4000.0	.10000E-02	.10000E-03
2	400.00	400.00	.70000E-03	.50000E-03
3	1000.0	1000.0	.15000E-02	-.10000E-03

NODE NO.S

IA	JA	IB	JB
1	1	1	4
1	4	1	7
8	1	16	1

INITIAL SPECIFIED FLOW DATA

PAR. NO.	FLOW PARAMETER	STD. DEV.	MULTIPLIER
10	8.0000	.80000	1.0000
11	.80000	.00000	1.0000
12	1.0000	.00000	1.0000

IN

	IB =	JB =	IE =	JE =	IFACT =
1	16	16	16	1	7	-1
2	1	15	1	7	7	-1

IN ARRAY

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
7	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

INITIAL SPECIFIED HEAD DATA

NO. OF NODES IN ZONE   1 = 16
PAR. NO. HEAD A =   13   PAR. NO. HEAD B =   14
STD. DEV. HEAD A = .00000   STD. DEV. HEAD B = .00000

INITIAL VALUES OF HEAD

I	J	HEAD
1	7	40.000
2	7	.00000
3	7	.00000
4	7	.00000
5	7	.00000
6	7	.00000
7	7	.00000
8	7	.00000
9	7	.00000
10	7	.00000
11	7	.00000
12	7	.00000
13	7	.00000
14	7	.00000
15	7	.00000
16	7	10.000

Figure 4.3-10—Continued

```
NO. OF NODES IN ZONE    2 = 7     PAR. NO. HEAD B = 15
PAR. NO. HEAD A = 14    .00000    STD. DEV. HEAD B = .00000
STD. DEV. HEAD A = .00000
                        INITIAL VALUES OF HEAD
                        I    J      HEAD
                        16   7     10.000
                        16   6      .00000
                        16   5      .00000
                        16   4      .00000
                        16   3      .00000
                        16   2      .00000
                        16   1     16.000
```

SOLUTION BY LDU FACTORIZATION ASSUMING D4 ORDERING
MINIMUM DIMENSIONS FOR ARRAYS USED BY THIS METHOD ARE AS FOLLOWS
AU: 5 BY 45
AL: 9 BY 45
IC: 5 BY 45
V: 90

INITIAL SOLUTION

	1	2	3	4	5	6	7	8	9	10
7	40.000	38.000	36.000	34.000	32.000	30.000	28.000	26.000	24.000	22.000
6	37.229	36.064	35.263	34.409	32.890	28.278	26.913	25.554	24.162	22.673
5	35.441	34.373	33.789	33.201	31.928	27.398	26.273	25.255	24.188	22.940
4	31.311	29.853	28.947	28.280	27.600	26.598	25.818	25.043	24.168	23.046
3	30.504	28.992	28.034	27.353	26.759	26.128	25.538	24.917	24.161	23.113
2	30.074	28.548	27.576	26.905	26.369	25.875	25.397	24.882	24.212	23.216
1	29.940	28.411	27.438	26.771	26.255	25.800	25.368	24.947	24.366	23.409

	11	12	13	14	15	16
7	20.000	18.000	16.000	14.000	12.000	10.000
6	20.982	17.072	15.556	14.585	13.453	11.000
5	21.315	16.957	15.500	14.807	14.007	12.000
4	21.447	16.977	15.532	14.946	14.378	13.000
3	21.534	17.036	15.605	15.088	14.724	14.000
2	21.660	17.180	15.772	15.308	15.118	15.000
1	21.916	17.623	16.251	15.816	15.719	16.000

NO. OF PARAMETERS HAVING PRIOR INFORMATION = 1

OBSERVATION NUMBER AND SENSITIVITIES FOR EACH PARAMETER
```
1 -.38228E-03  .50620E-03 -.10080E-04 -2586.5      83.948     -4.3781     887.72
   131.42      2.7446      .45909       .47201E-01  .63299E-02  .10367E-01  .57848E-02
   .15307E-04
2  .60319E-04  .48072E-03 -.16301E-04 -1946.1      111.72     -7.0915     885.92
   122.18      4.4396      .15965       .30392E-01  .10512E-01  .13276E-01  .85986E-02
   .24853E-04
```

3	.84556E-04 91.807 .65461E-04	.35640E-03 11.550	-.42379E-04 .54188E-01	-1104.0 .12744E-01	124.56 .29967E-01	-18.535 .22834E-01	877.71 .17803E-01
4	.20950E-04 49.653 .18996E-03	.19118E-03 32.568	-.11934E-03 .17697E-01	-377.48 .44771E-02	78.699 .10882	-52.851 .33990E-01	851.92 .30295E-01
5	-.44029E-04 21.594 .57555E-03	.83352E-04 93.554	-.34274E-03 .58100E-02	195.90 .15130E-02	36.560 .37073	-155.50 .34274E-01	774.03 .35670E-01
6	.26304E-03 9.4671 .17904E-02	.36640E-04 265.38	-.98740E-03 .22527E-02	64.792 .59308E-03	16.440 .49558	-468.15 .23666E-01	542.22 .27636E-01
7	.28057E-04 .92848 .25686E-01	.35984E-05 605.89	.30068E-03 .21021E-03	7.7542 .55570E-04	1.6281 .67776	-698.55 .30985E-02	52.278 .94553E-02
8	.25379E-05 .81492E-01 .29071	.31618E-06 453.52	-.12566E-06 .17832E-04	.74317 .47267E-05	.14383 .53296	-249.14 .40054E-03	4.5450 .24008E-01
9	-.83160E-04 147.71 .17175E-04	.57567E-03 3.0815	-.11317E-04 .26739	-2294.8 .48691E-01	99.794 .70743E-02	-4.9144 .11698E-01	873.71 .65538E-02
10	.74473E-04 126.28 .38574E-04	.49706E-03 6.8912	-.25302E-04 .91556E-01	-1475.1 .21604E-01	149.77 .16239E-01	-11.007 .18505E-01	869.67 .13061E-01
11	.47626E-04 78.330 .10790E-03	.30114E-03 19.029	-.69809E-04 .30075E-01	-697.32 .76372E-02	120.92 .47836E-01	-30.536 .33684E-01	855.15 .27683E-01
12	-.78560E-05 35.078 .31695E-03	.13518E-03 54.179	-.19833E-03 .96800E-02	-50.049 .25227E-02	58.932 .14360	-87.936 .41772E-01	812.07 .40022E-01
13	.66592E-04 14.175 .96774E-03	.54863E-04 156.44	-.56959E-03 .33205E-02	331.28 .87589E-03	24.679 .23450	-258.98 .34954E-01	679.82 .39677E-01
14	.93956E-04 2.9720 .71191E-02	.11523E-04 533.15	.78272E-03 .66123E-03	26.881 .17509E-03	5.2279 .21476	-919.11 .99077E-02	166.18 .14989E-01
15	.90142E-05 .28467 .78345E-01	.11049E-05 587.43	.25398E-03 .61321E-04	2.7048 .16276E-04	.50386 .19256	-286.38 .14460E-02	15.800 .23067E-01
16	.00000 .00000 .83333	.00000 .00000	.00000 .00000	.00000 .00000	.00000 .00000	.00000 .00000	.00000 .16667
17	-.45126E-03 211.03 .14073E-04	.77585E-03 2.5325	-.93026E-05 .43630	-2315.3 .10673	79.365 .56983E-02	-4.0344 .12821E-01	829.78 .58605E-02
18	-.30579E-04 199.47 .22863E-04	.80329E-03 4.1228	-.15146E-04 .14804	-1711.6 .43134E-01	182.15 .91359E-02	-6.5629 .16340E-01	828.39 .94350E-02
19	.28901E-05 150.39 .59929E-04	.57360E-03 10.867	-.39933E-04 .49074E-01	-925.31 .13689E-01	239.20 .22888E-01	-17.263 .32544E-01	821.58 .24242E-01

Figure 4.3-10—Continued

20	.74527E-05 65.606 .16546E-03	.25177E-03 30.027	-.11028E-03 .14925E-01	-317.08 .40309E-02	113.08 .56894E-01	-47.651 .59526E-01	798.48 .50865E-01
21	.97692E-06 23.765 .48724E-03	.92081E-04 87.594	-.32071E-03 .48754E-02	231.85 .13058E-02	42.275 .11083	-139.00 .58707E-01	730.34 .62438E-01
22	.31325E-03 9.5215 .14719E-02	.37011E-04 258.23	-.93635E-03 .19044E-02	105.13 .50911E-03	17.063 .11689	-407.58 .40150E-01	513.48 .50880E-01
23	.29513E-04 .89858 .19946E-01	.34962E-05 601.69	.55627E-03 .17858E-03	9.8336 .47720E-04	1.6132 .66737E-01	-466.70 .59267E-02	48.803 .22167E-01
24	.25224E-05 .76760E-01 .20807	.29879E-06 451.40	-.48766E-04 .15211E-04	.83892 .40639E-05	.13791 .37527E-01	29.768 .10161E-02	4.1862 .10958
25	.30166E-03 327.31 .14409E-04	.13148E-02 2.6005	-.95542E-05 .22761	-1554.3 .11082	139.64 .57427E-02	-4.1385 .16999E-01	743.34 .66367E-02
26	.18475E-03 295.75 .32589E-04	.11880E-02 5.9472	-.21865E-04 .76582E-01	-861.49 .25953E-01	420.38 .12207E-01	-9.4268 .26347E-01	741.46 .16217E-01
27	.87979E-04 146.91 .92238E-04	.55544E-03 17.470	-.64379E-04 .19856E-01	-282.46 .57530E-02	259.77 .27894E-01	-27.335 .81928E-01	726.82 .62115E-01
28	.91868E-05 44.467 .22814E-03	.17295E-03 44.753	-.16514E-03 .64338E-02	-11.250 .17865E-02	82.965 .48881E-01	-69.148 .10402	700.61 .96892E-01
29	.71576E-04 14.070 .67123E-03	.55001E-04 138.63	-.51144E-03 .22930E-02	364.06 .62403E-03	26.031 .64502E-01	-210.11 .83153E-01	594.94 .10344
30	.92326E-04 2.7081 .46455E-02	.10587E-04 517.97	.76168E-03 .46383E-03	34.853 .12548E-03	4.9790 .31186E-01	-781.62 .25569E-01	143.17 .47757E-01
31	.83915E-05 .24514 .48951E-01	.95788E-06 574.43	.84438E-04 .43482E-04	3.0762 .11721E-04	.44848 .17367E-01	-55.665 .57210E-02	13.157 .70359E-01
32	.00000 .00000 .50000	.00000 .00000	.00000 .00000	.00000 .00000	.00000 .00000	.00000 .00000	.00000 .50000
33	-.13585E-03 1002.5 .40076E-05	-.25598E-02 .72464	-.26627E-05 .95183E-01	-477.25 1.2178	-301.31 .15875E-02	-1.1524 .78076E-01	213.74 .53671E-02
34	-.66372E-04 990.80 .68844E-05	-.22910E-02 1.2640	-.46499E-05 .40683E-01	-345.96 .73922E-01	901.99 .25531E-02	-1.9997 .73485E-01	216.49 .14555E-01
35	-.79757E-04 808.93 .26192E-04	-.36705E-02 5.0881	-.18796E-04 .13808E-01	-139.57 .71506E-02	2002.1 .75221E-02	-7.8992 .94268E-01	282.90 .48924E-01
36	-.37833E-04 89.202 .99179E-04	.35423E-03 20.368	-.75545E-04 .61410E-02	-169.17 .18139E-02	184.86 .21123E-01	-31.070 .18904	572.98 .14489

Row	Values
37	-.65253E-05, 21.894, .27081E-03, 7.3247, .78789E-03, .86770E-04, 61.963, -.23158E-03, .24664E-02, 219.88, .69171E-03, 42.569, .31053E-01, -91.671, .16382, 547.07, .18057
38	.28872E-04, 216.36, -.81642E-03, .10252E-02, 112.84, .28178E-03, 13.856, .28030E-01, -307.78, .11166, 389.90, .16798
39	.23516E-04, .64136, .10225E-01, .25208E-05, 552.79, .43777E-03, .98203E-04, 9.3753, .26781E-04, 1.1989, .79262E-02, -333.52, .26022E-01, 34.509, .99611E-01
40	.18717E-05, .52032E-01, .10455, .20407E-06, 421.64, -.36846E-03, .84765E-05, .72193, .22995E-05, .96428E-01, .30619E-02, 279.22, .70276E-02, 2.8109, .25312
41	-.25042E-04, 931.86, .14039E-05, -.45069E-03, .25718, -.94596E-06, .17418E-01, -112.33, .26527, -613.46, .52938E-03, -.40719, .26812, 57.617, .21579E-01
42	-.21672E-04, 893.31, .46073E-05, -.22841E-02, .89634, -.33126E-05, .68573E-02, -63.908, .15118E-01, 782.39, .13611E-02, -1.3912, .23864, 70.690, .64002E-01
43	-.14945E-03, 153.08, .38360E-04, .67722E-03, 8.0378, -.29884E-04, .31654E-02, 23.624, .11004E-02, 298.22, .80304E-02, -12.191, .34417, 327.26, .20931
44	-.28535E-05, 28.327, .82432E-04, .11684E-03, 19.149, -.71883E-04, .18309E-02, -35.333, .53855E-03, 56.396, .11594E-01, -28.215, .31371, 358.20, .27945
45	.19335E-04, 6.9334, .23516E-03, .27719E-04, 70.983, -.27348E-03, .76354E-03, 275.39, .21377E-03, 13.427, .13371E-01, -98.168, .23796, 314.97, .33550
46	.52326E-04, 1.2223, .15742E-02, .48415E-05, 374.84, .67366E-03, .16018E-03, 21.772, .44215E-04, 2.3279, .49386E-02, -589.79, .10476, 70.055, .26050
47	.40914E-05, .10474, .16394E-01, .41281E-06, 418.28, -.12829E-03, .15304E-04, 1.6529, .41870E-05, .19695, .16250E-02, 99.608, .43177E-01, 5.7928, .30714
48	.00000, .00000, .16667, .00000, .00000, .00000, .00000, .00000, .00000, .00000, .00000, .83333
49	.00000, .00000, .00000, .00000, .00000, .00000, .00000, .00000, 1.0000, .00000, .00000
50	.00000, .00000, .00000, .00000, .00000, .00000, .00000, .86667, .00000, .13333
51	.00000, .00000, .00000, .00000, .00000, .00000, .00000, .73333, .00000, .26667
52	.00000, .00000, .00000, .00000, .00000, .00000, .00000, .60000, .00000, .40000

Figure 4.3-10—Continued

53	.00000	.00000	.00000	.00000	.00000	.00000	.46667	.53333	
54	.00000	.00000	.00000	.00000	.00000	.33333	.00000	.66667	
55	.00000	.00000	.00000	.00000	.00000	.20000	.00000	.80000	
56	.00000	.00000	.00000	.00000	.00000	.66667E-01	.00000	.93333	

SCALED LEAST SQUARES MATRIX

	1	2	3	4	5	6	7	8	9	10
1	1.0000	-.69544E-01	-.26635	.53102	-.17273	-.34007	-.17912	-.31368	.23649	-.35370
2	-.69544E-01	1.0000	-.30813E-01	.25996	.53051	-.20677E-01	.23497	.63216	.13554E-01	.11366
3	-.26635	-.30813E-01	1.0000	-.25211E-01	-.44226E-01	-.20498	-.34146	-.31616E-01	.16613	-.13671E-01
4	.53102	.25996	-.25211E-01	1.0000	.20068	-.23735E-01	.67194	.33069	.16667E-01	-.48758
5	-.17273	.53051	-.44226E-01	.20068	1.0000	-.28414E-01	.30415	.50312	-.18664E-01	-.76453E-01
6	-.34007	-.20677E-01	-.20498	-.23735E-01	-.28414E-01	1.0000	.26777	-.20134E-01	.73701	.84692E-02
7	-.17912	.23497	-.34146	.67194	.30415	.26777	1.0000	.33113	.17797	.31876
8	-.31368	.63216	-.31616E-01	.33069	.50312	-.20134E-01	.33113	1.0000	.13167E-01	.16830
9	.23649	.13554E-01	.16613	.16667E-01	-.18664E-01	.73701	.17797	.13167E-01	1.0000	.54464E-02
10	-.35370	.11366	-.13671E-01	-.48758	-.76453E-01	.84692E-02	.31876	.16830	.54464E-02	1.0000
11	-.24478	-.39631	-.52388E-02	-.20846	-.11478	-.31911E-02	.14497	.60353	.20517E-02	.13317
12	.20587	.32808E-01	-.21300	-.34027E-02	.43106E-01	.63606	.33531	.32631E-01	.58178	.17099E-01
13	-.27537E-01	-.66174E-01	-.73124E-01	-.29440E-01	.10554	-.78825E-01	.17203	.20337	.63744E-01	.20165E-01
14	.27664E-01	.45383E-02	-.77518E-01	-.33928E-02	-.65180E-01	-.10046	.14428	.55433E-01	.18631	.90366E-02
15	.67823E-02	.21096E-01	-.14278E-01	-.37787E-03	.26532E-03	-.52616E-03	.35636E-02	.18356E-03	.20016	.71068E-04

	11	12	13	14	15
1	-.24478	.20587	-.27537E-01	.27664E-01	.67823E-02
2	-.39631	.32808E-01	-.66174E-01	.45383E-02	.21096E-01
3	-.52388E-02	-.21300	-.73124E-01	-.77518E-01	-.14278E-01
4	-.20846	-.34027E-02	-.29440E-01	-.33928E-02	-.37787E-03
5	-.11478	.43106E-01	.10554	-.65180E-01	.26532E-03
6	-.31911E-02	.63606	-.78825E-01	-.10046	-.52616E-03
7	.14497	.33531	.17203	.14428	.35636E-02
8	.60353	.32631E-01	.20337	.55433E-01	.18356E-03
9	.20517E-02	.58178	.63744E-01	.18631	.20016
10	.13317	.17099E-01	.20165E-01	.90366E-02	.71068E-04
11	1.0000	.64661E-02	-.77897E-01	.82024E-02	.25505E-04
12	.64661E-02	1.0000	-.47384E-01	.59304E-01	.16373
13	-.77897E-01	-.47384E-01	1.0000	.43839	.17691E-02
14	.82024E-02	.59304E-01	.43839	1.0000	.27922
15	.25505E-04	.16373	.17691E-02	.27922	1.0000

```
SCALED GRADIENT VECTOR
 .51601    -.99079    -.45346    1.2969    -1.2398    -.22907    -1.3453    .30868
 .67924    -.37087     .90839     .69633     .16964     .49968    1.6963

ITERATION NO.   1
YSQ =  62.212    DET(C) = .31948E-03    AMP = .00000    AP = .46254
REGRESSION PARAMETERS
 3400.8     220.54     701.64    .98853E-03    .14786E-03    .12781E-02    -.50000E-04    .68114E-03
 .20223E-04    8.0000    .40847    1.0376         40.039        10.019        16.657
        * * * * * * * * *

    Output for iterations 2 through 6 is analogous to output for iteration 1,
and, thus, is omitted. The scaled least squares matrix and gradient vector
following is for iteration 6.

        * * * * * * * * *

SCALED LEAST SQUARES MATRIX
```

	1	2	3	4	5	6	7	8	9	10
1	1.0000	.17180	-.17219	.58498	.82455E-02	-.22676	-.19756	-.25747	.15229	-.45891
2	.17180	1.0000	.19293E-03	-.43378E-01	-.17720	-.38005E-03	-.83920E-01	-.95324	-.24789E-03	-.56820E-01
3	-.17219	.19293E-03	1.0000	-.51165E-01	-.11965E-01	-.38104	-.25862	-.91209E-02	.30206	-.72582E-02
4	.58498	-.43378E-01	-.51165E-01	1.0000	.86505E-02	-.34320E-01	-.55596	-.18514	.22358E-01	-.57010
5	.82455E-02	-.17720	-.11965E-01	.86505E-02	1.0000	-.72451E-02	.12575	.45636E-01	.45959E-02	-.10614E-01
6	-.22676	-.38005E-03	-.38104	-.34320E-01	-.72451E-02	1.0000	.20760	-.54336E-02	.80292	-.44164E-02
7	-.19756	-.83920E-01	-.25862	-.55596	.12575	.20760	1.0000	.22633	.13295	.35646
8	-.25747	-.95324	-.91209E-02	-.18514	.45636E-01	-.54336E-02	.22633	1.0000	.34231E-02	.13920
9	.15229	-.24789E-03	.30206	.22358E-01	.45959E-02	.80292	.13295	.34231E-02	1.0000	.27229E-02
10	-.45891	-.56820E-01	-.72582E-02	-.57010	-.10614E-01	-.44164E-02	.35646	.13920	.27229E-02	1.0000
11	-.26481	.63943	-.18290E-02	-.17598	-.46230	-.10535E-02	.12938	.64480	.65011E-03	.14295
12	.12516	.34937E-03	-.11514	-.29284E-01	.11169E-01	.67278	.25625	.93450E-02	.54316	.98485E-02
13	-.38769E-01	-.22608	.41526E-01	-.97069E-02	-.52273E-02	-.53213E-01	.16347	.26989	.43188E-01	.28550E-01
14	-.18259E-01	-.40783E-01	-.44719E-01	.18788E-01	.36389E-01	-.84249E-01	.12543	.53228E-01	.14376	.77942E-02
15	.20587E-02	.63503E-05	-.58981E-02	.14509E-03	.25097E-04	-.88041E-01	.12579E-02	.17694E-04	.16090	.14932E-04

Figure 4.3-10—Continued

	11	12	13	14	15
1	-.26481	.12516	-.38769E-01	.18259E-01	.20587E-02
2	-.63943	.34937E-03	-.22608	-.40783E-01	.63503E-05
3	-.18290E-02	-.11514	-.41526E-01	-.44719E-01	-.58981E-02
4	-.17598	.29284E-01	-.97069E-02	-.18788E-01	.14509E-03
5	-.46230	.11169E-01	-.52273E-02	-.36389E-01	.25097E-04
6	-.10535E-02	-.67278	-.53213E-01	-.84249E-01	-.88041E-01
7	.12938	.25625	.16347	.12543	.12579E-02
8	.64480	.93450E-02	.26989	.53228E-01	.17694E-04
9	.65011E-03	.54316	.43188E-01	.14376	.16090
10	.14295	.98485E-02	.28550E-01	.77942E-02	.14932E-04
11	1.0000	.23673E-02	.13444	.12050E-01	.29220E-05
12	-.23673E-02	1.0000	.25953E-01	.33840E-01	.14324
13	.13444	.25953E-01	1.0000	.42776	.70059E-03
14	-.12050E-01	-.33840E-01	.42776	1.0000	.27776
15	.29220E-05	.14324	.70059E-03	.27776	1.0000

SCALED GRADIENT VECTOR
-.51657E-03   .11084E-03   -.20590E-03   -.28675E-03   -.93586E-03
-.47620E-03   .19812E-03   -.25707E-03   -.15761E-02   -.30471E-03

ITERATION NO.   6
YSQ = 51.970   DET(C) = .45800E-04   AMP = .00000   AP = 1.0000
REGRESSION PARAMETERS
2865.3   117.67   497.85   .99782E-03   .96559E-04   .10785E-02   -.17358E-03
.14220E-03   8.0000   .22321   .89599   40.109   10.042   17.454

SOLUTION CONVERGED IN   6   ITERATIONS

ERROR VARIANCE = 1.2374
ESTIMATED SUM OF SQUARED ERRORS = 51.970
CORRELATION COEFFICIENT = .99033

ESTIMATED PARAMETER DATA

PAR. NO.	PAR.	STD. DEV.
1	2865.3	1293.6
2	117.67	266.83
3	497.85	426.73
4	.99782E-03	.34440E-03
5	.96559E-04	.22643E-03
6	.10785E-02	.99504E-03
7	-.17358E-03	.50163E-03
8	.33062E-03	.73586E-03
9	.14220E-03	.71033E-03

10	8.0000	.88990
11	.22321	.59676
12	.89599	1.1288
13	40.109	.68267
14	10.042	.66620
15	17.454	1.1536

VARIANCE-COVARIANCE MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.16734E+07	92126.	.19755E+06	-.12629	.81201E-01	.47075	-.10060	.25244	.54361E-01	283.68	235.55	437.83	22.301	26.987	-23.100
2	92126.	71196.	-12931.	.18984E-01	.52827E-01	-.17544E-01	-.84450E-01	.19080	-.55455E-02	11.633	129.38	9.4207	-5.5050	1.8507	-5.3819
3	.19755E+06	-12931.	.18210E+06	.17659E-01	-.10515E-01	.27608	.88590E-01	-.35979E-01	-.22092E-01	49.298	-15.206	274.15	1.1342	16.104	-10.743
4	-.12629	.18984E-01	.17659E-01	.11861E-06	-.10370E-07	.20261E-07	.47110E-07	.51140E-07	.75899E-08	-.98759E-04	.29647E-04	-.12778E-04	-.19591E-04	-.12935E-04	.12131E-04
5	.81201E-01	.52827E-01	-.10515E-01	.10370E-07	.51269E-07	-.13461E-07	-.69361E-07	.13232E-06	-.42375E-08	.95671E-05	.12079E-03	.88531E-05	.36933E-05	-.16303E-05	-.29205E-05
6	.47075	-.17544E-01	.27608	.20261E-07	-.13461E-07	.99011E-06	.14834E-06	-.49792E-07	.41018E-06	.10675E-03	-.11589E-03	.72802E-03	.19509E-04	.10117E-04	-.99534E-04
7	-.10060	-.84450E-01	.88590E-01	.47110E-07	-.69361E-07	.14834E-06	.25163E-06	-.22998E-06	.34130E-07	-.17185E-04	-.15827E-03	.25834E-04	-.17699E-04	-.21829E-04	.28963E-04
8	.25244	.19080	-.35979E-01	.51140E-07	.13232E-06	-.49792E-07	-.22998E-06	.54149E-06	-.16027E-07	-.32659E-04	-.30324E-03	.24921E-04	-.43658E-04	.13911E-04	-.19299E-04
9	.54361E-01	-.55455E-02	-.22092E-01	.75899E-08	-.42375E-08	.41018E-06	.34130E-07	-.16027E-07	.50456E-06	.14038E-04	-.73381E-05	.24132E-04	.18016E-04	-.49939E-04	-.89174E-04
10	283.68	11.633	49.298	-.98759E-04	.95671E-05	.10675E-03	-.17185E-04	-.32659E-04	.14038E-04	.79192	.22128E-01	.88817E-01	-.50764E-16	-.52352E-16	-.19250E-15
11	235.55	129.38	-15.206	.29647E-04	.12079E-03	-.11589E-03	-.15827E-03	-.30324E-03	-.73381E-05	.22128E-01	.35613	.35953E-01	.12982E-01	-.20156E-02	-.74691E-02
12	437.83	9.4207	274.15	-.12778E-04	.88531E-05	.72802E-03	.25834E-04	.24921E-04	.24132E-04	.88817E-01	.35953E-01	1.2742	.59038E-02	.80067E-01	.16715
13	22.301	-5.5050	1.1342	-.19591E-04	.36933E-05	.19509E-04	-.17699E-04	-.43658E-04	.18016E-04	-.50764E-16	.12982E-01	.59038E-02	.46604	.19602	.99711E-01
14	26.987	1.8507	16.104	-.12935E-04	-.16303E-05	.10117E-04	-.21829E-04	.13911E-04	-.49939E-04	-.52352E-16	-.20156E-02	.80067E-01	.19602	.44382	-.23097
15	-23.100	-5.3819	-10.743	.12131E-04	-.29205E-05	-.99534E-04	.28963E-04	-.19299E-04	-.89174E-04	-.19250E-15	-.74691E-02	.16715	.99711E-01	-.23097	1.3307

Figure 4.3-10—Continued

CORRELATION MATRIX

	1	2	3	4	5	6	7	8	9	10
1	1.0000									
2	.26690	1.0000								
3	.35786	.11357	1.0000							
4	-.28346	.20659	.12016	1.0000						
5	.27722	.87438	-.10882	.13298	1.0000					
6	.36572	-.66079E-01	.65018	.59123E-01	-.59744E-01	1.0000				
7	-.15503	.63094	.41385	.27269	.61067	.29720	1.0000			
8	.26519	.97176	.11458	.20179	.79416	-.68002E-01	-.62304	1.0000		
9	.59160E-01	-.29259E-01	-.72883E-01	.31026E-01	-.26347E-01	.58033	.95784E-01	-.30661E-01	1.0000	
10	.24642	.48992E-01	.12982	.32224	.47480E-01	.12056	-.38497E-01	.49873E-01	.22208E-01	1.0000
11	.30513	.81254	.59713E-01	.14425	.89391	-.19516E-01	-.52871	.69054	-.17311E-01	.41668E-01
12	.29984	.31278E-01	.56915	-.32869E-01	.34638E-01	.64818	.45625E-01	.30003E-01	.30096E-01	.88418E-01
13	.25253E-01	-.30222E-01	.38934E-02	-.83327E-01	-.23894E-01	.28720E-01	-.51685E-01	-.86908E-01	.37153E-01	-.83562E-16
14	.31315E-01	.10412E-01	.56647E-01	-.56376E-01	-.10808E-01	.15261E-01	-.65321E-01	-.28376E-01	-.10553	.88305E-16
15	-.15480E-01	-.17485E-01	-.21823E-01	-.30535E-01	-.11181E-01	-.86713E-01	-.50051E-01	-.22735E-01	-.10883	-.18752E-15

	11	12	13	14	15
1	.30513	.29984	.25253E-01	.31315E-01	-.15480E-01
2	.81254	.31278E-01	-.30222E-01	.10412E-01	-.17485E-01
3	-.59713E-01	.56915	-.38934E-02	.56647E-01	-.21823E-01
4	.14425	-.32869E-01	-.83327E-01	-.56376E-01	-.30535E-01
5	.89391	.34638E-01	-.23894E-01	-.10808E-01	-.11181E-01
6	-.19516E-01	.64818	-.28720E-01	.15261E-01	-.86713E-01
7	-.52871	.45625E-01	-.51685E-01	-.65321E-01	-.50051E-01
8	.69054	.30003E-01	-.86908E-01	-.28376E-01	-.22735E-01
9	-.17311E-01	.30096E-01	.37153E-01	-.10553	-.10883
10	.41668E-01	.88418E-01	-.83562E-16	.88305E-16	-.18752E-15
11	1.0000	.53372E-01	-.31867E-01	-.50698E-02	-.10850E-01
12	.53372E-01	1.0000	.76614E-02	.10647	-.12836
13	-.31867E-01	.76614E-02	1.0000	-.43102	.12662
14	-.50698E-02	.10647	-.43102	1.0000	-.30054
15	-.10850E-01	-.12836	.12662	-.30054	1.0000

HEAD RESIDUALS

OBS. NO.	PREDICTED VALUE	OBSERVED VALUE	WEIGHTED RESIDUAL
1	29.988	29.510	.47777
2	26.826	26.380	.44644
3	25.687	25.160	.52684
4	25.055	26.810	-1.7551
5	24.409	23.820	.58942
6	22.257	23.590	-1.3331
7	16.544	16.680	-.13645
8	16.426	15.310	1.1157
9	28.095	28.270	-.17495

10	26.249	25.170	1.0794
11	25.409	26.940	-1.5305
12	24.720	23.590	1.1296
13	23.386	22.530	.85609
14	17.048	17.890	-.84160
15	15.527	15.870	-.34305
16	16.218	15.980	.23841
17	30.499	31.250	-.75148
18	27.340	27.500	-.15980
19	26.114	26.220	-.10621
20	25.199	25.650	-.45050
21	24.155	24.350	-.19520
22	21.790	22.480	.69007
23	15.618	15.850	-.23236
24	15.152	15.710	-.55827
25	29.315	29.360	-.45113E-01
26	27.443	27.760	-.31699
27	26.121	25.510	.61129
28	24.861	24.430	.43100
29	23.194	22.460	.73447
30	16.819	17.160	-.34134
31	15.154	15.740	-.58562
32	13.748	14.690	-.94206
33	36.304	36.050	.25364
34	33.923	31.640	2.2829
35	31.303	32.910	-1.6070
36	25.933	25.050	.88337
37	24.185	25.470	-1.2851
38	21.558	20.390	1.1678
39	15.529	14.400	1.1286
40	14.428	14.370	.57542E-01
41	37.152	38.660	-1.5078
42	34.692	34.510	.18165
43	27.906	27.810	.96020E-01
44	25.453	25.980	-.52672
45	22.809	22.160	.64860
46	16.934	15.730	1.2037
47	14.804	14.790	-.14347E-01
48	11.277	10.480	.79746
49	40.109	41.280	-1.1707
50	36.100	34.820	1.2804
51	32.091	32.000	.91448E-01
52	28.082	26.970	1.1125
53	24.074	25.670	-1.5965
54	20.065	18.120	1.9446
55	16.056	18.400	-2.3444
56	12.047	11.900	.14670

Figure 4.3-10—Continued

FINAL COMPUTED NODAL HEAD ARRAY

	1	2	3	4	5	6	7	8	9	10
7	40.109	38.105	36.100	34.096	32.091	30.087	28.082	26.078	24.074	22.069
6	38.593	37.152	35.970	34.092	32.677	27.906	26.682	25.453	24.186	22.809
5	36.304	34.897	33.923	32.961	31.303	26.879	25.933	25.086	24.185	23.088
4	31.294	29.315	28.170	27.443	26.852	26.121	25.473	24.861	24.157	23.194
3	30.499	28.493	27.340	26.630	26.114	25.640	25.199	24.741	24.155	23.267
2	30.105	28.095	26.943	26.249	25.782	25.409	25.074	24.720	24.220	23.386
1	29.988	27.977	26.826	26.139	25.687	25.345	25.055	24.806	24.409	23.621

	11	12	13	14	15	16
7	20.065	18.060	16.056	14.051	12.047	10.042
6	21.189	16.934	15.585	14.804	13.839	11.277
5	21.558	16.804	15.529	15.033	14.428	12.513
4	21.698	16.819	15.552	15.154	14.800	13.748
3	21.790	16.875	15.618	15.282	15.152	14.983
2	21.935	17.048	15.814	15.527	15.585	16.218
1	22.257	17.722	16.544	16.282	16.426	17.454

NODAL SENSITIVITY ARRAYS

PARAMETER NUMBER 1

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	-.88137E-04	-.66179E-04	-.45801E-04	-.39830E-04	-.55354E-04	-.12696E-03	.28778E-04	.23754E-04	-.13650E-04	.49991E-05
5	-.29267E-03	-.18520E-03	-.11486E-03	-.90919E-04	-.10016E-03	-.17191E-03	.28266E-05	.11947E-04	-.10330E-04	.42481E-04
4	-.95285E-03	-.41941E-03	-.23200E-03	-.18360E-03	-.16481E-03	-.78069E-04	.89593E-05	.41699E-05	-.77975E-05	.64948E-04
3	-.74906E-03	-.18793E-03	.80211E-05	.52822E-04	.47692E-04	.42387E-04	.31497E-04	.77544E-06	-.11412E-04	.69590E-04
2	-.66908E-03	-.10218E-03	.96560E-04	.13846E-03	.12054E-03	.84288E-04	.39456E-04	-.10571E-04	-.30645E-04	.51074E-04
1	-.64887E-03	-.80986E-04	.11830E-03	.15921E-03	.13755E-03	.93228E-04	.36182E-04	-.42638E-04	-.88759E-04	-.12934E-04

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	.23800E-03	.53870E-04	.12424E-04	.28988E-05	.64893E-06	.00000
5	.33628E-03	.81713E-04	.19828E-04	.47999E-05	.11021E-05	.00000
4	.37635E-03	.93820E-04	.23314E-04	.57642E-05	.13465E-05	.00000
3	.38737E-03	.97069E-04	.24326E-04	.60791E-05	.14352E-05	.00000
2	.37098E-03	.92960E-04	.23522E-04	.59547E-05	.14231E-05	.00000
1	.29523E-03	.81573E-04	.21784E-04	.56894E-05	.13844E-05	.00000

PARAMETER NUMBER

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	-.14387E-01	-.10074E-01	-.11709E-01	-.14373E-01	-.14515E-01	.92843E-03	.34211E-03	.13590E-03	.58013E-04	.27361E-04
5	-.18117E-01	-.13644E-01	-.14908E-01	-.17039E-01	-.16410E-01	.95459E-03	.42323E-03	.19079E-03	.88987E-04	.44680E-04
4	.14446E-02	.19155E-02	.19316E-02	.16777E-02	.12208E-02	.66243E-03	.35277E-03	.18145E-03	.93448E-04	.50526E-04
3	.97440E-03	.10626E-02	.10513E-02	.92156E-03	.70059E-03	.44828E-03	.26676E-03	.15198E-03	.85361E-04	.49435E-04
2	.66706E-03	.67908E-03	.65545E-03	.57752E-03	.45560E-03	.31938E-03	.20690E-03	.12724E-03	.76302E-04	.46532E-04
1	.56792E-03	.56764E-03	.54208E-03	.47852E-03	.38354E-03	.27795E-03	.18624E-03	.11806E-03	.72642E-04	.45190E-04

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	.16279E-04	.38632E-05	.92257E-06	.22088E-06	.50347E-07	.00000
5	.27398E-04	.66173E-05	.16039E-05	.38889E-06	.89535E-07	.00000
4	.32134E-04	.79335E-05	.19603E-05	.48333E-06	.11279E-06	.00000
3	.32529E-04	.82024E-05	.20656E-05	.51793E-06	.12253E-06	.00000
2	.31401E-04	.80433E-05	.20547E-05	.52184E-06	.12476E-06	.00000
1	.30787E-04	.79331E-05	.20376E-05	.52000E-06	.12481E-06	.00000

PARAMETER NUMBER 3

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	-.15078E-05	-.18450E-05	-.30940E-05	-.62466E-05	-.14283E-04	-.35985E-04	-.48512E-04	-.88405E-04	-.18033E-03	-.39214E-03
5	-.35813E-05	-.42956E-05	-.68287E-05	-.12747E-04	-.26645E-04	-.62814E-04	-.86546E-04	-.15556E-03	-.30354E-03	-.60987E-03
4	-.71714E-05	-.84601E-05	-.12794E-04	-.21753E-04	-.38648E-04	-.69937E-04	-.10943E-03	-.19791E-03	-.37406E-03	-.71454E-03
3	-.73464E-05	-.86613E-05	-.13074E-04	-.22145E-04	-.39026E-04	-.69289E-04	-.12142E-03	-.22150E-03	-.41048E-03	-.76120E-03
2	-.74492E-05	-.87801E-05	-.13247E-04	-.22436E-04	-.39611E-04	-.70893E-04	-.12773E-03	-.23331E-03	-.42806E-03	-.78264E-03
1	-.74834E-05	-.88203E-05	-.13308E-04	-.22552E-04	-.39879E-04	-.71624E-04	-.12977E-03	-.23716E-03	-.43483E-03	-.79741E-03

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	-.91488E-03	.14872E-02	.63461E-03	-.31842E-03	-.10832E-02	.00000
5	-.12416E-02	.16125E-02	.77206E-03	-.90615E-04	-.73796E-03	.00000
4	-.13618E-02	.16511E-02	.87756E-03	.15751E-03	-.31308E-03	.00000
3	-.14022E-02	.17147E-02	.10011E-02	.39898E-03	.92917E-04	.00000
2	-.14164E-02	.17411E-02	.10674E-02	.55652E-03	.40007E-03	.00000
1	-.14671E-02	.10314E-02	.31150E-03	-.15526E-03	-.15560E-03	.00000

Figure 4.3-10—Continued

PARAMETER NUMBER 4

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	-215.11	-188.73	-131.28	-60.679	28.425	190.68	-87.083	20.249	224.47	344.43
5	-659.90	-563.73	-383.68	-189.76	7.0766	295.72	-2.5536	108.84	327.25	435.34
4	-1839.7	-1486.2	-965.48	-477.82	-100.25	58.649	-2.9031	158.15	369.78	454.73
3	-2381.7	-2014.5	-1475.8	-969.80	-559.28	-265.57	-62.017	163.44	373.52	442.85
2	-2568.6	-2196.9	-1650.5	-1135.3	-707.56	-368.73	-95.016	150.65	353.11	411.00
1	-2611.5	-2238.8	-1690.3	-1172.7	-740.48	-392.23	-107.12	135.69	328.13	378.88

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	73.063	16.265	3.6788	.84037	.18435	.00000
5	99.174	23.537	5.5765	1.3181	.29626	.00000
4	99.187	24.104	5.8472	1.4138	.32409	.00000
3	81.573	20.043	4.9562	1.2274	.28812	.00000
2	49.088	12.940	3.4397	.90946	.22489	.00000
1	16.186	7.2124	2.4016	.71519	.18895	.00000

PARAMETER NUMBER 5

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	-10145.	-6628.8	-1888.6	3026.2	6864.6	566.80	227.02	94.150	40.969	19.439
5	-5092.0	-1632.6	2740.1	6990.7	9748.2	796.11	326.18	141.39	64.549	31.984
4	-45.157	223.72	574.68	847.68	907.44	468.64	253.76	129.89	66.328	35.590
3	98.952	180.21	316.85	422.38	428.36	290.68	178.68	103.29	58.378	33.893
2	115.00	144.04	200.46	243.68	242.05	188.26	129.19	82.183	50.321	31.085
1	112.99	130.65	167.03	194.69	192.14	156.68	112.60	74.543	47.156	29.857

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	11.572	2.7438	.65424	.15633	.35567E-01	.00000
5	19.493	4.6912	1.1335	.27406	.62952E-01	.00000
4	22.546	5.5529	1.3691	.33692	.78502E-01	.00000
3	22.320	5.6298	1.3691	.35558	.84129E-01	.00000
2	21.092	5.4194	1.3881	.35330	.84608E-01	.00000
1	20.494	5.3038	1.3673	.35003	.84219E-01	.00000

PARAMETER NUMBER 6

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	-.56423	-.68739	-1.1401	-2.2681	-5.1014	-12.653	-16.846	-29.872	-58.764	-121.97
5	-1.3462	-1.6105	-2.5426	-4.6962	-9.6800	-22.485	-30.726	-54.278	-103.67	-203.50
4	-2.7065	-3.1904	-4.8147	-8.1564	-14.398	-25.743	-39.990	-71.738	-134.24	-253.57
3	-2.7752	-3.2709	-4.9332	-8.3460	-14.685	-26.034	-45.675	-83.418	-154.71	-286.92
2	-2.8186	-3.3234	-5.0189	-8.5148	-15.074	-27.098	-49.159	-90.592	-168.12	-312.00
1	-2.8337	-3.3421	-5.0515	-8.5846	-15.246	-27.559	-50.386	-93.209	-173.72	-325.89

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	-268.07	-914.35	-415.10	52.738	427.32	.00000
5	-405.64	-1071.0	-523.05	-80.647	249.85	.00000
4	-477.90	-1137.6	-596.95	-212.15	43.367	.00000
3	-527.28	-1215.7	-699.11	-364.37	-172.51	.00000
2	-575.68	-1366.5	-893.16	-597.67	-438.41	.00000
1	-619.29	-1626.4	-1207.6	-930.66	-757.77	.00000

PARAMETER NUMBER 7

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	134.36	135.15	139.27	155.18	213.78	423.78	430.34	425.60	411.17	380.06
5	377.63	378.13	381.27	395.29	451.98	671.75	672.85	660.10	633.20	580.30
4	930.45	929.45	925.93	917.84	898.80	846.13	814.48	789.59	752.72	686.06
3	957.35	956.35	952.96	945.73	931.75	907.47	883.96	857.00	815.12	740.83
2	970.56	969.64	966.62	960.66	950.44	935.00	915.68	888.75	844.74	766.78
1	974.52	973.65	970.80	965.33	956.30	942.95	924.83	898.02	853.43	774.40

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	312.13	72.205	16.967	4.0234	.91257	.00000
5	474.93	116.08	28.379	6.9270	1.6030	.00000
4	560.30	140.16	34.999	8.7025	2.0437	.00000
3	604.32	152.75	38.551	9.6850	2.2947	.00000
2	625.13	158.73	40.256	10.164	2.4192	.00000
1	631.24	160.48	40.760	10.306	2.4564	.00000

Figure 4.3–10—Continued

PARAMETER NUMBER 8

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	4732.5	4716.7	4642.4	4379.9	3464.1	252.34	97.684	40.087	17.476	8.3539
5	4879.2	4861.7	4781.4	4505.7	3563.7	306.14	132.31	59.145	27.548	13.847
4	564.00	558.38	538.54	493.77	400.12	214.34	112.34	57.222	29.313	15.808
3	326.10	321.44	305.87	274.50	220.26	142.40	84.596	48.013	26.879	15.531
2	211.05	207.33	195.47	173.67	140.69	99.968	65.085	40.069	24.018	14.635
1	176.90	173.55	163.05	144.48	117.83	86.505	58.366	37.109	22.851	14.213

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	5.0002	1.1905	.28503	.68379E-01	.15610E-01	.00000
5	8.5005	2.0548	.49846	.12095	.27865E-01	.00000
4	10.045	2.4792	.61250	.15101	.35241E-01	.00000
3	10.210	2.5731	.64770	.16235	.38401E-01	.00000
2	9.8706	2.5275	.64543	.16387	.39165E-01	.00000
1	9.6806	2.4939	.64036	.16338	.39203E-01	.00000

PARAMETER NUMBER 9

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	.32358	.39574	.66278	1.3357	3.0478	7.6628	10.311	18.707	37.918	81.735
5	.76895	.92207	1.4647	2.7309	5.6992	13.413	18.463	33.114	64.432	129.04
4	1.5405	1.8172	2.7476	4.6701	8.2923	14.988	23.441	42.381	80.082	152.97
3	1.5782	1.8606	2.8084	4.7565	8.3819	14.882	26.092	47.641	88.419	164.39
2	1.6005	1.8865	2.8464	4.8218	8.5150	15.246	27.487	50.256	92.352	169.39
1	1.6080	1.8953	2.8600	4.8477	8.5749	15.407	27.925	51.041	93.496	170.79

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	188.43	579.24	657.93	650.43	537.26	.00000
5	261.96	716.90	818.77	807.03	654.01	.00000
4	291.64	752.02	858.23	844.50	679.98	.00000
3	303.99	761.79	868.07	853.47	685.86	.00000
2	308.99	764.70	870.57	855.62	687.20	.00000
1	310.34	765.36	871.07	856.02	687.45	.00000

PARAMETER NUMBER 10

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	.53074E-01	.41784E-01	.27227E-01	.16052E-01	.84549E-02	.29026E-02	.22072E-02	.14298E-02	.86627E-03	.52587E-03
5	.17238	.12120	.73462E-02	.41729E-01	.21818E-01	.77685E-02	.52650E-02	.31434E-02	.18110E-02	.10671E-02
4	.53580	.29685	.16411	.89902E-01	.47263E-01	.19187E-01	.97739E-02	.51620E-02	.27977E-02	.15951E-02
3	.55281	.30544	.16869	.92611E-01	.49752E-01	.25101E-01	.12884E-01	.67301E-02	.35968E-02	.20262E-02
2	.55909	.30976	.17133	.94342E-01	.51351E-01	.27318E-01	.14418E-01	.76205E-02	.40852E-02	.22995E-02
1	.56074	.31104	.17217	.94920E-01	.51870E-01	.27914E-01	.14869E-01	.79014E-02	.42464E-02	.23920E-02

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	.35319E-03	.90109E-04	.22920E-04	.57962E-05	.13805E-05	.00000
5	.70802E-03	.17946E-03	.45409E-04	.11435E-04	.27152E-05	.00000
4	.10445E-02	.26290E-03	.66152E-04	.16585E-04	.39246E-05	.00000
3	.13190E-02	.33078E-03	.82964E-04	.20741E-04	.48968E-05	.00000
2	.14952E-02	.37456E-03	.93835E-04	.23431E-04	.55261E-05	.00000
1	.15554E-02	.38960E-03	.97581E-04	.24359E-04	.57435E-05	.00000

PARAMETER NUMBER 11

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	5.2629	1.5043	.43378	.12627	.35285E-01	.21657E-02	.12477E-02	.70540E-03	.39528E-03	.22865E-03
5	5.3522	1.5578	.46126	.13987	.41723E-01	.42803E-02	.25542E-02	.14242E-02	.78468E-03	.44855E-03
4	.41301	.19460	.92953E-01	.45246E-01	.21772E-01	.83311E-02	.41539E-02	.21481E-02	.11439E-02	.64351E-03
3	.16773	.11507	.67522E-01	.37391E-01	.19890E-01	.98661E-02	.50281E-02	.26184E-02	.13976E-02	.78690E-03
2	.86163E-01	.70502E-01	.48177E-01	.29924E-01	.17453E-01	.96505E-02	.52243E-02	.28116E-02	.15276E-02	.86830E-03
1	.65914E-01	.57139E-01	.41532E-01	.27093E-01	.16425E-01	.94182E-02	.52256E-02	.28549E-02	.15648E-02	.89380E-03

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	.15038E-03	.37906E-04	.95429E-05	.23921E-05	.56579E-06	.00000
5	.29351E-03	.73788E-04	.18538E-04	.46398E-05	.10963E-05	.00000
4	.41870E-03	.10498E-03	.26325E-04	.65802E-05	.15534E-05	.00000
3	.51215E-03	.12843E-03	.32210E-04	.80521E-05	.19010E-05	.00000
2	.56718E-03	.14248E-03	.35782E-04	.89538E-05	.21153E-05	.00000
1	.58495E-03	.14709E-03	.36966E-04	.92548E-05	.21873E-05	.00000

Figure 4.3-10—Continued

PARAMETER NUMBER 12

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	.76330E-03	.88767E-03	.13018E-02	.21588E-02	.38679E-02	.75714E-02	.85541E-02	.10375E-01	.12020E-01	.12432E-01
5	.19056E-02	.22156E-02	.32313E-02	.52410E-02	.89224E-02	.16324E-01	.19248E-01	.24539E-01	.29456E-01	.31222E-01
4	.39952E-02	.46600E-02	.68241E-02	.10962E-01	.17595E-01	.26030E-01	.34276E-01	.47619E-01	.60297E-01	.66227E-01
3	.41516E-02	.48696E-02	.72470E-02	.11991E-01	.20371E-01	.34243E-01	.56139E-01	.87944E-01	.11888	.13594
2	.43173E-02	.51151E-02	.78265E-02	.13558E-01	.24748E-01	.46355E-01	.87636E-01	.15976	.23271	.27533
1	.43904E-02	.52278E-02	.81107E-02	.14386E-01	.27324E-01	.54928E-01	.11880	.28637	.45789	.54126

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	.10815E-01	.37378E-02	.17059E-02	.10118E-02	.54211E-03	.00000
5	.27643E-01	.10525E-01	.57682E-02	.39903E-02	.23306E-02	.00000
4	.60505E-01	.27741E-01	.19343E-01	.15492E-01	.98366E-02	.00000
3	.13066	.80666E-01	.70257E-01	.62343E-01	.42825E-01	.00000
2	.28726	.26868	.27081	.25580	.19186	.00000
1	.64700	1.0178	1.0750	1.0521	.88426	.00000

PARAMETER NUMBER 13

	1	2	3	4	5	6	7	8	9	10
7	1.0000	.93333	.86667	.80000	.73333	.66667	.60000	.53333	.46667	.40000
6	.39285	.37685	.35397	.33205	.31917	.33899	.32197	.29224	.25767	.22049
5	.14081	.13716	.13142	.12816	.13480	.16994	.16873	.15773	.14100	.12079
4	.11845E-01	.12381E-01	.14471E-01	.19753E-01	.32807E-01	.68297E-01	.84042E-01	.83858E-01	.76883E-01	.66329E-01
3	.82604E-02	.88185E-02	.10777E-01	.15034E-01	.23211E-01	.36528E-01	.44537E-01	.45972E-01	.43113E-01	.37559E-01
2	.64353E-02	.69259E-02	.85357E-02	.11629E-01	.16555E-01	.22785E-01	.27109E-01	.28385E-01	.27049E-01	.23781E-01
1	.58695E-02	.63253E-02	.77821E-02	.10441E-01	.14358E-01	.18881E-01	.22168E-01	.23290E-01	.22332E-01	.19716E-01

	11	12	13	14	15	16
7	.33333	.26667	.20000	.13333	.66667E-01	.00000
6	.18025	.87171E-01	.55050E-01	.34922E-01	.17220E-01	.00000
5	.97407E-01	.35487E-01	.17317E-01	.97065E-02	.45785E-02	.00000
4	.53174E-01	.16900E-01	.65239E-02	.30275E-02	.13025E-02	.00000
3	.30087E-01	.90087E-02	.29788E-02	.11334E-02	.42423E-03	.00000
2	.19080E-01	.55774E-02	.16997E-02	.55725E-03	.17977E-03	.00000
1	.15836E-01	.45992E-02	.13659E-02	.42289E-03	.12691E-03	.00000

PARAMETER NUMBER 14

	1	2	3	4	5	6	7	8	9	10
7	.00000	.66667E-01	.13333	.20000	.26667	.33333	.40000	.46667	.53333	.60000
6	.15372E-01	.32033E-01	.57705E-01	.88387E-01	.12946	.20556	.22813	.25909	.28951	.31132
5	.10066E-01	.14731E-01	.24523E-01	.39070E-01	.63709E-01	.11582	.12728	.14225	.15507	.15964
4	.37082E-02	.44186E-02	.67529E-02	.11795E-01	.22878E-01	.50658E-01	.67216E-01	.77109E-01	.82863E-01	.82525E-01
3	.34160E-02	.39940E-02	.59039E-02	.97854E-02	.16880E-01	.28298E-01	.37221E-01	.42946E-01	.45599E-01	.44328E-01
2	.31570E-02	.36280E-02	.51388E-02	.79695E-02	.12434E-01	.18283E-01	.23385E-01	.26804E-01	.28133E-01	.26878E-01
1	.30552E-02	.34851E-02	.48427E-02	.72941E-02	.10934E-01	.15379E-01	.19372E-01	.22085E-01	.23046E-01	.21857E-01

	11	12	13	14	15	16
7	.66667	.73333	.80000	.86667	.93333	1.0000
6	.30453	.21532	.21569	.24872	.35986	.83333
5	.14521	.74009E-01	.65795E-01	.91276E-01	.20329	.66667
4	.72001E-01	.29976E-01	.24680E-01	.44954E-01	.13558	.50000
3	.37743E-01	.14128E-01	.11438E-01	.25624E-01	.87696E-01	.33333
2	.22552E-01	.79464E-02	.60884E-02	.13898E-01	.46144E-01	.16667
1	.18234E-01	.62253E-02	.42550E-02	.78324E-02	.16239E-01	.00000

PARAMETER NUMBER 15

	1	2	3	4	5	6	7	8	9	10
7	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
6	.45762E-06	.55363E-06	.90191E-06	.17496E-05	.38206E-05	.91972E-05	.11942E-04	.19965E-04	.36036E-04	.66041E-04
5	.10996E-05	.13103E-05	.20462E-05	.37146E-05	.74778E-05	.16918E-04	.22763E-04	.38833E-04	.70682E-04	.13024E-03
4	.22244E-05	.26194E-05	.39408E-05	.66398E-05	.11607E-04	.20362E-04	.31282E-04	.55442E-04	.10223E-03	.18999E-03
3	.22841E-05	.26909E-05	.40539E-05	.68467E-05	.12022E-04	.21278E-04	.37452E-04	.68727E-04	.12839E-03	.24109E-03
2	.23253E-05	.27432E-05	.41486E-05	.70553E-05	.12539E-04	.22685E-04	.41559E-04	.77546E-04	.14622E-03	.27719E-03
1	.23404E-05	.27628E-05	.41863E-05	.71427E-05	.12760E-04	.23263E-04	.43019E-04	.80673E-04	.15267E-03	.29070E-03

	11	12	13	14	15	16
7	.00000	.00000	.00000	.00000	.00000	.00000
6	.12088E-03	.69575E-03	.27812E-02	.10897E-01	.42624E-01	.16667
5	.23958E-03	.13876E-02	.55542E-02	.21781E-01	.85235E-01	.33333
4	.35253E-03	.20657E-02	.82948E-02	.32603E-01	.12777	.50000
3	.45274E-03	.27011E-02	.10918E-01	.43167E-01	.16994	.66667
2	.52725E-03	.32166E-02	.13153E-01	.52679E-01	.21017	.83333
1	.55694E-03	.34503E-02	.14283E-01	.58303E-01	.23983	1.0000

Figure 4.3–10—Continued

SCALED AND ORTHOCONALIZED SENSITIVITY MATRIX
OBSERVATION NUMBER AND VALUES FOR EACH PARAMETER

PARAMETER NOS.

OBS. NO.	1	2	3	4	5	6	7	8
1	-.48486	.10012	-.88313E-01	-.19353	.28397E-01	.11681	-.71251E-01	-.30847E-01
2	-.88399E-01	.86688E-03	-.12077E-01	-.36799	.81978E-02	.92902E-01	-.37863E-01	-.10784E-01
3	.10279	-.62999E-02	.85497E-02	.19925	.10257E-01	.59571E-01	.13319	-.11047E-01
4	-.27036E-01	.87074E-03	-.25771E-01	-.34406E-01	.77748E-02	-.16027E-01	.17937	-.18262E-01
5	-.66324E-01	.13546E-01	.11370	.10690	.69793E-02	.14516	.16006	-.18556E-01
6	.22061	-.36989E-01	.30457	.11209	-.98094E-02	.27365	.87243E-01	.38504E-02
7	.16278E-01	-.27362E-02	.75860E-01	-.13250E-01	-.27104E-03	.37791	.12464	.92882E-02
8	-.10345E-02	-.17403E-03	-.36267E-01	.13396E-02	-.21064E-03	-.27755	.12024	.90602E-02
9	-.76352E-01	.33228E-01	-.16223E-01	-.36282	.13719E-01	.41578E-01	-.19559E-01	-.10764E-01
10	.10347	-.67224E-03	.12581E-01	-.27327	-.14384E-01	.77535E-01	.93922E-01	-.28296E-02
11	.62983E-01	-.13619E-02	-.57208E-02	-.10549	.12057E-01	-.22372E-01	.16289	-.11724E-01
12	.78991E-02	.51252E-02	-.56173E-01	.36031E-01	.71357E-02	-.64501E-01	.18224	-.18494E-01
13	.38164E-01	-.51786E-02	.17661	.63478E-01	.72899E-03	-.18764	.96243E-01	-.98763E-02
14	.69463E-01	-.11696E-01	.42019	.61148E-01	-.73223E-03	-.26174	.30864E-01	.25158E-02
15	.44496E-02	-.74899E-03	.13116	-.94200E-02	-.44404E-03	-.14667	-.22347E-01	.16658E-02
16	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
17	.55973	.12502	.10194	-.10463	.33366E-01	.15744	.51240E-01	-.41809E-02
18	.59936E-02	.30105E-01	-.29543E-02	-.27720	.26133E-01	.49428E-01	.54632E-01	.19606E-01
19	.35638E-01	.14626E-01	-.34547E-02	.12426	.32773E-01	.24332E-01	.14430	-.11134E-01
20	.23535E-01	.38568E-02	-.24510E-01	.23826E-01	.13341E-01	-.16764E-01	.17587	-.11128E-01
21	-.85274E-02	.39930E-01	-.97751E-01	.80158E-01	.53928E-02	.11314	.16245	-.14950E-01
22	.28946	.48766E-01	.27714	.14230	.12564E-01	.20817	.65493E-01	.56575E-02
23	.18178E-01	-.30618E-02	.23775	-.22473E-01	.41211E-03	.12972	.16959E-01	-.12634E-02
24	-.10724E-02	-.18062E-03	.21957E-01	-.17465E-02	.61644E-04	.49475E-01	.13621E-01	-.10264E-02
25	.31340	.11057	.59339E-01	-.30343E-02	-.38305E-01	.86424E-01	.44078E-01	.60891E-01
26	-.13719	.73254E-01	-.30966E-01	.48271E-01	.76989E-01	.49781E-01	.14654	.66978E-01
27	-.58336E-01	.29640E-01	-.27338E-01	.30433E-01	.40859E-01	.44313E-01	.18463	.15586E-01
28	.31159E-02	.48384E-02	-.45976E-01	.64542E-01	.10511E-01	.49994E-01	.16651	-.11471E-01
29	.48532E-01	.68414E-02	.15882	-.58337E-01	-.77344E-03	.15751	.97825E-01	-.73610E-02
30	.70106E-01	-.11809E-01	.39923	.37663E-01	-.83040E-03	.19214	.50137E-01	-.34587E-02
31	.43072E-02	-.72567E-03	.37663E-01	-.42920E-02	-.10301E-04	.55111E-01	-.10748E-01	.82561E-03
32	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
33	-.21870	.49895	-.23188E-01	-.23453E-01	.46622	.35522E-01	.35154E-01	-.10854
34	-.85831E-01	.42675	-.32858E-02	-.46393E-01	.12394	.38490E-02	.39904E-02	.60535E-01
35	-.74841E-01	.47312	-.46130E-01	.17499E-01	.63320	-.72934E-02	-.21677E-02	-.28154E-01
36	.21121E-01	.12171E-01	-.20284E-01	.77644E-04	.26193E-01	.19469E-01	.13814	.63711E-01
37	-.77194E-02	.39616E-02	-.72558E-01	.69708E-01	.58646E-02	-.82403E-01	.13392	-.10142E-01
38	.25128	-.42359E-01	-.24629	-.11793	-.10878E-01	.16625	-.60479E-01	.64329E-02
39	.14816E-01	-.24980E-02	.18349	-.17486E-01	-.28823E-03	.94202E-02	-.15200E-01	-.10613E-02
40	.82357E-03	-.13884E-03	.17273	.87086E-02	-.84439E-03	.98069E-02	.41607E-01	.31532E-02
41	-.49452E-01	.28984	-.54730E-04	-.23377E-01	-.54244	.48665E-02	.55888E-01	.10402
42	-.29762E-01	.42055	.63147E-02	-.19010E-01	.14605	.15402E-01	-.13436E-01	.42213E-01
43	-.94867E-01	.43794E-01	-.26109E-01	.95067E-01	.51478E-01	-.56935E-01	.10631	-.42625E-02
44	.17750E-01	.97528E-03	-.17684E-01	.56082E-02	-.68348E-02	-.12966E-01	.85974E-01	-.42298E-02
45	-.37355E-02	.16855E-03	-.91226E-01	-.66980E-01	.16181E-02	-.93569E-01	.66688E-01	-.32339E-02
46	.40254E-01	-.68013E-02	.35554	-.39731E-01	.85119E-04	-.14449	.47001E-01	.18819E-02
47	.21661E-02	-.36559E-03	-.74211E-01	.27965E-02	-.43380E-03	-.14051E-01	-.24688E-01	

.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000
.00000	.00000	.00000	.00000	.00000	.00000	.00000

PARAMETER NOS.

OBS. NO.	9	10	11	12	13	14	15
1	-.18116E-01	.18461E-01	-.10261	-.76670E-01	.15602E-01	-.10705E-02	.26210E-02
2	.15398E-01	.63714E-02	-.10520E-01	.60021E-01	.21545E-01	-.88367E-02	-.23863E-02
3	.19270E-01	-.59081E-01	-.13502E-01	.58394E-01	-.31557E-02	-.19161E-02	-.43963E-02
4	-.13865E-01	.39926E-02	-.40248E-01	.62201E-01	-.17986E-01	-.80041E-02	.29120E-02
5	-.64447E-02	-.13709E-01	-.54199E-01	.15221	-.22743E-01	-.26276E-02	-.26176E-01
6	.63039E-01	-.82059E-02	.13557E-01	.57499E-01	-.49339E-02	-.38820E-02	-.35640E-01
7	-.28073E-01	-.34999E-01	.11925E-01	.27877	.19645E-01	.15303E-01	-.88377E-01
8	.28890E-01	-.31691E-01	.11163E-01	.25216	.19334E-01	.12377E-01	.13325
9	.42650E-02	.89289E-03	-.22234E-01	.23578E-01	.25553E-01	.96000E-02	-.40331E-01
10	.17661E-01	.79392E-02	.70300E-02	.60078E-01	.49967E-02	.29181E-02	.85570E-03
11	.17141E-01	-.40038E-02	-.20132E-01	.47596E-01	.11892E-01	-.62497E-02	-.64073E-02
12	.76332E-02	-.17338E-02	-.45887E-01	.47766E-01	-.23701E-01	-.10628E-01	.23060E-02
13	-.19936E-01	.49384E-01	-.28424E-01	-.19798E-01	-.33736E-01	-.19195E-01	-.18410E-03
14	.89682E-01	-.64821E-02	.17682E-02	-.51953E-01	-.13593E-01	-.21140E-03	.70513E-01
15	.14885	-.35952E-02	-.44827E-03	.29859E-01	-.36843E-02	-.31528E-01	.11362E-02
16	.00000	.00000	.00000	.00000	.00000	.84057E-01	.77757
17	-.23283E-01	.15133E-01	-.42296E-01	-.93402E-01	.39154E-02	-.50069E-02	.19005E-02
18	.93527E-02	-.94857E-02	.50630E-01	.40927E-01	.14658E-02	.49220E-02	-.33358E-02
19	.14382E-01	-.76122E-02	.32460E-01	.38813E-01	-.14822E-01	-.47574E-02	.42417E-02
20	-.13359E-01	-.11690E-02	.23750E-01	.26954E-01	-.11451E-01	-.73606E-02	-.76480E-02
21	-.74947E-04	.63984E-02	-.40223E-01	-.23171E-01	-.26848E-01	-.14388E-01	.96146E-02
22	-.33999E-01	.20836E-01	.25131E-01	-.18951	-.19809E-01	-.26975E-01	.16407E-01
23	.13087	.67467E-02	-.38542E-02	-.52603E-01	-.11616E-01	-.37463E-01	-.12748E-01
24	.20500	.18080E-02	-.12905E-02	-.13909E-01	-.60042E-02	-.59968E-02	.11174
25	-.11093E-01	-.64651E-02	.11763	-.32837E-01	-.33475E-01	-.37928E-02	-.35805E-02
26	.16055E-02	-.13031E-01	.14458	.14100E-02	-.48216E-01	-.10543E-01	.12397E-02
27	.90939E-02	-.33097E-01	.31903E-01	.38786E-02	-.15396E-01	.88702E-02	.77372E-02
28	.89084E-02	.31763E-02	-.27959E-01	.56644E-02	.30900E-02	-.11232E-02	.75317E-02
29	-.88583E-02	.15650E-01	-.20747E-01	-.11340	-.18460E-01	-.10425E-01	.13955E-01
30	-.28167E-01	.15862E-01	-.26993E-02	-.12832	-.14381E-01	-.93444E-02	-.16983E-01
31	.25336	-.41461E-02	-.20409E-02	-.32566E-01	-.77660E-02	-.41361E-01	-.18186E-01
32	.00000	.00000	.00000	.00000	.00000	.25217	.39943
33	.62420E-02	-.13050E-01	.14321	.25669E-02	.17396E-02	.13249E-02	-.41588E-03
34	-.34508E-02	.36447E-02	-.38942E-01	-.14874E-02	-.69499E-01	.13692E-01	-.53401E-02
35	.91670E-03	-.23169E-02	.29096E-01	-.22257E-02	.10733E-01	.86787E-02	.30763E-02
36	.90702E-02	-.22451E-02	.16153E-01	.66404E-02	.53810E-01	.12316E-01	.12429E-03
37	.34392E-02	.82785E-02	-.27664E-01	-.47790E-01	.30662E-01	.21147E-01	.18821E-02
38	-.16103E-01	.22265E-01	.25521E-01	-.20154	.17508E-01	.11819E-01	.65712E-02

Figure 4.3-10—Continued

39	.16117	.78650E-02	-.36457E-02	-.61941E-01	-.41596E-02	-.18909E-01	-.26062E-01
40	.30681	.62477E-04	.14940E-02	-.16047E-02	-.15188E-02	.32843E-01	-.63188E-02
41	-.20001E-02	.76411E-02	-.88409E-01	.18257E-02	.14546E-01	-.79979E-02	.33027E-02
42	-.28171E-02	.69553E-02	-.77255E-01	-.69040E-03	.48823E-01	-.47372E-03	.17000E-03
43	-.83165E-03	-.62057E-02	.89476E-01	-.14396E-01	.13127	.15071E-01	-.57999E-02
44	.64958E-02	.10056E-02	-.87212E-02	-.18886E-02	.13289	.55158E-01	-.14467E-01
45	-.12977E-02	.10628E-01	-.14025E-01	-.77051E-01	.80307E-01	.83613E-01	-.20443E-01
46	-.31529E-01	.12352E-01	-.35402E-02	-.98877E-01	.28424E-01	.76494E-01	-.11890E-01
47	.25299	.18326E-02	-.45556E-04	-.15020E-01	.12539E-01	.58409E-01	-.72207E-01
48	.00000	.00000	.00000	.00000	.00000	.42028	.21297E-01
49	.00000	.00000	.00000	.00000	.52187	-.22325	.72429E-01
50	.00000	.00000	.00000	.00000	.45229	-.12624	.40402E-01
51	.00000	.00000	.00000	.00000	.38271	-.29228E-01	.83755E-02
52	.00000	.00000	.00000	.00000	.31312	.67784E-01	-.23651E-01
53	.00000	.00000	.00000	.00000	.24354	.16480	-.55678E-01
54	.00000	.00000	.00000	.00000	.17396	.26181	-.87705E-01
55	.00000	.00000	.00000	.00000	.10437	.35882	-.11973
56	.00000	.00000	.00000	.00000	.34791E-01	.45583	-.15176
57	.79278	.12415E-01	.52711E-01	.41851E-16	.18357E-16	.61881E-16	

Figure 4.3-10—Continued

Program Listing

```
C       FINITE DIFFERENCE PROGRAM FOR NONLINEAR REGRESSION SOLUTION
C       OF TWO-DIMENSIONAL, STEADY-STATE, GROUND-WATER FLOW PROBLEMS
C       BY R. L. COOLEY, USGS, DENVER, COLO.
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        DIMENSION TITLE(20),DX(30),DY(30),CX(500),CY(500),VL(500)
       1,QR(500),WELL(500),HR(500),HC(500),HCI(70),BK(70),BL(70)
       2,BM(70),BN(70),HO(70),W(70),PAR(4),QBF(50),PLA(50),PLB(50)
       3,CXHR(50),CXHL(50),CYHT(50),CYHB(50),AU(5,250),AL(20,250)
       4,V(500),X(20,70),S(20,90),XV(500),XS(4),P(20),WP(20)
       5,A(20,20),B(50)
        DIMENSION JPOS(30),IZN(500),IBZN(50),IPRM(4,20),IBNA(50)
       1,IBNB(50),IBPA(10),IBPB(10),IBHN(50),IHSN(100),KOBS(70)
       2,LN(4),NCBA(20),NCEA(20),NCBF(20),NCEF(20),NCBH(20)
       3,NCEH(20),ILOC(500),JLOC(500),IN(500),IC(5,250)
        COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NOBS,NQSD,NBH
       1,NVAR,NVX2,KOUNT,INDT,IPO
        COMMON/LOC/ILOC,JLOC
        COMMON/TNME/IIN,IOUT
        COMMON/FLT/CX,CY,VL,QR,WELL,HR,HC,BK,BL,BM,BN,HO,W
        COMMON/SOLV/AU,AL
        COMMON/REG/DMX,ADMX,AP,CSA,AMP,RP,BP,YSQ
        EQUIVALENCE (TITLE(1),A(1,1),S(1,1),AU(1,1)),(HC(1),XV(1))
       1,(CX(1),HCI(1)),(ILOC(1),IC(1,1))
        OPEN (5,FILE='INVFD.DAT',STATUS='OLD',ACCESS='SEQUENTIAL'
       1,FORM='FORMATTED')
        OPEN (6,FILE='INVFD.OUT',STATUS='NEW',ACCESS='SEQUENTIAL'
       1,FORM='FORMATTED')
        OPEN (7,STATUS='NEW',ACCESS='SEQUENTIAL',FORM='UNFORMATTED')
C**DEFINE INPUT FILE, OUTPUT FILE, SCRATCH FILE, AND ARRAY DIMENSIONS
        IIN=5
        IOUT=6
        ITA=7
        NVD=20
        NAD=20
C**READ THREE TITLE LINES
        WRITE(IOUT,804)
        DO 5 I=1,3
        READ(IIN,801) (TITLE(J),J=1,20)                          SET A
      5 WRITE(IOUT,803) (TITLE(J),J=1,20)
C**READ JOB SPECIFICATION DATA
        READ(IIN,800) ID,JD,NZNS,NOBS,NPAR,NVAR,NWELS,NQBZ,NHBZ,NUM,IPRX
       1,IPO,ISO                                                 SET B
        WRITE(IOUT,802) ID,JD,NZNS,NOBS,NPAR,NVAR,NWELS,NQBZ,NHBZ,NUM
       1,IPRX,IPO,ISO
        READ(IIN,820) DMX,CSA,RP,BP,EV                           SET C
        WRITE(IOUT,806) DMX,CSA,RP,BP,EV
C**READ INITIAL ARRAY DATA
        IM=ID-1
        JM=JD-1
        CALL ARRAY(DX,IM,1,0)                                    SET D
```

Program Listing—Continued

```
      CALL ARRAY(DY,JM,1,0)                                      SET E
      CALL ARRAY(CX,IM,JM,0)                                     SET F
      CALL ARRAY(CY,IM,JM,0)                                     SET G
      CALL ARRAY(VL,IM,JM,0)                                     SET H
      CALL ARRAY(HR,ID,JD,0)                                     SET I
      CALL ARRAY(QR,IM,JM,0)                                     SET J
      CALL ARRAY(HC,ID,JD,0)                                     SET K
C**READ GRID ZONATION
      CALL ARRAYI(IZN,IM,JM,0)                                   SET L
C**READ OBSERVED HEAD DATA
      IF(NOBS.LT.1) GO TO 15
      CALL HOBS(DX,DY,V(1),V(ID+1),BK,BL,BM,BN,HO,W,KOBS)        SET M
C**INITIALIZE PARAMETER NUMBERS AND STANDARD DEVIATIONS TO ZERO
   15 DO 25 I=1,NZNS
      DO 20 K=1,4
   20 IPRM(K,I)=0
   25 CONTINUE
      DO 27 I=1,NVAR
   27 WP(I)=0.
C**READ AQUIFER PARAMETER NUMBERS
      IF(NPAR.LT.1) GO TO 35
      WRITE(IOUT,816)
      DO 30 J=1,NZNS
      READ(IIN,800) I,(IPRM(K,I),K=1,4)                          SET N
      WRITE(IOUT,818) I,(IPRM(K,I),K=1,4)
   30 CONTINUE
C**READ AQUIFER PARAMETER STANDARD DEVIATIONS
      WRITE(IOUT,822)
      DO 32 J=1,NPAR
      READ(IIN,812) K,WP(K)                                      SET O
   32 WRITE(IOUT,823) K,WP(K)
C**READ INITIAL AQUIFER PARAMETERS BY ZONE, AND LOAD THEM INTO THE
C    PARAMETER VECTOR
   35 WRITE(IOUT,810)
      M=NVAR
      DO 45 J=1,NZNS
      READ(IIN,812) I,PAR(1),PAR(2),PAR(3),PAR(4)                SET P
      WRITE(IOUT,814) I,PAR(1),PAR(2),PAR(3),PAR(4)
      DO 40 K=1,4
      L=IPRM(K,I)
      IF(L.GT.0) GO TO 40
      M=M+1
      L=M
      IPRM(K,I)=L
   40 B(L)=PAR(K)
   45 CONTINUE
C**DEFINE JPOS ARRAY SUCH THAT COLUMN+JPOS(ROW)=NODE NUMBER
      JPOS(1)=0
      DO 50 J=2,JD
   50 JPOS(J)=JPOS(J-1)+ID
C**READ POINT FLOW DATA
      NIJ=ID*JD
```

Program Listing—Continued

```
            DO 55 N=1,NIJ
    55 WELL(N)=0.
            IF(NWELS.LT.1) GO TO 61
            WRITE(IOUT,824)
            DO 60 K=1,NWELS
            READ(IIN,826) I,J,TMP                              SET Q
            WRITE(IOUT,828) I,J,TMP
            L=I+JPOS(J)
    60 WELL(L)=TMP
C**READ AND FORM ARRAYS FOR SPECIFIED POINT OR LINE FLOWS
    61 NQSD=0
            IF(NQBZ.LT.1) GO TO 85
            WRITE(IOUT,830)
            N=0
            DO 80 J=1,NQBZ
            READ(IIN,832) IA,JA,IB,JB,IP,QB,SDQB,QBM            SET R
            WRITE(IOUT,831) IA,JA,IB,JB,IP,QB,SDQB,QBM
            M=1
            K=IA-1
            IF(JA.EQ.JB) GO TO 62
            M=ID
            K=JA-1
    62 MA=IA+JPOS(JA)
            MB=IB+JPOS(JB)-M
            IF(MB.GE.MA) GO TO 64
            IF(IP.LT.1) GO TO 63
            N=N+1
            IBNA(N)=MA
            IBNB(N)=MA
            QBF(N)=.5*QBM
            IBZN(N)=IP
            GO TO 68
    63 WELL(MA)=QB*QBM
            GO TO 80
    64 QBM=.5*QBM
            IF(IP.LT.1) GO TO 70
            DO 66 L=MA,MB,M
            N=N+1
            IBNA(N)=L
            IBNB(N)=L+M
            K=K+1
            TEMP=DX(K)
            IF(M.EQ.ID) TEMP=DY(K)
            QBF(N)=QBM*TEMP
    66 IBZN(N)=IP
    68 B(IP)=QB
            WP(IP)=SDQB
            GO TO 80
    70 TMP=QB*QBM
            DO 75 L=MA,MB,M
            K=K+1
            TEMP=DX(K)
```

Program Listing—Continued

```
        IF(M.EQ.ID) TEMP=DY(K)
        TEMP=TMP*TEMP
        WELL(L)=WELL(L)+TEMP
   75 WELL(L+M)=WELL(L+M)+TEMP
   80 CONTINUE
        NQSD=N
C**READ SPECIFIED BOUNDARY HEAD POSITIONS AS -1'S
   85 CALL ARRAYI(IN,ID,JD,0)                                    SET S
C**READ DATA AND FORM ARRAYS FOR SPECIFIED HEADS AND PARAMETERS
        IF(NHBZ.LT.1) GO TO 110
        WRITE(IOUT,833)
        NBH=0
        DO 108 KK=1,NHBZ
        READ(IIN,834) IZ,NN,M,N,SDHA,SDHB                        SET T
        WRITE(IOUT,836) IZ,NN,M,N,SDHA,SDHB
        DO 95 J=1,NN
        READ(IIN,826) ILOC(J),JLOC(J),V(J)                       SET T
   95 WRITE(IOUT,840) ILOC(J),JLOC(J),V(J)
        IBPA(IZ)=M
        IBPB(IZ)=N
        IF(M.LT.1) GO TO 97
        B(M)=V(1)
        WP(M)=SDHA
   97 IF(N.LT.1) GO TO 98
        B(N)=V(NN)
        WP(N)=SDHB
   98 J=JLOC(1)
        K=ILOC(1)+JPOS(J)
        M=M+N
        NBHS=NBH
        IF(IN(K).LT.-1) GO TO 100
        NBH=NBH+1
        IF(M.GT.0) IN(K)=-NBH-1
        IBZN(NBH+NQSD)=IZ
        IBHN(NBH)=K
        PLA(NBH)=1.
        PLB(NBH)=0.
  100 IF(NN.LT.2) GO TO 107
        DIST=0.
        DO 102 KNT=2,NN
        J=JLOC(KNT)
        L=ILOC(KNT)+JPOS(J)
        NBH=NBH+1
        IF(M.GT.0) IN(L)=-NBH-1
        IBZN(NBH+NQSD)=IZ
        IBHN(NBH)=L
        JM1=JLOC(KNT-1)
        IF(J.EQ.JM1) GO TO 101
        J=MIN0(J,JM1)
        DIST=DIST+DY(J)
        GO TO 102
  101 I=MIN0(ILOC(KNT),ILOC(KNT-1))
```

Program Listing—Continued

```
        DIST=DIST+DX(I)
    102 PLB(NBH)=DIST
        N=NBH-NN+1
        DO 106 KNT=2,NN
        J=JLOC(KNT)
        L=ILOC(KNT)+JPOS(J)
        N=N+1
        TMPA=PLB(N)/DIST
        TMPB=1.-TMPA
        TMPC=TMPA*V(NN)+TMPB*V(1)
        IF(DABS(V(KNT)).LE.0.) GO TO 104
        TMP=V(KNT)/TMPC
        TMPA=TMPA*TMP
        TMPB=TMPB*TMP
        TMPC=V(KNT)
    104 PLA(N)=TMPB
        PLB(N)=TMPA
    106 HC(L)=TMPC
    107 IF(M.LT.1) NBH=NBHS
    108 HC(K)=V(1)
C**COMPARE CX AND CY WITH IZN FOR CONFLICT
    110 IER=0
        N=0
        DO 115 J=1,JM
        DO 115 I=1,IM
        N=N+1
        IF(IZN(N).LT.1) GO TO 115
        IF(CX(N).GT.0..OR.CY(N).GT.0.) GO TO 115
        IER=1
        WRITE(IOUT,842) I,J
    115 CONTINUE
        IF(IER.LT.1) GO TO 120
        WRITE(IOUT,844)
        STOP
C**TRANSFER DOMAIN GEOMETRY TO IN(M) AND COMPUTE CELL FLOW-COEFFICIENTS
    120 N=0
        DO 122 J=1,JM
        DYN=.5*DY(J)
        DO 122 I=1,IM
        N=N+1
        IF(IZN(N).LT.1) GO TO 122
        M=N+J
        IF(IN(M).GT.-1) IN(M)=1
        IF(IN(M-1).GT.-1) IN(M-1)=1
        IF(IN(M+ID-1).GT.-1) IN(M+ID-1)=1
        IF(IN(M+ID).GT.-1) IN(M+ID)=1
        CX(N)=CX(N)*DYN/DX(I)
        DXN=.5*DX(I)
        CY(N)=CY(N)*DXN/DY(J)
        AREA=DXN*DYN
        VL(N)=VL(N)*AREA
        QR(N)=QR(N)*AREA
```

Program Listing—Continued

```
    122 CONTINUE
C**SET UP D4 ORDERING
        CALL ORDER(JPOS,IN,IC)
C**COMPUTE INITIAL SOLUTION
        CALL COEF(WELL,HR,HC,CX,CY,VL,QR,CXHR,CXHL,CYHT,CYHB,QBF,B,AU,AL
      1,V,IZN,IBZN,IPRM,IBNA,IBNB,IN,IC,NAD)
        CALL D4SOLV(HC,AU,AL,V,IN,IC,NAD)
        WRITE(IOUT,846)
        CALL ARRAY(HC,ID,JD,1)
        IF(ISO.EQ.1) GO TO 640
C**COMPUTE AND COUNT PRIOR INFORMATION DATA
        NPRIR=0
        DO 137 I=1,NVAR
        P(I)=B(I)
        IF(WP(I).LE.0.) GO TO 137
        WP(I)=EV/(WP(I)*WP(I))
        NPRIR=NPRIR+1
    137 CONTINUE
        WRITE(IOUT,848) NPRIR
C**INITIALIZE BEGINNING AND END POINT ARRAYS
        DO 148 I=1,NVAR
        NCBA(I)=0
        NCEA(I)=0
        NCBF(I)=0
        NCEF(I)=0
        NCBH(I)=0
    148 NCEH(I)=0
C**DEFINE BEGINNING AND END POINT ARRAYS FOR AQUIFER PARAMETERS
        IF(NPAR.LT.1) GO TO 154
        N=0
        DO 152 J=1,JM
        DO 152 I=1,IM
        N=N+1
        L=IZN(N)
        IF(L.LT.1) GO TO 152
        DO 150 M=1,4
        K=IPRM(M,L)
        IF(K.GT.NVAR) GO TO 150
        NCEA(K)=N
        IF(NCBA(K).LT.1) NCBA(K)=N
    150 CONTINUE
    152 CONTINUE
C**ORDER IBZN AND CORRESPONDING ARRAYS FOR LINE FLOW PARAMETERS
C   FROM SMALLEST TO LARGEST
    154 IF(NQSD.LT.1) GO TO 162
        DO 158 I=1,NQSD
        DO 156 J=I,NQSD
        IF(IBZN(J).GE.IBZN(I)) GO TO 156
        ITMP=IBZN(I)
        IBZN(I)=IBZN(J)
        IBZN(J)=ITMP
        ITMP=IBNA(I)
```

*Program Listing—*Continued

```
          IBNA(I)=IBNA(J)
          IBNA(J)=ITMP
          ITMP=IBNB(I)
          IBNB(I)=IBNB(J)
          IBNB(J)=ITMP
          TMP=QBF(I)
          QBF(I)=QBF(J)
          QBF(J)=TMP
      156 CONTINUE
      158 CONTINUE
C**DEFINE BEGINNING AND END POINT ARRAYS FOR LINE FLOW PARAMETERS
          DO 160 I=1,NQSD
          K=IBZN(I)
          NCEF(K)=I
          IF(NCBF(K).LT.1) NCBF(K)=I
      160 CONTINUE
C**DEFINE SEQUENCE NUMBERS, AND BEGINNING AND END POINT ARRAYS
C  FOR SPECIFIED HEAD PARAMETERS
      162 IF(NBH.LT.1) GO TO 174
          DO 164 I=1,NBH
          J=IBZN(I+NQSD)
          K=IBPA(J)
          IF(K.EQ.0) GO TO 163
          NCEH(K)=I
          IF(NCBH(K).LT.1) NCBH(K)=I
      163 K=IBPB(J)
          IF(K.EQ.0) GO TO 164
          NCEH(K)=I
          IF(NCBH(K).LT.1) NCBH(K)=I
      164 CONTINUE
          L=0
          DO 172 K=1,NVAR
          IF(NCBH(K).LT.1) GO TO 172
          JB=NCBH(K)
          JE=NCEH(K)
          NCBH(K)=L+1
          DO 170 J=JB,JE
          I=IBZN(J+NQSD)
          IF(IBPA(I).EQ.K.OR.IBPB(I).EQ.K) GO TO 168
          GO TO 170                                    .
      168 L=L+1
          IHSN(L)=J
      170 CONTINUE
          NCEH(K)=L
      172 CONTINUE
C**BEGIN ITERATIONS
      174 INDT=0
          ER=.01
          ERP=.001
          AMP=0.
          NVX2=NVAR+NVAR
          KOUNT=0
```

Program Listing—Continued

```
  176 KOUNT=KOUNT+1
      REWIND ITA
C**SOLVE FOR SENSITIVITIES:
      DO 260 N=1,NVAR
      DO 178 I=1,NEQ
  178 V(I)=0.
      IF(NCBA(N).LT.1) GO TO 208
C**ASSEMBLE R.H.S. FOR AQUIFER PARAMETERS
      LB=NCBA(N)
      LE=NCEA(N)
      DO 200 L=LB,LE
      J=IZN(L)
      IF(J.LT.1) GO TO 200
      NA=L+(L-1)/IM
      NB=NA+1
      NC=NB+ID
      ND=NA+ID
      INA=IN(NA)
      INB=IN(NB)
      INC=IN(NC)
      IND=IN(ND)
      IF(IPRM(1,J).NE.N) GO TO 180
      IF(INA.GT.0) V(INA)=V(INA)+CX(L)*(HC(NB)-HC(NA))
      IF(INB.GT.0) V(INB)=V(INB)+CX(L)*(HC(NA)-HC(NB))
      IF(INC.GT.0) V(INC)=V(INC)+CX(L)*(HC(ND)-HC(NC))
      IF(IND.GT.0) V(IND)=V(IND)+CX(L)*(HC(NC)-HC(ND))
  180 IF(IPRM(2,J).NE.N) GO TO 185
      IF(INA.GT.0) V(INA)=V(INA)+CY(L)*(HC(ND)-HC(NA))
      IF(INB.GT.0) V(INB)=V(INB)+CY(L)*(HC(NC)-HC(NB))
      IF(INC.GT.0) V(INC)=V(INC)+CY(L)*(HC(NB)-HC(NC))
      IF(IND.GT.0) V(IND)=V(IND)+CY(L)*(HC(NA)-HC(ND))
  185 IF(IPRM(3,J).NE.N) GO TO 190
      IF(INA.GT.0) V(INA)=V(INA)+VL(L)*(HR(NA)-HC(NA))
      IF(INB.GT.0) V(INB)=V(INB)+VL(L)*(HR(NB)-HC(NB))
      IF(INC.GT.0) V(INC)=V(INC)+VL(L)*(HR(NC)-HC(NC))
      IF(IND.GT.0) V(IND)=V(IND)+VL(L)*(HR(ND)-HC(ND))
  190 IF(IPRM(4,J).NE.N) GO TO 200
      IF(INA.GT.0) V(INA)=V(INA)+QR(L)
      IF(INB.GT.0) V(INB)=V(INB)+QR(L)
      IF(INC.GT.0) V(INC)=V(INC)+QR(L)
      IF(IND.GT.0) V(IND)=V(IND)+QR(L)
  200 CONTINUE
C**ASSEMBLE R.H.S. FOR SPECIFIED LINE FLOW PARAMETERS
  208 IF(NCBF(N).LT.1) GO TO 212
      LB=NCBF(N)
      LE=NCEF(N)
      DO 210 L=LB,LE
      I=IBNA(L)
      J=IN(I)
      IF(J.GT.0) V(J)=V(J)+QBF(L)
      I=IBNB(L)
      J=IN(I)
```

Program Listing—Continued

```
          IF(J.GT.0) V(J)=V(J)+QBF(L)
      210 CONTINUE
C**ASSEMBLE R.H.S. FOR SPECIFIED HEAD PARAMETERS
      212 IF(NCBH(N).LT.1) GO TO 216
          LB=NCBH(N)
          LE=NCEH(N)
          DO 214 L=LB,LE
          K=IHSN(L)
          I=IBZN(K+NQSD)
          TMP=0.
          IF(IBPA(I).EQ.N) TMP=PLA(K)
          IF(IBPB(I).EQ.N) TMP=TMP+PLB(K)
          I=IBHN(K)+1
          IF(I.LE.NIJ) J=IN(I)
          IF(J.GT.0) V(J)=V(J)+CXHR(K)*TMP
          I=IBHN(K)-1
          IF(I.GT.0) J=IN(I)
          IF(J.GT.0) V(J)=V(J)+CXHL(K)*TMP
          I=IBHN(K)+ID
          IF(I.LE.NIJ) J=IN(I)
          IF(J.GT.0) V(J)=V(J)+CYHT(K)*TMP
          I=IBHN(K)-ID
          IF(I.GT.0) J=IN(I)
          IF(J.GT.0) V(J)=V(J)+CYHB(K)*TMP
      214 CONTINUE
C**MODIFY R.H.S.--UPPER HALF
      216 DO 220 J=1,ICR1
          II=IC(1,J)
          DO 218 I=2,II
          LR=IC(I,J)
          V(LR)=V(LR)-AU(I,J)*V(J)
      218 CONTINUE
      220 V(J)=V(J)/AU(1,J)
C**MODIFY R.H.S.--LOWER HALF
          JJ=NEQ-ICR
          DO 224 J=1,JJ
          JR=J+ICR1
          LR=JR
          DO 222 I=2,IB1
          LR=LR+1
          IF(AL(I,J).NE.0.) V(LR)=V(LR)-AL(I,J)*V(JR)
      222 CONTINUE
      224 V(JR)=V(JR)/AL(1,J)
C**BACK SOLVE--LOWER HALF
          V(NEQ)=V(NEQ)/AL(1,NEQ-ICR1)
          DO 230 J=1,JJ
          KK=NEQ-J
          KL=KK-ICR1
          L=KK
          DO 226 I=2,IB1
          L=L+1
          IF(AL(I,KL).NE.0.) V(KK)=V(KK)-AL(I,KL)*V(L)
```

Program Listing—Continued

```
  226 CONTINUE
  230 CONTINUE
C**BACK SOLVE--UPPER HALF
      DO 250 J=1,ICR1
      KK=ICR-J
      II=IC(1,KK)
      DO 240 I=2,II
      L=IC(I,KK)
      V(KK)=V(KK)-AU(I,KK)*V(L)
  240 CONTINUE              .
  250 CONTINUE
      WRITE(ITA) (V(I),I=1,NEQ)
C**COMPUTE SENSITIVITIES AT OBSERVATION POINTS
      DO 255 I=1,NOBS
      K=KOBS(I)
      LN(1)=IN(K)
      LN(2)=IN(K+1)
      LN(3)=IN(K+ID)
      LN(4)=IN(K+ID+1)
      DO 253 J=1,4
      L=LN(J)
      IF(L.GT.0) GO TO 252
      XS(J)=0.
      IF(L.GT.-2) GO TO 253
      L=-L-1
      IZ=IBZN(L+NQSD)
      TMP=0.
      IF(IBPA(IZ).EQ.N) TMP=PLA(L)
      IF(IBPB(IZ).EQ.N) TMP=TMP+PLB(L)
      XS(J)=TMP
      GO TO 253
  252 XS(J)=V(L)
  253 CONTINUE
  255 X(N,I)=BK(I)*XS(1)+BL(I)*XS(2)+BM(I)*XS(3)+BN(I)*XS(4)
  260 CONTINUE
      IF(IPO.NE.1) GO TO 270
      WRITE(IOUT,850)
      DO 265 N=1,NOBS
      WRITE(IOUT,852) N,(X(K,N),K=1,NVAR)
  265 CONTINUE
C**CALL LEAST SQUARES
  270 CALL LSTSQ(HC,BK,BL,BM,BN,HO,W,P,WP,X,A,B,V,KOBS,IN,NVD)
      IF(INDT.GT.0) GO TO 515
C**COMPUTE NEW SPECIFIED HEADS
      IF(NBH.LT.1) GO TO 310
      DO 300 N=1,NBH
      M=IBZN(N+NQSD)
      K=IBPA(M)
      TMPA=0.
      IF(K.GT.0) TMPA=PLA(N)*V(K)
      L=IBPB(M)
      TMPB=0.
```

Program Listing—Continued

```
        IF(L.GT.0) TMPB=PLB(N)*V(L)
        J=IBHN(N)
        HC(J)=HC(J)+TMPA+TMPB
  300 CONTINUE
C**CHECK FOR CONVERGENCE
  310 IF(ADMX.LT.ER) GO TO 350
C**CHECK FOR PARAMETERS GOING TO ZERO
        IND=0
        DO 335 I=1,NVAR
        IF(DABS(B(I)).GT.DABS(ERP*P(I))) GO TO 335
        WRITE(IOUT,858) I
        IND=1
  335 CONTINUE
        IF(IND.GT.0) GO TO 515
        IF(KOUNT.EQ.NUM) GO TO 340
C**COMPUTE NEW HEADS AT GRID POINTS
        CALL COEF(WELL,HR,HC,CX,CY,VL,QR,CXHR,CXHL,CYHT,CYHB,QBF,B,AU,AL
      1,V,IZN,IBZN,IPRM,IBNA,IBNB,IN,IC,NAD)
        CALL D4SOLV(HC,AU,AL,V,IN,IC,NAD)
        GO TO 176
  340 WRITE(IOUT,860) NUM
        GO TO 515
  350 WRITE(IOUT,862) KOUNT
C**COMPUTE FINAL ESTIMATES OF HEAD
        REWIND ITA
        DO 366 K=1,NVAR
        READ(ITA) (CX(I),I=1,NEQ)
        DO 364 J=1,NIJ
        L=IN(J)
        IF(L.GT.0) HC(J)=HC(J)+CX(L)*V(K)
  364 CONTINUE
  366 CONTINUE
C**CORRECT A FOR MARQUARDT PARAMETER
        IF(NVAR.EQ.1) GO TO 420
        IF(AMP.LE.0.) GO TO 385
        DO 380 I=1,NVAR
        A(I,I)=1.+RP
        DO 375 J=I,NVAR
  375 A(J,I)=A(I,J)
  380 CONTINUE
        AMP=-1.
        CALL LSTSQ(HC,BK,BL,BM,BN,HO,W,P,WP,X,A,B,V,KOBS,IN,NVD)
        IF(INDT.GT.0) GO TO 515
C**COMPUTE A-INVERSE
  385 A(NVAR,NVAR)=1./A(NVAR,NVAR)
        NM1=NVAR-1
        DO 410 K=1,NM1
        KP1=K+1
        DO 395 I=KP1,NVAR
        SUM=0.
        IM1=I-1
        DO 390 J=K,IM1
```

Program Listing—Continued

```
  390 SUM=SUM+A(I,J)*A(J,K)
      A(K,I)=-SUM
  395 A(I,K)=-SUM*A(I,I)
      DO 405 I=1,K
      SUM=A(K,I)
      DO 400 J=KP1,NVAR
  400 SUM=SUM+A(J,K)*A(I,J)
      A(K,I)=SUM
  405 A(I,K)=A(K,I)
  410 CONTINUE
      DO 415 J=1,NVAR
  415 A(J,NVAR)=A(NVAR,J)
      GO TO 425
  420 A(1,1)=1./(1.+RP)
C**COMPUTE TR((A-INVERSE)**2) AND A-INVERSE - RP*(A-INVERSE)**2
  425 TRACE=0.
      IF(RP.LE.0.) GO TO 448
      DO 445 N=1,NVAR
      DO 430 J=1,NVAR
  430 V(J)=A(J,N)
      SUMA=0.
      DO 440 J=N,NVAR
      SUM=0.
      DO 435 I=1,NVAR
  435 SUM=SUM+V(I)*A(I,J)
      V(J+NVAR)=SUM
  440 A(J,N)=A(J,N)-RP*SUM
  445 TRACE=TRACE+V(N+NVAR)
C**COMPUTE SUM OF SQUARED ERRORS
  448 YSQ=0.
      DO 450 N=1,NOBS
      K=KOBS(N)
      HCI(N)=BK(N)*HC(K)+BL(N)*HC(K+1)+BM(N)*HC(K+ID)+BN(N)*HC(K+ID+1)
  450 YSQ=YSQ+(HO(N)-HCI(N))*W(N)*(HO(N)-HCI(N))
      DO 455 I=1,NVAR
  455 YSQ=YSQ+(P(I)-B(I))*WP(I)*(P(I)-B(I))
C**COMPUTE ERROR VARIANCE
      TEMP=NPRIR-NVAR
      OBS=NOBS
      VAR=YSQ/(OBS+TEMP+RP*RP*TRACE)
C**COMPUTE CORRELATION COEFFICIENT
      SUMA=0.
      SUMB=0.
      SUMC=0.
      SUMD=0.
      SUM=0.
      DO 460 N=1,NOBS
      TMP=W(N)**.5
      W(N)=TMP
      TEMP=TMP*HO(N)
      TMP=TMP*HCI(N)
      SUMA=SUMA+TEMP
```

Program Listing—Continued

```
        SUMB=SUMB+TMP
        SUMC=SUMC+TEMP*TEMP
        SUMD=SUMD+TMP*TMP
        SUM=SUM+TEMP*TMP
    460 CONTINUE
        R=(OBS*SUM-SUMA*SUMB)/((OBS*SUMC-SUMA*SUMA)*(OBS*SUMD-SUMB*SUMB))
       1**.5
C**PRINT ERROR VARIANCE, ESTIMATED SUM OF SQUARED ERRORS, AND
C   CORRELATION COEFFICIENT
        WRITE(IOUT,864) VAR,YSQ,R
C**COMPUTE VARIANCE-COVARIANCE MATRIX
        DO 463 J=1,NVAR
        TEMP=V(J+NVX2)
        DO 462 I=J,NVAR
        A(I,J)=VAR*A(I,J)/(V(I+NVX2)*TEMP)
    462 A(J,I)=A(I,J)
    463 V(J)=A(J,J)**.5
C**PRINT PARAMETERS AND STANDARD ERRORS
        WRITE(IOUT,870)
        DO 480 J=1,NVAR
    480 WRITE(IOUT,856) J,B(J),V(J)
C**PRINT VARIANCE-COVARIANCE MATRIX
    490 WRITE(IOUT,874)
        CALL PRTOT(A,NVAR,NVD,0)
C**COMPUTE AND PRINT CORRELATION MATRIX
        DO 510 J=1,NVAR
        TEMP=V(J)
        DO 500 I=J,NVAR
        A(I,J)=A(I,J)/(V(I)*TEMP)
    500 A(J,I)=A(I,J)
    510 CONTINUE
        WRITE(IOUT,876)
        CALL PRTOT(A,NVAR,NVD,0)
C**PRINT COMPUTED AND OBSERVED HEADS, AND COMPUTE AND PRINT RESIDUALS
        GO TO 518
    515 DO 516 N=1,NOBS
        K=KOBS(N)
        HCI(N)=BK(N)*HC(K)+BL(N)*HC(K+1)+BM(N)*HC(K+ID)+BN(N)*HC(K+ID+1)
    516 W(N)=W(N)**.5
    518 WRITE(IOUT,878)
        DO 520 N=1,NOBS
        RES=W(N)*(HCI(N)-HO(N))
        WRITE(IOUT,880) N,HCI(N),HO(N),RES
    520 CONTINUE
C**PRINT HYDRAULIC HEADS FOR EACH NODE
        WRITE(IOUT,881)
        CALL ARRAY(HC,ID,JD,1)
C**PRINT SENSITIVITIES FOR EACH NODE
        IF(IPRX.LT.1.AND.KOUNT.LT.NUM) STOP
        WRITE(IOUT,882)
        REWIND ITA
        DO 530 KK=1,NVAR
```

*Program Listing—*Continued

```
      READ(ITA) (CX(I),I=1,NEQ)
      WRITE(IOUT,884) KK
      DO 525 N=1,NIJ
      L=IN(N)
      IF(L.GT.0) GO TO 523
      XV(N)=0.
      IF(L.GT.-2) GO TO 525
      L=-L-1
      IZ=IBZN(L+NQSD)
      TMP=0.
      IF(IBPA(IZ).EQ.KK) TMP=PLA(L)
      IF(IBPB(IZ).EQ.KK) TMP=TMP+PLB(L)
      XV(N)=TMP
      GO TO 525
  523 XV(N)=CX(L)
  525 CONTINUE
  530 CALL ARRAY(XV,ID,JD,1)
      IF(NVAR.LT.2) STOP
C**SCALE AND ORTHOGONALIZE COLUMNS OF SENSITIVITY MATRIX, X:
C**SCALE X AND AUGMENT X TO INCLUDE PRIOR
      DO 535 N=1,NOBS
      DO 535 K=1,NVAR
  535 S(K,N)=X(K,N)*W(N)/V(K+NVX2)
      IF(NPRIR.LT.1) GO TO 539
      N=NOBS
      DO 538 I=1,NVAR
      IF(WP(I).LT.1.E-10) GO TO 538
      N=N+1
      DO 537 J=1,NVAR
  537 S(J,N)=0.
      S(I,N)=WP(I)**.5/V(I+NVX2)
  538 CONTINUE
C**ORTHOGONALIZE S
  539 NTMP=NOBS+NPRIR
      DO 540 I=1,NTMP
  540 CY(I)=S(1,I)
      DO 600 N=2,NVAR
      NM1=N-1
      SUM=0.
      DO 550 I=1,NTMP
      SUM=SUM+CY(I)*CY(I)
      S(NM1,I)=CY(I)
  550 CONTINUE
      IF(SUM.LT.1.E-20) GO TO 610
      V(NM1)=1./SUM
      DO 570 J=1,NM1
      SUM=0.
      DO 560 K=1,NTMP
  560 SUM=SUM+V(J)*S(J,K)*S(N,K)
  570 CX(J)=SUM
      DO 590 K=1,NTMP
      SUM=0.
```

Program Listing—Continued

```
        DO 580 I=1,NM1
  580 SUM=SUM+S(I,K)*CX(I)
  590 CY(K)=S(N,K)-SUM
  600 CONTINUE
C**PRINT ORTHOGONALIZED S
  610 WRITE(IOUT,886)
        K=1
        L=8
        DO 630 M=1,NVAR,8
        IF(L.GT.NVAR) L=NVAR
        WRITE(IOUT,888) (I,I=K,L)
        DO 620 J=1,NTMP
        S(NVAR,J)=CY(J)
        WRITE(IOUT,890) J,(S(I,J),I=K,L)
  620 CONTINUE
        WRITE(IOUT,890)
        K=K+8
        L=L+8
  630 CONTINUE
        STOP
C**READ, PRINT, AND EXECUTE FOR ALTERNATE SOLUTIONS
  640 READ(IIN,800) N                                          SET U
        IF(N.LT.1) STOP
        DO 690 KNT=1,N
        WRITE(IOUT,891) KNT
        DO 670 L=1,NVAR
        READ(IIN,812) I,PR                                     SET V
        V(I)=PR-B(I)
  670 B(I)=PR
        WRITE(IOUT,892)
        CALL PRTOT(B,NVAR,0,1)
        IF(NBH.LT.1) GO TO 685
        DO 680 N=1,NBH
        M=IBZN(N+NQSD)
        K=IBPA(M)
        TMPA=0.
        IF(K.GT.0) TMPA=PLA(N)*V(K)
        L=IBPB(M)
        TMPB=0.
        IF(L.GT.0) TMPB=PLB(N)*V(L)
        J=IBHN(N)
  680 HC(J)=HC(J)+TMPA+TMPB
  685 CALL COEF(WELL,HR,HC,CX,CY,VL,QR,CXHR,CXHL,CYHT,CYHB,QBF,B,AU,AL
      1,V,IZN,IBZN,IPRM,IBNA,IBNB,IN,IC,NAD)
        CALL D4SOLV(HC,AU,AL,V,IN,IC,NAD)
        WRITE(IOUT,898)
        CALL ARRAY(HC,ID,JD,1)
  690 CONTINUE
        STOP
C
  800 FORMAT (16I5)
  801 FORMAT (20A4)
```

Program Listing—Continued

```
802 FORMAT (56HONUMBER OF COLUMNS (ID) -------------------------------
   $=,I5
      $/56H NUMBER OF ROWS (JD) -------------------------------- =,I5
      $/56H NUMBER OF AQUIFER ZONES (NZNS) ---------------------- =,I5
      $/56H NUMBER OF OBSERVATIONS (NOBS) ----------------------- =,I5
      $/56H NUMBER OF AQUIFER PARAMETERS (NPAR) ---------------- =,I5
      $/56H TOTAL NUMBER OF PARAMETERS (NVAR) ------------------ =,I5
      $/56H NUMBER OF KNOWN POINT FLOWS (NWELS) ---------------- =,I5
      $/56H NUMBER OF SPECIFIED FLOW ZONES (NQBZ) -------------- =,I5
      $/56H NUMBER OF SPECIFIED HEAD ZONES (NHBZ) -------------- =,I5
      $/56H MAXIMUM NUMBER OF ITERATIONS (NUM) ----------------- =,I5
      $/56H SENSITIVITY PRINT AND ORTHOGONALIZATION OPTION (IPRX) =,I5
      $/56H ADDITIONAL PRINTOUT OPTION (IPO) ------------------- =,I5
      $/56H SOLUTION ONLY OPTION (ISO) ------------------------- =,I5)
803 FORMAT (1H ,20A4)
804 FORMAT (1H1)
806 FORMAT (50H MAXIMUM ALLOWABLE PARAMETER CORRECTION (DMX) - =
   $,G11.5
      $/50H SEARCH DIRECTION ADJUSTMENT PARAMETER (CSA) -- = ,G11.5
      $/50H RIDGE PARAMETER FOR REGRESSION (RP) ---------- = ,G11.5
      $/50H BIAS PARAMETER FOR REGRESSION (BP) ----------- = ,G11.5
      $/50H ESTIMATED ERROR VARIANCE (EV) --------------- = ,G11.5)
810 FORMAT (1H0,12X,34HINITIAL AQUIFER PARAMETERS BY ZONE/6H  ZONE
   1,5X,5HTRANX,8X,5HTRANY,8X,5HVLEAK,8X,5HQDIST)
812 FORMAT (I5,4F10.0)
814 FORMAT (1H ,I4,2X,4(2X,G11.5))
816 FORMAT (1H0,11X,25HAQUIFER PARAMETER NUMBERS/1H ,5X,4HZONE,4X
   1,5HTRANX,3X,5HTRANY,3X,5HVLEAK,3X,5HQDIST)
818 FORMAT (1H ,8I8)
820 FORMAT (8F10.0)
822 FORMAT (1H0,4X,19HSTANDARD DEVIATIONS/1H ,3X,22HFOR AQUIFER PARAME
   1TERS/1H ,6X,4HPAR.,6X,4HSTD./1H ,7X,3HNO.,6X,4HDEV.)
823 FORMAT (1H ,5X,I4,4X,G11.5)
824 FORMAT (1H0,12X,11HPOINT FLOWS/1H ,7X,1HI,7X,1HJ,4X,9HVOL. RATE)
826 FORMAT (2I5,F10.0)
828 FORMAT (1H ,2I8,4X,G11.5)
830 FORMAT (1H0,22X,27HINITIAL SPECIFIED FLOW DATA/1H ,6X,9HNODE NO.S
   1,7X,4HPAR.,6X,4HFLOW,10X,4HSTD./1H ,19H IA   JA   IB   JB,4X
   2,3HNO.,4X,9HPARAMETER,7X,4HDEV.,5X,10HMULTIPLIER)
831 FORMAT (1H ,4(1X,I3,1X),2X,I3,3X,3(2X,G11.5))
832 FORMAT (5I5,3F10.0)
833 FORMAT (1H0,18X,27HINITIAL SPECIFIED HEAD DATA)
834 FORMAT (4I5,2F10.0)
836 FORMAT (1H0,22H  NO. OF NODES IN ZONE,I4,3H = ,I3/1H
   1,20H  PAR. NO. HEAD A - ,I3,12X,18HPAR. NO. HEAD B = ,I3/1H
   2,21H  STD. DEV. HEAD A - ,G11.5,22H   STD. DEV. HEAD B = ,G11.5
   3/1H ,20X,22HINITIAL VALUES OF HEAD/1H ,21X,1HI,5X,1HJ,8X,4HHEAD)
840 FORMAT (1H ,19X,2(I3,3X),2X,G11.5)
842 FORMAT (9HOAT CELL ,1H(,I3,1H,,I3,1H),23H, IZN>0, CX=0, AND CY=0)
844 FORMAT (60HOPROGRAM ABORTED BECAUSE OF CONFLICT BETWEEN IZN, CX, A
   1ND CY)
846 FORMAT (18H0 INITIAL SOLUTION)
```

Program Listing—Continued

```
848 FORMAT (46H0NO. OF PARAMETERS HAVING PRIOR INFORMATION = ,I4)
850 FORMAT (1H0,3X,55HOBSERVATION NUMBER AND SENSITIVITIES FOR EACH PA
   1RAMETER)
852 FORMAT (1H ,1X,I4,7(1X,G11.5)/(1H ,(5X,7(1X,G11.5))))
856 FORMAT (1H ,1X,I4,3X,4(G11.5,4X))
858 FORMAT (11H0PARAMETER ,I3,17H EFFECTIVELY ZERO)
860 FORMAT (//32H0SOLUTION FAILED TO CONVERGE IN ,I3,11H ITERATIONS)
862 FORMAT (//23H0SOLUTION CONVERGED IN ,I3,11H ITERATIONS)
864 FORMAT (18H0ERROR VARIANCE = ,G11.5/35H ESTIMATED SUM OF SQUARED E
   1RRORS = ,G11.5/27H·CORRELATION COEFFICIENT = ,G11.5)
870 FORMAT (1H0,5X,24HESTIMATED PARAMETER DATA/1H ,6H  PAR.
   1/1H ,6H   NO.,5X,4HPAR.,9X,9HSTD. DEV.)
874 FORMAT (28H0 VARIANCE-COVARIANCE MATRIX)
876 FORMAT (20H0 CORRELATION MATRIX)
878 FORMAT (1H0,21X,14HHEAD RESIDUALS/1H ,7H   OBS.,5X,9HPREDICTED
   1,7X,8HOBSERVED,8X,8HWEIGHTED/1H ,4X,3HNO.,7X,5HVALUE,10X
   2,5HVALUE,10X,8HRESIDUAL)
880 FORMAT (1H ,2X,I4,1X,3(5X,G11.5))
881 FORMAT (33H0 FINAL COMPUTED NODAL HEAD ARRAY)
882 FORMAT (26H0 NODAL SENSITIVITY ARRAYS)
884 FORMAT (19H0 PARAMETER NUMBER ,I5)
886 FORMAT (1H0,3X,44HSCALED AND ORTHOGONALIZED SENSITIVITY MATRIX
   1/1H ,49H OBSERVATION NUMBER AND VALUES FOR EACH PARAMETER)
888 FORMAT (7H   OBS.,13X,14HPARAMETER NOS./1H ,3X,3HNO.,4X,8(I3,9X))
890 FORMAT (1H ,1X,I4,8(1X,G11.5))
891 FORMAT (24H0ADDITIONAL SOLUTION NO.,I5)
892 FORMAT (1H0,30X,14HNEW PARAMETERS/1H ,3(4X,3HNO.,8X,5HVALUE,4X))
898 FORMAT (16H0 COMPUTED HEADS)
    END
```

Program Listing—Continued

```
      SUBROUTINE ARRAY(A,IND,JND,IT)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION A(IND,JND)
      COMMON/TNME/IIN,IOUT
C**IF IT=0, LOAD 1 AND 2 DIMENSIONAL ARRAYS
C**IF IT=1, PRINT 2 DIMENSIONAL ARRAYS
      IF(IT.EQ.1) GO TO 55
      DO 5 J=1,JND
      DO 5 I=1,IND
    5 A(I,J)=0.
      READ(IIN,65) NME,NOBL,IPRN
      WRITE(IOUT,75) NME
      DO 50 K=1,NOBL
      READ(IIN,70) IB,IE,JB,JE,FACT,IVAR
      WRITE(IOUT,80) K,IB,IE,JB,JE,FACT
      IF(IVAR.GT.0) GO TO 20
      DO 10 J=JB,JE
      DO 10 I=IB,IE
   10 A(I,J)=FACT
      GO TO 50
   20 DO 40 J=JB,JE
      READ(IIN,90) (A(I,J),I=IB,IE)
      DO 40 I=IB,IE
   40 A(I,J)=A(I,J)*FACT
   50 CONTINUE
      IF(IPRN.GT.0) RETURN
      WRITE(IOUT,100) NME
   55 DO 60 K=1,IND,10
      I10=K+9
      IF(I10.GT.IND) I10=IND
      WRITE(IOUT,110) (I,I=K,I10)
      WRITE(IOUT,105)
      DO 60 J=1,JND
      JR=JND-J+1
   60 WRITE(IOUT,120) JR,(A(I,JR),I=K,I10)
      RETURN
C
   65 FORMAT (A4,1X,2I5)
   70 FORMAT (4I5,F10.0,3I5)
   75 FORMAT (1H0,A4)
   80 FORMAT (1H ,I3,2X,5HIB = ,I5,2X,5HIE = ,I5,2X,5HJB = ,I5,2X
     1,5HJE = ,I5,2X,7HFACT = ,G11.5)
   90 FORMAT (8F10.0)
  100 FORMAT (1H0,1X,A4,6H ARRAY)
  105 FORMAT (1H )
  110 FORMAT (1H0,10(9X,I3))
  120 FORMAT (1H ,I3,1X,10(1X,G11.5))
      END
      SUBROUTINE ARRAYI(INT,IND,JND,IT)
      DIMENSION INT(IND,JND)
      COMMON/TNME/IIN,IOUT
C**IF IT=0, LOAD 1 AND 2 DIMENSIONAL INTEGER ARRAYS
C**IF IT=1, PRINT 2 DIMENSIONAL INTEGER ARRAYS
```

Program Listing—Continued

```
      IF(IT.EQ.1) GO TO 45
      DO 5 J=1,JND
      DO 5 I=1,IND
    5 INT(I,J)=0
      READ(IIN,55) NME,NOBL,IPRN
      WRITE(IOUT,65) NME
      DO 40 K=1,NOBL
      READ(IIN,60) IB,IE,JB,JE,IFACT,IVAR
      WRITE(IOUT,70) K,IB,IE,JB,JE,IFACT
      IF(IVAR.GT.0) GO TO 20
      DO 10 J=JB,JE
      DO 10 I=IB,IE
   10 INT(I,J)=IFACT
      GO TO 40
   20 DO 30 J=JB,JE
      READ(IIN,60) (INT(I,J),I=IB,IE)
   30 CONTINUE
   40 CONTINUE
      IF(IPRN.GT.0) RETURN
      WRITE(IOUT,80) NME
   45 DO 50 K=1,IND,30
      I30=K+29
      IF(I30.GT.IND) I30=IND
      WRITE(IOUT,90) (I,I=K,I30)
      WRITE(IOUT,100)
      DO 50 J=1,JND
      JR=JND-J+1
   50 WRITE(IOUT,110) JR,(INT(I,JR),I=K,I30)
      RETURN
C
   55 FORMAT (A4,1X,2I5)
   60 FORMAT (16I5)
   65 FORMAT (1H0,A4)
   70 FORMAT (1H ,I3,2X,5HIB = ,I5,2X,5HIE = ,I5,2X,5HJB = ,I5,2X
     1,5HJE = ,I5,2X,8HIFACT = ,I5)
   80 FORMAT (1H0,1X,A4,6H ARRAY)
   90 FORMAT (1H0,3X,30(1X,I3))
  100 FORMAT (1H )
  110 FORMAT (1H ,31(I3,1X))
      END
      SUBROUTINE ORDER(JPOS,IN,IC)
      DIMENSION JPOS(1),IN(1),IC(5,1)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NOBS,NQSD,NBH
     1,NVAR,NVX2,KOUNT,INDT,IPO
      COMMON/TNME/IIN,IOUT
C**COMPUTE EQUATION NUMBERS FOR D4 ORDERING:
      NXP=ID+JD-1
      K=0
C**ORDER--LEFT TO RIGHT, BOTTOM TO TOP
      DO 20 I=1,NXP,2
      DO 20 J=1,JD
      IK=I-J+1
      IF(IK.LT.1.OR.IK.GT.ID) GO TO 20
```

Program Listing—Continued

```
        N=IK+JPOS(J)
        IF(IN(N).LT.1) GO TO 20
        K=K+1
        IN(N)=K
   20 CONTINUE
        ICR=K+1
        DO 30 I=2,NXP,2
        DO 30 J=1,JD
        IK=I-J+1
        IF(IK.LT.1.OR.IK.GT.ID) GO TO 30
        N=IK+JPOS(J)
        IF(IN(N).LT.1) GO TO 30
        K=K+1
        IN(N)=K
   30 CONTINUE
C**COMPUTE BAND WIDTH AND DETERMINE CONNECTING EQUATION NUMBERS:
        MNO=9999
        MXO=0
        N=0
        IND=0
        DO 80 J=1,JD
        DO 80 I=1,ID
        N=N+1
        JR=IN(N)
        IF(JR.LT.1.OR.JR.GE.ICR) GO TO 80
        IU=1
C**BELOW
        IF((J-1).LT.1.OR.IN(N-ID).LT.1) GO TO 40
        IU=IU+1
        IC(IU,JR)=IN(N-ID)
        MM=IN(N-ID)-JR
        MXO=MAXO(MM,MXO)
        MNO=MINO(MM,MNO)
C**LEFT
   40 IF((I-1).LT.1.OR.IN(N-1).LT.1) GO TO 50
        IU=IU+1
        IC(IU,JR)=IN(N-1)
        MM=IN(N-1)-JR
        MNO=MINO(MM,MNO)
        MXO=MAXO(MM,MXO)
C**RIGHT
   50 IF((I+1).GT.ID.OR.IN(N+1).LT.1) GO TO 60
        IU=IU+1
        IC(IU,JR)=IN(N+1)
        MM=IN(N+1)-JR
        MXO=MAXO(MM,MXO)
        MNO=MINO(MM,MNO)
C**ABOVE
   60 IF((J+1).GT.JD.OR.IN(N+ID).LT.1) GO TO 70
        IU=IU+1
        IC(IU,JR)=IN(N+ID)
        MM=IN(N+ID)-JR
        MXO=MAXO(MM,MXO)
```

Program Listing—Continued

```
              MNO=MINO(MM,MNO)
       70 IC(1,JR)=IU
              IF(IU.GT.1) GO TO 80
              WRITE(IOUT,110) I,J
              IND=1
       80 CONTINUE
              IF(IND.GT.0) STOP
              NEQ=K
              ICR1=ICR-1
              IB1=MXO-MNO+1
              LH1=NEQ-ICR1
              WRITE(IOUT,90)
              WRITE(IOUT,100) ICR1,IB1,LH1,ICR1,NEQ
              RETURN
C
       90 FORMAT (51H0SOLUTION BY LDU FACTORIZATION ASSUMING D4 ORDERING)
      100 FORMAT (65H MINIMUM DIMENSIONS FOR ARRAYS USED BY THIS METHOD ARE
          1AS FOLLOWS/1H ,12H AU:     5 BY,I5/1H ,4H AL:,I5,3H BY,I5/1H
          2,12H IC:     5 BY,I5/1H ,4H  V:,I5)
      110 FORMAT (1H0,13HACTIVE NODE (,I3,1H,,I3,20H) CANNOT BE ISOLATED)
              END
              SUBROUTINE COEF(WELL,HR,HC,CX,CY,VL,QR,CXHR,CXHL,CYHT,CYHB,QBF
          1,B,AU,AL,V,IZN,IBZN,IPRM,IBNA,IBNB,IN,IC,NAD)
              IMPLICIT DOUBLE PRECISION (A-H,O-Z)
              DIMENSION WELL(1),HR(1),HC(1),CX(1),CY(1),VL(1),QR(1),CXHR(1)
          1,CXHL(1),CYHT(1),CYHB(1),QBF(1),B(1),AU(5,1),AL(NAD,1),V(1)
              DIMENSION IZN(1),IBZN(1),IPRM(4,1),IBNA(1),IBNB(1),IN(1),IC(5,1)
              COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NOBS,NQSD,NBH
          1,NVAR,NVX2,KOUNT,INDT,IPO
C**INITIALIZE ARRAYS
              DO 10 J=1,ICR1
              DO 10 I=1,5
       10 AU(I,J)=0.
              DO 20 J=1,LH1
              DO 20 I=1,IB1
       20 AL(I,J)=0.
              DO 40 I=1,NIJ
              N=IN(I)
              IF(N.GT.0) V(N)=WELL(I)
       40 CONTINUE
              IF(NBH.LT.1) GO TO 45
              DO 42 I=1,NBH
              CXHR(I)=0.
              CXHL(I)=0.
              CYHT(I)=0.
       42 CYHB(I)=0.
C**CALCULATE V FOR SPECIFIED FLOW PARAMETERS
       45 IF(NQSD.LT.1) GO TO 52
              DO 50 I=1,NQSD
              IZ=IBZN(I)
              TMP=B(IZ)*QBF(I)
              INA=IBNA(I)
              L=IN(INA)
```

Program Listing—Continued

```
         IF(L.GT.0) V(L)=V(L)+TMP
         INB=IBNB(I)
         L=IN(INB)
         IF(L.GT.0) V(L)=V(L)+TMP
      50 CONTINUE
C**BEGIN MAIN LOOP
      52 N=0
         DO 150 J=1,JM
         DO 150 I=1,IM
         N=N+1
         M=IZN(N)
         IF(M.LT.1) GO TO 150
         LTX=IPRM(1,M)
         LTY=IPRM(2,M)
         LVL=IPRM(3,M)
         LQD=IPRM(4,M)
         NB=N+J
         NA=NB-1
         NC=NB+ID
         ND=NA+ID
         INA=IN(NA)
         INB=IN(NB)
         INC=IN(NC)
         IND=IN(ND)
         CXT=B(LTX)*CX(N)
         CYT=B(LTY)*CY(N)
         VLT=B(LVL)*VL(N)
         QRT=B(LQD)*QR(N)
         E=CXT+CYT+VLT
C**CALCULATE AU, AL, V, AND COEFFICIENT ARRAYS FOR SPECIFIED HEAD
C  PARAMETERS
         K=-INA-1
         IF(K) 60,75,53
      53 CXHR(K)=CXHR(K)+CXT
         CYHT(K)=CYHT(K)+CYT
         GO TO 75
      60 IF(INA.GE.ICR) GO TO 65
         AU(1,INA)=AU(1,INA)+E
         AU(4,INA)=AU(4,INA)-CXT
         AU(5,INA)=AU(5,INA)-CYT
         GO TO 70
      65 AL(1,INA-ICR1)=AL(1,INA-ICR1)+E
      70 V(INA)=V(INA)+QRT+VLT*(HR(NA)-HC(NA))+CXT*(HC(NB)-HC(NA))
        1+CYT*(HC(ND)-HC(NA))
      75 K=-INB-1
         IF(K) 85,100,77
      77 CXHL(K)=CXHL(K)+CXT
         CYHT(K)=CYHT(K)+CYT
         GO TO 100
      85 IF(INB.GE.ICR) GO TO 90
         AU(1,INB)=AU(1,INB)+E
         AU(3,INB)=AU(3,INB)-CXT
         AU(5,INB)=AU(5,INB)-CYT
```

Program Listing—Continued

```
          GO TO 95
       90 AL(1,INB-ICR1)=AL(1,INB-ICR1)+E
       95 V(INB)=V(INB)+QRT+VLT*(HR(NB)-HC(NB))+CXT*(HC(NA)-HC(NB))
         1+CYT*(HC(NC)-HC(NB))
      100 K=-INC-1
          IF(K) 110,125,102
      102 CXHL(K)=CXHL(K)+CXT
          CYHB(K)=CYHB(K)+CYT
          GO TO 125
      110 IF(INC.GE.ICR) GO TO 115
          AU(1,INC)=AU(1,INC)+E
          AU(2,INC)=AU(2,INC)-CYT
          AU(3,INC)=AU(3,INC)-CXT
          GO TO 120
      115 AL(1,INC-ICR1)=AL(1,INC-ICR1)+E
      120 V(INC)=V(INC)+QRT+VLT*(HR(NC)-HC(NC))+CXT*(HC(ND)-HC(NC))
         1+CYT*(HC(NB)-HC(NC))
      125 K=-IND-1
          IF(K) 135,150,127
      127 CXHR(K)=CXHR(K)+CXT
          CYHB(K)=CYHB(K)+CYT
          GO TO 150
      135 IF(IND.GE.ICR) GO TO 140
          AU(1,IND)=AU(1,IND)+E
          AU(2,IND)=AU(2,IND)-CYT
          AU(4,IND)=AU(4,IND)-CXT
          GO TO 145
      140 AL(1,IND-ICR1)=AL(1,IND-ICR1)+E
      145 V(IND)=V(IND)+QRT+VLT*(HR(ND)-HC(ND))+CXT*(HC(NC)-HC(ND))
         1+CYT*(HC(NA)-HC(ND))
      150 CONTINUE
C**COMPRESS AU
          N=0
          DO 190 J=1,JD
          DO 190 I=1,ID
          N=N+1
          K=IN(N)
          IF(K.LT.1.OR.K.GT.ICR1.OR.IC(1,K).EQ.5) GO TO 190
          IU=1
          IF((J-1).LT.1.OR.IN(N-ID).LT.1) GO TO 160
          IU=IU+1
          AU(IU,K)=AU(2,K)
      160 IF((I-1).LT.1.OR.IN(N-1).LT.1) GO TO 170
          IU=IU+1
          AU(IU,K)=AU(3,K)
      170 IF((I+1).GT.ID.OR.IN(N+1).LT.1) GO TO 180
          IU=IU+1
          AU(IU,K)=AU(4,K)
      180 IF((J+1).GT.JD.OR.IN(N+ID).LT.1) GO TO 190
          IU=IU+1
          AU(IU,K)=AU(5,K)
      190 CONTINUE
          RETURN
```

Program Listing—Continued

```
            END
            SUBROUTINE D4SOLV(HC,AU,AL,V,IN,IC,NAD)
            IMPLICIT DOUBLE PRECISION (A-H,O-Z)
            DIMENSION HC(1),AU(5,1),AL(NAD,1),V(1)
            DIMENSION IN(1),IC(5,1)
            COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NOBS,NQSD,NBH
           1,NVAR,NVX2,KOUNT,INDT,IPO
C**DECOMPOSE TO FILL AL
            DO 280 J=1,ICR1
            II=IC(1,J)
            DO 270 I=2,II
            LR=IC(I,J)
            L=LR-ICR1
            C=AU(I,J)/AU(1,J)
            DO 260 K=I,II
            KL=IC(K,J)-LR+1
            AL(KL,L)=AL(KL,L)-C*AU(K,J)
        260 CONTINUE
            AU(I,J)=C
            V(LR)=V(LR)-C*V(J)
        270 CONTINUE
        280 V(J)=V(J)/AU(1,J)
C**DECOMPOSE AL
            JJ=NEQ-ICR
            DO 310 J=1,JJ
            JR=J+ICR1
            L=J
            DO 300 I=2,IB1
            L=L+1
            IF(AL(I,J).EQ.0.) GO TO 300
            LR=L+ICR1
            C=AL(I,J)/AL(1,J)
            KL=0
            DO 290 K=I,IB1
            KL=KL+1
            IF(AL(K,J).NE.0.) AL(KL,L)=AL(KL,L)-C*AL(K,J)
        290 CONTINUE
            AL(I,J)=C
            V(LR)=V(LR)-C*V(JR)
        300 CONTINUE
        310 V(JR)=V(JR)/AL(1,J)
C**BACK SOLVE--LOWER HALF
            V(NEQ)=V(NEQ)/AL(1,NEQ-ICR1)
            DO 330 J=1,JJ
            K=NEQ-J
            KL=K-ICR1
            L=K
            DO 320 I=2,IB1
            L=L+1
            IF(AL(I,KL).NE.0.) V(K)=V(K)-AL(I,KL)*V(L)
        320 CONTINUE
        330 CONTINUE
C**BACK SOLVE--UPPER HALF
```

Program Listing—Continued

```
      DO 350 J=1,ICR1
      K=ICR-J
      II=IC(1,K)
      DO 340 I=2,II
      L=IC(I,K)
      V(K)=V(K)-AU(I,K)*V(L)
  340 CONTINUE
  350 CONTINUE
C**COMPUTE HC+DELTHC
      DO 360 N=1,NIJ
      L=IN(N)
      IF(L.LT.1) GO TO 360
      HC(N)=HC(N)+V(L)
  360 CONTINUE
      RETURN
      END
      SUBROUTINE LSTSQ(HC,BK,BL,BM,BN,HO,W,P,WP,X,C,B,V,KOBS,IN,NVD)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION HC(1),BK(1),BL(1),BM(1),BN(1),HO(1),W(1),P(1),WP(1)
     1,X(NVD,1),C(NVD,1),B(1),V(3)
      DIMENSION KOBS(1),IN(1)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NOBS,NQSD,NBH
     1,NVAR,NVX2,KOUNT,INDT,IPO
      COMMON/TNME/IIN,IOUT
      COMMON/REG/DMX,ADMX,AP,CSA,AMP,RP,BP,YSQ
C**CHECK FOR NONZERO MARQUARDT PARAMETER
      NM1=NVAR-1
      IF(AMP.LT.-.5) GO TO 105
C**INITIALIZE
      DO 20 J=1,NVAR
      DO 10 I=1,NVAR
   10 C(I,J)=0.
   20 V(J)=0.
      YSQ=0.
C**FORM COEFFICIENT MATRIX AND R.H.S. VECTOR
      DO 70 N=1,NOBS
      K=KOBS(N)
      TEMP=HO(N)-BK(N)*HC(K)-BL(N)*HC(K+1)-BM(N)*HC(K+ID)-BN(N)
     1*HC(K+ID+1)
      DO 60 J=1,NVAR
      TMP=W(N)*X(J,N)
      DO 50 I=J,NVAR
   50 C(I,J)=X(I,N)*TMP+C(I,J)
   60 V(J)=TMP*TEMP+V(J)
      YSQ=YSQ+TEMP*W(N)*TEMP
   70 CONTINUE
      IF(NVAR.EQ.1) GO TO 190
      DO 80 I=1,NVAR
      TEMP=C(I,I)+WP(I)
      IF(TEMP.GT.1.E-10) GO TO 78
      WRITE(IOUT,260) I
      INDT=1
      GO TO 80
```

Program Listing—Continued

```
   78 C(I,I)=TEMP**.5
   80 CONTINUE
      IF(INDT.GT.0) RETURN
      DO 100 J=1,NM1
      TEMP=C(J,J)
      JP1=J+1
      DO 90 I=JP1,NVAR
      C(I,J)=C(I,J)/(C(I,I)*TEMP)
   90 C(J,I)=C(I,J)
      V(J)=(V(J)+WP(J)*(P(J)-B(J)))/TEMP+RP*TEMP*(BP*P(J)-B(J))
      V(J+NVAR)=V(J)
      V(J+NVX2)=TEMP
  100 C(J,J)=1.+RP+AMP
      TEMP=C(NVAR,NVAR)
      V(NVAR)=(V(NVAR)+WP(NVAR)*(P(NVAR)-B(NVAR)))/TEMP
     1+RP*TEMP*(BP*P(NVAR)-B(NVAR))
      V(NVX2)=V(NVAR)
      V(NVAR+NVX2)=TEMP
      C(NVAR,NVAR)=1.+RP+AMP
      IF(IPO.NE.1) GO TO 105
      WRITE(IOUT,250)
      CALL PRTOT(C,NVAR,NVD,0)
      WRITE(IOUT,255)
      WRITE(IOUT,230) (V(I),I=1,NVAR)
C**SOLVE FOR V USING LDU FACTORIZATION:
C**DECOMPOSITION AND FORWARD SUBSTITUTION
  105 DET=1.
      DO 140 K=1,NM1
      PIV=C(K,K)
      DET=DET*PIV
      IF(DABS(PIV).GT.1.E-10) GO TO 110
      WRITE(IOUT,210)
      INDT=1
      RETURN
  110 PIV=1./PIV
      KP1=K+1
      DO 130 J=KP1,NVAR
      TMP=C(J,K)*PIV
      DO 120 I=J,NVAR
  120 C(I,J)=C(I,J)-TMP*C(I,K)
  130 V(J)=V(J)-TMP*V(K)
      C(K,K)=PIV
  140 CONTINUE
      DET=DET*C(NVAR,NVAR)
      IF(DABS(C(NVAR,NVAR)).GT.1.E-10) GO TO 150
      WRITE(IOUT,210)
      INDT=1
      RETURN
  150 IF(AMP.LT.-.5) RETURN
C**BACK SUBSTITUTION
      V(NVAR)=V(NVAR)/C(NVAR,NVAR)
      I=NVAR
  160 I=I-1
```

Program Listing—Continued

```
              IF(I.LT.1) GO TO 175
              IP1=I+1
              SUM=0.
              DO 170 J=IP1,NVAR
        170   SUM=SUM+C(J,I)*V(J)
              V(I)=(V(I)-SUM)*C(I,I)
              GO TO 160
     C**CHECK SOLUTION AND ADD MARQUARDT PARAMETER IF NEEDED
        175   TMPA=0.
              TMPB=0.
              TMPC=0.
              DO 176 I=1,NVAR
              TMPA=TMPA+V(I)*V(I)
              TMPB=TMPB+V(I+NVAR)*V(I+NVAR)
        176   TMPC=TMPC+V(I)*V(I+NVAR)
              IF(TMPC.GT.CSA*DSQRT(TMPA*TMPB)) GO TO 200
              AMP=1.5*AMP+.001
              IF(AMP.GT.1.) GO TO 200
              DO 180 I=1,NVAR
              V(I)=V(I+NVAR)
              C(I,I)=1.+RP+AMP
              DO 178 J=I,NVAR
        178   C(J,I)=C(I,J)
        180   CONTINUE
              GO TO 105
     C**SOLUTION WHEN NVAR=1
        190   TEMP=C(1,1)+WP(1)
              IF(TEMP.GT.1.E-10) GO TO 195
              I=1
              WRITE(IOUT,260) I
              INDT=1
              RETURN
        195   V(3)=TEMP**.5
              V(2)=(V(1)+WP(1)*(P(1)-B(1)))/V(3)+RP*V(3)*(BP*P(1)-B(1))
              C(1,1)=1.+RP
              DET=C(1,1)
              V(1)=V(2)/DET
              IF(IPO.NE.1) GO TO 200
              WRITE(IOUT,250)
              CALL PRTOT(C,1,NVD,0)
              WRITE(IOUT,255)
              WRITE(IOUT,230) V(1)
     C**COMPUTE AND PRINT PARAMETERS
        200   ADMX=0.
              DO 203 J=1,NVAR
              V(J)=V(J)/V(J+NVX2)
              TMPA=1.
              IF(B(J)) 201,202,201
        201   TMPA=B(J)
        202   TMP=DABS(V(J)/TMPA)
              IF(TMP.GT.ADMX) ADMX=TMP
        203   CONTINUE
              AP=1.
```

Program Listing—Continued

```
       IF(ADMX.GT.DMX) AP=DMX/ADMX
       DO 204 J=1,NVAR
       V(J)=AP*V(J)
   204 B(J)=V(J)+B(J)
       WRITE(IOUT,220) KOUNT,YSQ,DET,AMP,AP
       WRITE(IOUT,230) (B(J),J=1,NVAR)
       RETURN
C
   210 FORMAT (42H0LEAST SQUARES COEFFICIENT MATRIX SINGULAR
      1/35H SOLUTION FOR PARAMETERS NOT UNIQUE)
   220 FORMAT (1H0,14HITERATION NO. ,I3/1H ,6HYSQ = ,G11.5,2X
      1,9HDET(C) = ,G11.5,2X,6HAMP = ,G11.5,2X,5HAP = ,G11.5
      2/1H ,21HREGRESSION PARAMETERS)
   230 FORMAT ((1H ,8(G11.5,2X)))
   250 FORMAT (29H0 SCALED LEAST SQUARES MATRIX)
   255 FORMAT (24H0 SCALED GRADIENT VECTOR)
   260 FORMAT (29H0SENSITIVITIES FOR PARAMETER ,I4,17H EFFECTIVELY ZERO)
       END
       SUBROUTINE PRTOT(C,NO,NOD,IT)
       IMPLICIT DOUBLE PRECISION (A-H,O-Z)
       DIMENSION C(1)
       COMMON/TNME/IIN,IOUT
C**IF IT=0, PRINT SYMMETRIC MATRIX DIVIDED VERTICALLY INTO TEN-COLUMN
C  BLOCKS
C**IF IT=1, PRINT VECTOR IN THREE COLUMNS
       IF(IT.EQ.1) GO TO 25
       DO 20 L=1,NO,10
       J10=L+9
       IF(J10.GT.NO) J10=NO
       WRITE(IOUT,30) (J,J=L,J10)
       WRITE(IOUT,50)
       K=-NOD
       DO 10 I=1,NO
       K=K+NOD
    10 WRITE (IOUT,40) I,(C(J+K),J=L,J10)
       WRITE(IOUT,60)
    20 CONTINUE
       RETURN
    25 NR=NO/3
       IF((3*NR).NE.NO) NR=NR+1
       DO 26 K=1,NR
    26 WRITE(IOUT,80) (L,C(L),L=K,NO,NR)
       RETURN
C
    30 FORMAT (1H0,8X,I3,9(9X,I3))
    40 FORMAT (1H ,I3,10(1X,G11.5)
    50 FORMAT (1H )
    60 FORMAT (1H0)
    80 FORMAT (1H ,3X,3(I3,7X,G11.5,3X))
       END
```

Program Listing—Continued

```
      SUBROUTINE HOBS(DX,DY,X,Y,BK,BL,BM,BN,HO,W,KOBS)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION DX(1),DY(1),X(1),Y(1),BK(1),BL(1),BM(1),BN(1),HO(1)
     1,W(1)
      DIMENSION KOBS(1)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NOBS,NQSD,NBH
     1,NVAR,NVX2,KOUNT,INDT,IPO
      COMMON/TNME/IIN,IOUT
C**COMPUTE X-LOCATIONS OF NODE POINTS
      X(1)=0.
      DO 10 I=2,ID
   10 X(I)=X(I-1)+DX(I-1)
C**COMPUTE Y-LOCATIONS OF NODE POINTS
      Y(1)=0.
      DO 20 J=2,JD
   20 Y(J)=Y(J-1)+DY(J-1)
      WRITE(IOUT,40)
      DO 30 I=1,NOBS
C**READ OBSERVED HEAD DATA
      READ(IIN,50) N,IL,JL,XL,YL,HO(N),W(N)
      WRITE(IOUT,60) N,IL,JL,XL,YL,HO(N),W(N)
C**COMPUTE LOCATION OF FIRST NODE IN CELL (IL,JL)
      K=IL+IM*(JL-1)
      KOBS(N)=K+(K-1)/IM
C**COMPUTE WEIGHTS FOR BILINEAR INTERPOLATION
      AREA=DX(IL)*DY(JL)
      BK(N)=(X(IL+1)-XL)*(Y(JL+1)-YL)/AREA
      BL(N)=(XL-X(IL))*(Y(JL+1)-YL)/AREA
      BM(N)=(X(IL+1)-XL)*(YL-Y(JL))/AREA
   30 BN(N)=(XL-X(IL))*(YL-Y(JL))/AREA
C
   40 FORMAT (1H0,25X,18HOBSERVED HEAD DATA/1H ,6H  OBS.,2X,9HCELL LOC.
     1,3X,6HX LOC.,7X,6HY LOC.,8X,4HOBS.,8X,6HWEIGHT/1H ,6H   NO.,4X
     2,1HI,5X,1HJ,30X,4HHEAD)
   50 FORMAT (3I5,4F10.0)
   60 FORMAT (1H ,1X,3(I4,2X),4(G11.5,2X))
      RETURN
      END
```

References Cited

Cooley, R.L., 1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1—Theory: Water Resources Research, v. 18, no. 4, p. 965-976.

———1983, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 2—Applications: Water Resources Research, v. 19, no. 3, p. 662-676.

Price, H.S., and Coats, K.H., 1974, Direct methods in reservoir simulation: Society of Petroleum Engineers Journal, June 1974, p. 295-308.

Wang, H.F., and Anderson, M.P., 1982, Introduction to groundwater modeling—Finite difference and finite element methods: San Francisco, W.H. Freeman, 237 p.

Additional Reading

Cooley, R.L., 1977, A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 1—Theory and numerical properties: Water Resources Research, v. 13, no. 2, p. 318-324.

———1979, A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 2—Application of statistical analysis: Water Resources Research, v. 15, no. 3, p. 603-617.

———1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1—Theory: Water Resources Research, v. 18, no. 4, p. 965-976.

———1983, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 2—Applications: Water Resources Research, v. 19, no. 3, p. 662-676.

5 Elementary Analysis and Use of the Regression Model

5.1 Assumed Forms of Model Equations

As a purely algebraic process, regression contains no assumptions other than those already mentioned. However, to statistically analyze results of, and predictions to be made by, the method, additional assumptions must be made. Based on these assumptions, an effective methodology has been developed (see, for example, Draper and Smith, 1981) to analyze and use a linear regression model.

The statistical methods also may be applied to a nonlinear model, provided the model is close enough to being linear. Fortunately, whether or not the model is close enough can usually be determined. All statistics and procedures are, accordingly, derived for a linear, or effectively linear, model. To make the equations applicable for both a linear and nonlinear model, the basic types of models assumed are the incremental linear model and the nonlinear model as linearized using the Taylor series expansion.

The model assumed, then, is of the form

$$\underline{f}_\beta - \underline{f}_0 \cong \underline{\underline{X}}(\underline{\beta} - \underline{b}_0) \qquad (5.1-1)$$

where strict equality applies for a linear model and, for a nonlinear model, $\underline{\underline{X}}$ is assumed to be evaluated at $\underline{b}_0$. Also, for simplicity of notation, define

$$\underline{f}_\beta = \underline{f}(\underline{\xi}, \underline{\beta}) \qquad (5.1-2)$$

$$\underline{f}_0 = (\underline{\xi}, \underline{b}_0) \ . \qquad (5.1-3)$$

Based on equation 5.1-1, the true regression model is

$$\underline{Y} - \underline{f}_0 \cong \underline{\underline{X}}(\underline{\beta} - \underline{b}_0) + \underline{\epsilon} \qquad (5.1-4)$$

where strict equality only holds for a linear model because $\underline{\epsilon}$ is the true vector of disturbances. The estimated regression model derived from equation 5.1-4 is

$$\underline{Y} - \underline{f}_0 \cong \underline{\underline{X}}(\underline{b} - \underline{b}_0) + \underline{e} \qquad (5.1-5)$$

where, as for equation 5.1-4, strict equality only holds for a linear model because $\underline{e}$ is assumed to be the true vector of residuals defined by $\underline{e} = \underline{Y} - \underline{f}(\underline{\xi}, \underline{b})$.

By minimizing $S(\underline{b}) = \underline{e}^T \underline{\underline{\omega}} \underline{e}$ with respect to $\underline{b}$ using the standard procedure, exact best-fit estimates $\underline{b}$ of $\underline{\beta}$ and $\underline{\hat{e}} = \underline{Y} - \underline{f}(\underline{\xi}, \underline{\hat{b}})$ of $\underline{\epsilon}$ are obtained. For a linear model $\underline{\hat{b}}$ is obtained exactly by using equation 5.1-5 as the estimated regression model. For a nonlinear model, use of the linearized model leads to an approximate relationship to find $\underline{\hat{b}}$. Thus, by minimizing $S(\underline{b})$ using equation 5.1-5 as the estimated regression model, normal equations

$$\underline{\underline{X}}^T \underline{\underline{\omega}} \underline{\underline{X}}(\underline{\hat{b}} - \underline{b}_0) \cong \underline{\underline{X}}^T \underline{\underline{\omega}}(\underline{Y} - \underline{f}_0) \qquad (5.1-6)$$

that are approximate for a nonlinear model are derived. The regression model obtained by replacing general estimates $\underline{b}$ and $\underline{e}$ in equation 5.1-5 with best-fit estimates $\underline{\hat{b}}$ and $\underline{\hat{e}}$ is

$$\underline{Y} - \underline{f}_0 \cong \underline{\underline{X}}(\underline{\hat{b}} - \underline{b}_0) + \underline{\hat{e}} \ . \qquad (5.1-7)$$

By utilizing the definition of $\underline{\hat{e}}$ ($\underline{\hat{e}} = \underline{Y} - \underline{\hat{f}}$, where $\underline{\hat{f}} = \underline{f}(\underline{\xi}, \underline{\hat{b}})$), in equation 5.1-7, a predictive model

$$\underline{\hat{f}} - \underline{f}_0 \cong \underline{\underline{X}}(\underline{\hat{b}} - \underline{b}_0) \qquad (5.1-8)$$

is obtained.

A final point is the establishment of the general condition for a minimum in $S(\underline{b})$. If $\underline{b}_0$ is set equal to $\underline{\hat{b}}$ in equation 5.1-6, then $\underline{f}_0 = \underline{\hat{f}}$, and

$$\underline{\underline{X}}^T \underline{\underline{\omega}}(\underline{Y} - \underline{\hat{f}}) = \underline{0} \ . \qquad (5.1-9)$$

Because any approximation inherent in equation 5.1-6 is removed as $\underline{b}_0 \rightarrow \underline{\hat{b}}$, equation 5.1-9 is exact for both linear and nonlinear models. The left side of equation 5.1-9 is the negative of the gradient of $S(\underline{b})$. Thus, equation 5.1-9 states that the gradient of $S(\underline{b})$ is zero at a minimum point of $S(\underline{b})$.

From here on, for simplicity the approximate equality sign in regression models and normal equations (for example, 5.1-1, 5.1-4 through 5.1-8) is replaced by an equal sign. However,

remember that all relationships derived by using the linearized model are approximate for a nonlinear model.

5.2 Assumptions of Regression Modeling

Some of the assumptions listed below have already been mentioned; they are discussed more completely here.

1. A true model exists:

$$Y = f(\xi_1, \xi_2, ..., \xi_k; \underline{\beta}) + \epsilon \ . \qquad (5.2\text{-}1)$$

The model response, Y, consists of two parts, a deterministic part, f, and an additive stochastic part, ϵ.

2. The disturbances, $\underline{\epsilon}$, have the following properties

$$E(\underline{\epsilon}) = \underline{0} \qquad (5.2\text{-}2)$$

$$\text{Var}(\underline{\epsilon}) = \underline{V}\sigma^2 \qquad (5.2\text{-}3)$$

where the structure, $\underline{V}$, of the variance-covariance matrix $\underline{V}\sigma^2$ is assumed to be symmetric positive definite and known. Alternative forms for equations 5.2-2 and 5.2-3 are obtained by premultiplying equation 5.2-2, and pre- and postmultiplying equation 5.2-3, by $\underline{V}^{-\frac{1}{2}}$ to obtain

$$E(\underline{V}^{-\frac{1}{2}}\underline{\epsilon}) = \underline{0} \qquad (5.2\text{-}4)$$

$$\text{Var}(\underline{V}^{-\frac{1}{2}}\underline{\epsilon}) = \underline{I}\sigma^2 \ . \qquad (5.2\text{-}5)$$

The assumptions given by equations 5.2-2 through 5.2-5 indicate that $\underline{\epsilon}$ is considered to be a vector of random variables with zero mean and variance-covariance matrix $\underline{V}\sigma^2$. Furthermore, weighted disturbances $\underline{V}^{-\frac{1}{2}}\underline{\epsilon}$ have constant variance $\underline{I}\sigma^2$ and are uncorrelated. To require the expected value of $\underline{\epsilon}$ to be zero is to require that equation 5.2-1 be the true (or unbiased) model and to require in addition that $\underline{\epsilon}$ be unbiased. Although imperfections in most physical theories prevent the former assumption from holding strictly true, a model should be constructed so that the absolute value of any $E(\epsilon_j)$ is as small as possible. From the practical

point of view, it is required that the bias not be significant. Criteria for this are developed later on.

The full form of $\underline{V}$ is usually very difficult to obtain from the type of data usually available. However, if $\underline{V}$ is assumed to be diagonal so that there is no correlation among the ϵ_j, then $\underline{V}$ can often be found by using graphical methods of analyzing residuals, to be discussed later on.

3. The matrices $\underline{\omega}$ and $\underline{V}^{-\frac{1}{2}}$ are equivalent; that is,

$$\underline{\omega} = \underline{V}^{-1} \ . \qquad (5.2\text{-}6)$$

For a linear model at least, the Gauss-Markov theorem (Beck and Arnold, 1977, p. 232-234) establishes that the variance of $\hat{b}_j$, $\text{Var}(\hat{b}_j)$, is a minimum if equation 5.2-6 is true. Furthermore, to compute $\text{Var}(\hat{\underline{b}})$ correctly, whether or not equation 5.2-6 is true, $\underline{V}$ must be known. Hence, assumption of another form for $\underline{\omega}$ (such as $\underline{I}$, for example) would not avoid the problem of having to know $\underline{V}$ to analyze the model. However, for a linear model, equation 5.2-6 is not essential to compute an unbiased estimate of $\underline{\beta}$. This fact may be demonstrated as follows. Solve equation 5.1-6 for $\hat{\underline{b}} - \underline{b}_0$ and take the expected value of it to obtain:.

$$E(\hat{\underline{b}} - \underline{b}_0) = (\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}E(\underline{Y} - \underline{f}_0)$$

$$= (\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}\underline{X}(\underline{\beta} - \underline{b}_0) = \underline{\beta} - \underline{b}_0 \qquad (5.2\text{-}7)$$

where the fact that $E(\underline{\epsilon}) = 0$ was used. From equation 5.2-7 it is seen that

$$E(\hat{\underline{b}}) = \underline{\beta} \ . \qquad (5.2\text{-}8)$$

If $\underline{V}$ is diagonal, then

$$\underline{V} = \begin{bmatrix} 1/\omega_1 & & & \\ & 1/\omega_2 & & \\ & & \ddots & \\ & & & 1/\omega_n \end{bmatrix} \qquad (5.2\text{-}9)$$

where the double subscripts on ω have been replaced by single subscripts to indicate the diagonal nature of $\underline{V}$.

4. The disturbances are normally distributed:

$$\underline{\epsilon} \sim N(\underline{0}, \underline{V}\sigma^2) \qquad (5.2\text{-}10)$$

or

$$\underline{V}^{-\frac{1}{2}}\underline{\epsilon} \sim N(\underline{0}, \underline{I}\sigma^2) \ . \qquad (5.2\text{-}11)$$

Assumption of either equation 5.2-10 or 5.2-11 is only necessary if investigations using the F distribution are to be performed.

That $\underline{\epsilon}$ (or $\underline{V}^{-\frac{1}{2}}\underline{\epsilon}$) be normally distributed implies that the elements of $\underline{V}^{-\frac{1}{2}}\underline{\epsilon}$ are neither systematic nor constant but are equally likely to be positive or negative. In addition, small errors are more frequent than large ones. Many types of models are subject to a number of sources of error, any one of which may or may not be normally distributed. However, in the case where a resultant error is the sum of a number of components, Central Limit Theorem implies that $\underline{\epsilon}$ (or $\underline{V}^{-\frac{1}{2}}\underline{\epsilon}$) could be normally distributed even if its component vectors were not.

Because $\underline{\epsilon}$ and $\underline{\beta}$ are unknown, the assumptions (1 through 4) discussed cannot be checked directly. However, they may often be checked indirectly, which is a subject of model analysis.

5.3 Relationships Between Residuals and Disturbances

Many of the investigations involving the regression model are based either directly or indirectly on relationships between residuals $\underline{\hat{e}}$ and disturbances $\underline{\epsilon}$. Residuals may be written in terms of disturbances by employing equations 5.1-6 and 5.1-7. First, equation 5.1-7 is written in the form

$$\underline{\hat{e}} = \underline{Y} - \underline{f}_0 - \underline{X}(\underline{\hat{b}} - \underline{b}_0) \ . \qquad (5.3\text{-}1)$$

Then equation 5.1-6 is solved for $\underline{\hat{b}} - \underline{b}_0$ and substituted into equation 5.3-1 to obtain

$$\underline{\hat{e}} = \underline{Y} - \underline{f}_0 - \underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}(\underline{Y} - \underline{f}_0)$$

$$= (\underline{I} - \underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega})(\underline{Y} - \underline{f}_0) \ . \qquad (5.3\text{-}2)$$

If $\underline{b}_0$ is set equal to β, then $\underline{Y} - \underline{f}_0 = \underline{Y} - \underline{f}_\beta = \underline{\epsilon}$ and equation 5.3-2 gives

$$\underline{\hat{e}} = (\underline{I} - \underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega})\underline{\epsilon} \ . \qquad (5.3\text{-}3)$$

It is frequently more convenient to work with weighted residuals, $\underline{\omega}^{\frac{1}{2}}\underline{\hat{e}}$, and weighted disturbances, $\underline{\omega}^{\frac{1}{2}}\underline{\epsilon}$. In this case equation 5.3-3 becomes

$$\underline{\omega}^{\frac{1}{2}}\underline{\hat{e}} = (\underline{I} - \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}})\underline{\omega}^{\frac{1}{2}}\underline{\epsilon} \ . \qquad (5.3\text{-}4)$$

An interesting and useful property of the matrix appearing in either equation 5.3-3 or 5.3-4 is displayed, for equation 5.3-4 for example, as follows:

$$[\underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}][\underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}]$$

$$= \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}\underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}$$

$$= \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}} \ . \qquad (5.3\text{-}5)$$

In other words, the matrix times itself yields the original matrix. This result is true for both equations 5.3-3 and 5.3-4. For equation 5.3-4 it is also true that the matrix is symmetric (as can be seen in the derivation of equation 5.3-5), so that the matrix times its transpose yields the original matrix. This type of matrix is known as a symmetric idempotent matrix.

Another useful property can be derived based on the idempotency discussed above. For convenience let

$$\underline{R} = \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}} \qquad (5.3\text{-}6)$$

Then because $\underline{R}$ is idempotent,

$$(\underline{I} - \underline{R})(\underline{I} - \underline{R}) = \underline{I} - \underline{R} - \underline{R} + \underline{R}$$

$$= \underline{I} - \underline{R} \ . \qquad (5.3\text{-}7)$$

Hence, $\underline{I} - \underline{R}$ is also idempotent.

5.4 Some Statistical Measures

The first step in model analysis should always be to examine some statistical measures that indicate (1) goodness of fit of the model to the data and (2) model conditioning as it affects reliability of the computed parameters. With the background given above, useful statistical measures can be derived.

5.4.1 The Error Variance, s^2

For a linear model this measure is an unbiased estimate of σ^2. For a nonlinear model, s^2 is biased. A sketch of the derivation follows: The sum of squares, $S(\hat{\underline{b}})$, is defined as

$$S(\hat{\underline{b}})=\hat{\underline{e}}^T\underline{\omega}\hat{\underline{e}}$$

$$=(\underline{\omega}^{1/2}\hat{\underline{e}})^T(\underline{\omega}^{1/2}\hat{\underline{e}}) \quad . \tag{5.4-1}$$

Using equations 5.3-4, 5.3-7, and the fact that tr(scaler)=scaler, equation 5.4-1 becomes

$$S(\hat{\underline{b}})=(\underline{\omega}^{1/2}\underline{\epsilon})^T(\underline{I}-\underline{R})^T(\underline{I}-\underline{R})(\underline{\omega}^{1/2}\underline{\epsilon})$$

$$=tr[(\underline{\omega}^{1/2}\underline{\epsilon})^T(\underline{I}-\underline{R})(\underline{\omega}^{1/2}\underline{\epsilon})]$$

$$=tr[(\underline{I}-\underline{R})(\underline{\omega}^{1/2}\underline{\epsilon})(\underline{\omega}^{1/2}\underline{\epsilon})^T] \quad . \tag{5.4-2}$$

The expected value of equation 5.4-2 is

$$E(S(\hat{\underline{b}}))=tr\{(\underline{I}-\underline{R})E[(\underline{\omega}^{1/2}\underline{\epsilon})(\underline{\omega}^{1/2}\underline{\epsilon})^T]\}$$

$$=tr[(\underline{I}-\underline{R})Var(\underline{\omega}^{1/2}\underline{\epsilon})]$$

$$=tr[(\underline{I}-\underline{R})\sigma^2]$$

$$=(n-p)\sigma^2 \tag{5.4-3}$$

from which

$$\sigma^2=\frac{E(S(\hat{\underline{b}}))}{n-p} \quad . \tag{5.4-4}$$

The fact that $tr(\underline{R})=p$ can be demonstrated by rearranging the matrices within $\underline{R}$. The estimate, s^2, of σ^2 is

$$s^2=\frac{S(\hat{\underline{b}})}{n-p}=\frac{(\underline{Y}-\hat{\underline{f}})^T\underline{\omega}(\underline{Y}-\hat{\underline{f}})}{n-p} \quad . \tag{5.4-5}$$

A useful modification of equation 5.4-5 that is exact for a linear model and almost exact for a nonlinear one is obtained by choosing $\underline{b}_0$ in equation 5.1-7 to be very near $\hat{\underline{b}}$. Then

$$s^2=\frac{(\underline{Y}-\underline{f}_0-\underline{X}(\hat{\underline{b}}-\underline{b}_0))^T\underline{\omega}(\underline{Y}-\underline{f}_0-\underline{X}(\hat{\underline{b}}-\underline{b}_0))}{n-p}$$

$$=\frac{(\underline{Y}-\underline{f}_0)^T\underline{\omega}(\underline{Y}-\underline{f}_0)-2(\hat{\underline{b}}-\underline{b}_0)^T\underline{X}^T\underline{\omega}(\underline{Y}-\underline{f}_0)+(\hat{\underline{b}}-\underline{b}_0)^T\underline{X}^T\underline{\omega}\underline{X}(\hat{\underline{b}}-\underline{b}_0)}{n-p}$$

$$=\frac{(\underline{Y}-\underline{f}_0)^T\underline{\omega}(\underline{Y}-\underline{f}_0)-(\hat{\underline{b}}-\underline{b}_0)^T\underline{X}^T\underline{\omega}(\underline{Y}-\underline{f}_0)}{n-p} \tag{5.4-6}$$

where use was made of equation 5.1-6. For a linear model $\underline{b}_0$ may be chosen to be $\underline{0}$. For a nonlinear model a bias exists in equations 5.4-5 or 5.4-6 that results from the fact that development is based on assumption of a linear model.

Even when biased, s^2 gives a useful measure of overall goodness of fit of the model. The standard deviation or scatter is given by s. In general, $s/\Delta Y_s$ should be small, where ΔY_s is the difference between maximum and minimum values of Y_s.

5.4.2 The Correlation, R_y, Between $\underline{\omega}^{1/2}\underline{Y}$ and $\underline{\omega}^{1/2}\hat{\underline{f}}$

This measure is defined as

$$R_y=\frac{\underline{d}_y^T\hat{\underline{d}}_y}{[(\underline{d}_y^T\underline{d}_y)(\underline{d}_f^T\underline{d}_f)]^{1/2}} \tag{5.4-7}$$

where

$$\underline{d}_y=\underline{\omega}^{1/2}\underline{Y}-m_y\underline{l} \tag{5.4-8}$$

$$\underline{d}_f=\underline{\omega}^{1/2}\hat{\underline{f}}-m_f\underline{l} \tag{5.4-9}$$

$$m_y=\sum_{i=1}^{n}\underline{\omega}_i^{1/2}\underline{Y}/n \tag{5.4-10}$$

$$m_f=\sum_{i=1}^{n}\underline{\omega}_i^{1/2}\hat{\underline{f}}/n \tag{5.4-11}$$

$\underline{l}$=vector of ones, and $\underline{\omega}_i^{1/2}$ is row i of $\underline{\omega}^{1/2}$. The correlation R_y is another measure of goodness of fit. Usually it should be greater than about 0.9 to indicate a good fit to the reliable data.

5.4.3 The Variance-Covariance Matrix for $\hat{\underline{b}}$

This measure may be derived directly from equation 5.1-6 and is

$$\mathrm{Var}(\hat{\underline{b}}) = \mathrm{Var}[(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}(\underline{Y}-\underline{f}_0) + \underline{b}_0]$$

$$= (\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}\mathrm{Var}(\underline{Y})\underline{\omega}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}$$

$$= (\underline{X}^T\underline{\omega}\underline{X})^{-1}\sigma^2 \qquad (5.4\text{-}12)$$

where assumptions from equations 5.2-3 and 5.2-6 and the fact that $\mathrm{Var}(\underline{Y}) = \mathrm{Var}(\underline{Y}-\underline{f}_\beta) = \underline{\omega}^{-1}\sigma^2$ were employed. The estimate of $\mathrm{Var}(\hat{\underline{b}})$ is

$$\mathrm{Var}(\hat{\underline{b}}) = (\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2 \ . \qquad (5.4\text{-}13)$$

The standard error of the estimate for the ith parameter is given by the square root of the ith diagonal component of $(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2$. This estimate is a measure of the range over which the respective parameter may be varied to produce a similar solution for the dependent variable as that obtained using $\hat{\underline{b}}$.

5.4.4 The Correlation, r_{ij}, Between any Two Parameters $\hat{b}_i$ and $\hat{b}_j$

By definition

$$r_{ij} = \frac{\mathrm{Cov}(\hat{b}_i, \hat{b}_j)}{[\mathrm{Var}(\hat{b}_i)\mathrm{Var}(\hat{b}_j)]^{1/2}} \qquad (5.4\text{-}14)$$

where the variance and covariance terms are components of either $(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2$ or $(\underline{X}^T\underline{\omega}\underline{X})^{-1}\sigma^2$. This measure gives an estimate of the degree of linear dependence of one parameter on another throughout the course of repeated experiments if such experiments were to be carried out. As discussed earlier, it is an indication of the degree of linear dependency in the sensitivity matrix.

Problem 5.4-1

This problem is concerned with preliminary analysis of the linear regression solution of problem 3.2-1. The measures (except $s/\Delta Y_s$), to be computed in a, b, and c below, also are calculated by the computer model of problem 4.2-1. Check your computations against the computer-generated results.
  a. Using equation 5.4-6, compute s^2. Compute $s/\Delta Y_s$. Would you say that the fit is very good?

  b. Using equation 5.4-13, compute $\mathrm{Var}(\hat{\underline{b}})$. Are the parameters determined very precisely?
  c. Using equation 5.4-14, determine $\underline{r}$, the correlation matrix. Are there any evident problems with conditioning?

5.5 Analysis of Residuals

Examination of the statistics discussed in the previous section should give a preliminary indication of general model conditioning and model fit to the data. However, a thorough analysis of residuals is necessary in order to examine the validity of the assumptions given in section 5.2. Interest is focused primarily on indications of nonrandomness of the residuals and on indications that the residuals are not distributed normally. The analysis should include both sample and prior information partitions of the residuals so that any incompatibility between the two partitions can be detected as differences between the two partitions. Although the techniques given in the present section are usually adequate to detect any incompatibility, a formal test given in section 6.3 also can be applied if desired.

Analytical methods used here are graphical. Draper and Smith (1981, p. 141-192) give a number of methods for examining residuals, and they emphasize that graphical procedures involving visual analysis are the most valuable tools because violations of assumptions serious enough to require corrective action generally are apparent on the various plots. However, to use the procedures effectively it is necessary to determine the properties that the residuals should be expected to exhibit under ideal conditions.

5.5.1 Distribution of Residuals

Investigation of the distribution of weighted residuals $\underline{\omega}_j^{1/2}\hat{\underline{e}}$, where $\underline{\omega}_j^{1/2}$ is a row of $\underline{\omega}^{1/2}$, made in order to infer the distribution of weighted disturbances $\underline{\omega}_j^{1/2}\underline{\epsilon}$, is difficult because, even if the assumption given by equation 5.2-5 holds so that the elements $\underline{\omega}_j^{1/2}\underline{\epsilon}$ are uncorrelated and have equal variance, elements of $\underline{\omega}_j^{1/2}\hat{\underline{e}}$ are correlated and have unequal variance. To show this for the linearized model, equations 5.3-4 and 5.3-6 can be combined to give

$$\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}} = (\underline{I}-\underline{R})\underline{\omega}^{\frac{1}{2}}\underline{\epsilon} \qquad (5.5\text{-}1)$$

from which

$$E(\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}}) = (\underline{I}-\underline{R})\underline{\omega}^{\frac{1}{2}}E(\underline{\epsilon})$$

$$= \underline{0} \qquad (5.5\text{-}2)$$

and

$$\mathrm{Var}(\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}}) = \mathrm{Var}[(\underline{I}-\underline{R})\underline{\omega}^{\frac{1}{2}}\underline{\epsilon}]$$

$$= (\underline{I}-\underline{R})\mathrm{Var}(\underline{\omega}^{\frac{1}{2}}\underline{\epsilon})(\underline{I}-\underline{R})^T$$

$$= (\underline{I}-\underline{R})(\underline{I}-\underline{R})^T\sigma^2$$

$$= (\underline{I}-\underline{R})\sigma^2 \ . \qquad (5.5\text{-}3)$$

Hence, if assumptions given by equations 5.2–4 through 5.2–6, and 5.2–11 hold,

$$\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}} = \hat{\underline{u}} \sim N(\underline{0},(\underline{I}-\underline{R})\sigma^2) \qquad (5.5\text{-}4)$$

where for convenience, the definition is made that $\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}}=\hat{\underline{u}}$.

It can be shown that $\underline{I}\text{-}\underline{R}{\to}\underline{I}$ as $n\text{-}p{\to}\infty$. Whenever $n\text{-}p$ becomes small, correlation and unequal variance become significant. Most tests for distribution of residuals assume equal (or, a common) variance and no correlation because all residuals are assumed to have come from the same univariate distribution. Therefore, correlation and unequal variance of the $\hat{u}_j$ are serious problems with regard to testing for normality when the number of parameters is not small compared to the number of observations.

Another difficulty concerns the determination of whether or not the model fits the data. If the model fits the data and correlation of the values of $\hat{u}_j$ is not significant, then these residuals should appear to be nearly random. However, if correlation is significant, then the correlation will be reflected in the residual values. Patterns could develop in some of the plots (to be discussed), and these patterns could be mistakenly attributed to lack of model fit.

5.5.2 Graphical Procedures

The first step in using graphical procedures is to develop a control group. Several sets of simulated residuals distributed as in equation 5.5–4 form the control group. These sets are then compared graphically with the true weighted residuals $\hat{\underline{u}}$ to help decide whether the distribution of $\hat{\underline{u}}$ differs to a visually detectable extent from a normal distribution and whether correlation could cause an apparently nonrandom (or non-normal) pattern of residuals to develop in the residual plots.

A set of simulated residuals may be generated by generating a set of uncorrelated random normal deviates $\underline{d}$ so that $E(\underline{d})=\underline{0}$ and $\mathrm{Var}(\underline{d})=\underline{I}s^2$, then forming linear combinations of these deviates that have the covariance given by equation 5.5–3. The method of generating the simulated residuals from the uncorrelated random deviates can be derived as follows. Assume, as a working hypothesis, that

$$\underline{g} = \underline{\Omega}\underline{d} \qquad (5.5\text{-}5)$$

where $\underline{g}$ is the set of simulated residuals, and $\underline{\Omega}$ is a symmetric and nonstochastic matrix to be determined. Vector $\underline{g}$ must be generated so that $E(\underline{g})=\underline{0}$ and $\mathrm{Var}(\underline{g})=(\underline{I}-\underline{R})s^2$ (equations 5.5–2 and 5.5–3). From equation 5.5–5

$$E(\underline{g}) = \underline{\Omega}E(\underline{d})$$

$$= \underline{0} \qquad (5.5\text{-}6)$$

as required. By definition

$$\mathrm{Var}(\underline{g}) = \underline{\Omega}\mathrm{Var}(\underline{d})\underline{\Omega}^T$$

$$= \underline{\Omega}^2 s^2 \qquad (5.5\text{-}7)$$

where the definition of $\mathrm{Var}(\underline{d})$ and the symmetry of $\underline{\Omega}$ were used. Hence, $\underline{\Omega}$ must be defined so that

$$\underline{\Omega}^2 = \underline{I}\text{-}\underline{R} \ . \qquad (5.5\text{-}8)$$

However, because $\underline{I}\text{-}\underline{R}$ is idempotent, $(\underline{I}\text{-}\underline{R}) = (\underline{I}\text{-}\underline{R})^2$, and equation 5.5–8 may be simplified to become

$$\underline{\Omega} = \underline{I}\text{-}\underline{R} \qquad (5.5\text{-}9)$$

so that equation 5.5–5 assumes as its final form

$$\underline{g} = (\underline{I}\text{-}\underline{R})\underline{d} \ . \qquad (5.5\text{-}10)$$

To generate a set of simulated residuals, g, it is a simple matter to generate a set of uncorrelated random normal deviates, $\underline{d}$, then use equation 5.5-10. This procedure is followed for the number of sets (usually at least three) desired to form the control group.

Normal probability plots.—These are graphs of cumulative frequency, F, versus values of the elements of vectors $\underline{\hat{u}}$, $\underline{d}$, or $\underline{g}$. Cumulative frequency corresponding to the ith element of one of the vectors ($\underline{\hat{u}}$, for example) is computed from the formula

$$F_i = m_i/(n+1), i=1,2, ...,n \qquad (5.5\text{-}11)$$

where m_i is the number of values of $\underline{\hat{u}}$ (for example) smaller than or equal to $\hat{u}_i$, and n is the number of observations. Use of $n+1$ in the denominator adjusts for the fact that F_n cannot be equal to 1 because the tail of the normal distribution extends to infinity. An example of a normal probability plot for $\underline{\hat{u}}$ is illustrated in figure 5.5-1.

To determine the effects of correlation and unequal variance, normal probability plots for the sets $\underline{d}$ and the sets $\underline{g}$ may be compared. If the plots for $\underline{d}$ and $\underline{g}$ are very similar, then a normal probability plot of $\underline{\hat{u}}$ would not be expected to be affected to a great extent by
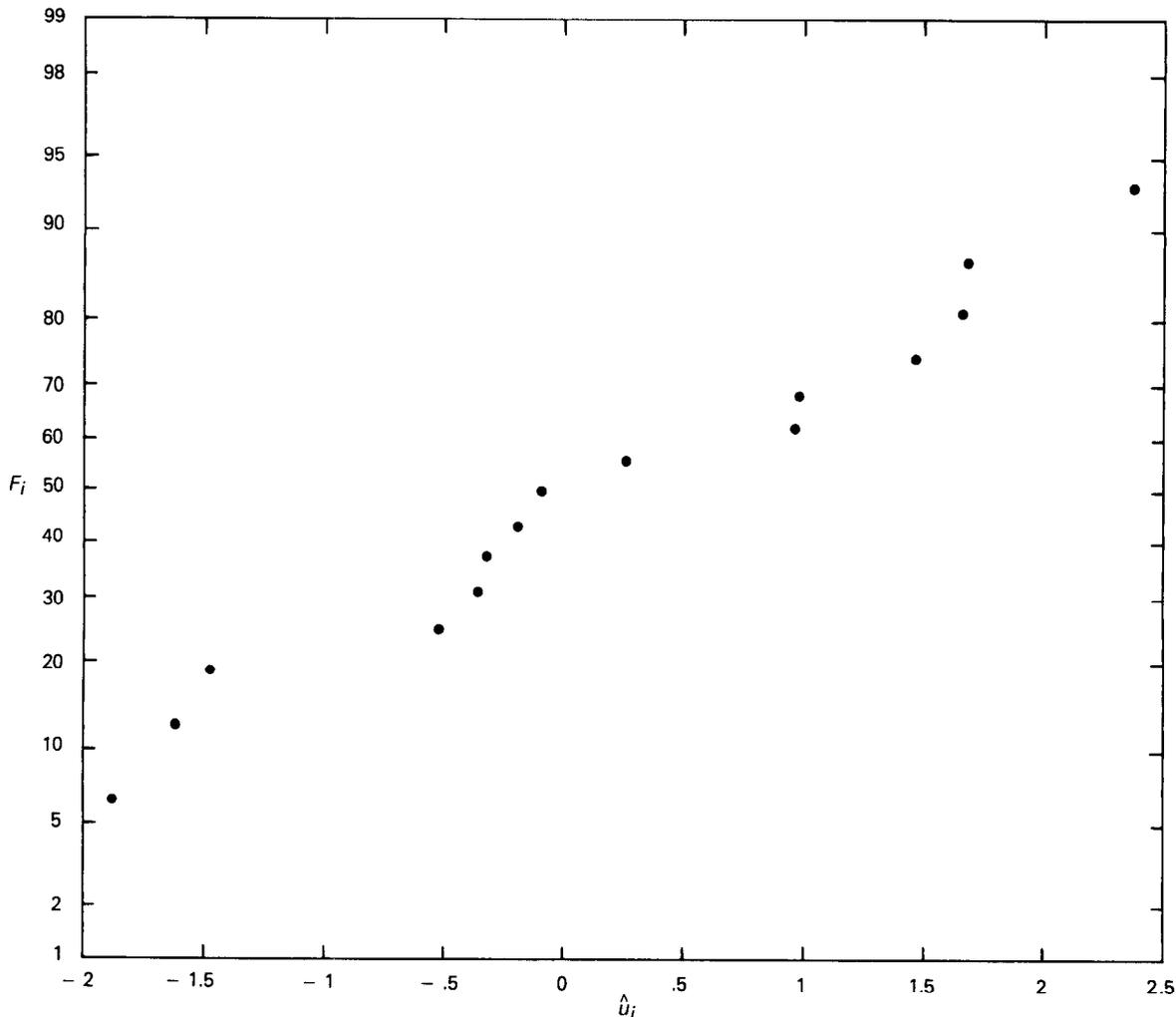

Figure 5.5-1

correlation and unequal variance. Other types of plots involving $\hat{\underline{u}}$ probably would not be affected much by correlation and unequal variance either. Significant departures of the plots for $\underline{g}$ from those for $\underline{d}$ suggest serious correlation and unequal variance effects in $\hat{\underline{u}}$ also.

Whether or not $\hat{\underline{u}}$ departs to a visually detectable extent (which is considered to be synonymous with significant here) from a normal distribution can be determined by comparing plots for the generated set $\underline{g}$ with the plot for $\hat{\underline{u}}$. If the plot for $\hat{\underline{u}}$ has a trend similar to the set of curves for $\underline{g}$, then the distribution for $\hat{\underline{u}}$ probably does not differ enough from equation 5.5-4 to consider abandoning the normality assumption. When examining the plots, it must be remembered that, because of correlation and unequal variance, the plots will not necessarily exhibit the linear trend expected for a univariate normal distribution.

Other residual plots.—In the following discussion it is assumed that the effects of correlation and unequal variance resulting from equation 5.5-3 are negligible so that other effects may be investigated. This assumption might hold true even if a normal probability plot is affected by correlation and unequal variance. However, if one or more patterns (or trends) appear to be present in one or more of the residual plots, then analogous plots using $\underline{g}$ instead of $\hat{\underline{u}}$ also should be prepared and examined. If the suspicious patterns also are present in the plots using $\underline{g}$, then the patterns probably result from correlation and unequal variance inherent in $(\underline{I}-\underline{R})s^2$ and not from model error.

Three types of plot are often useful: (1) Plot of $\hat{u}_j$ vs. $\hat{f}_j$; (2) plots of $\hat{u}_j$ vs. independent variables (ξ_i); (3) plot of $\hat{u}_j$ vs. Cartesian coordinates of point j. If $\underline{\omega}=\underline{I}$ was employed in the regression, then $\hat{u}_j=\hat{e}_j$. Additional discussion of the first two types of plots may be found in Draper and Smith (1981, p. 147-148).

1. Plot of $\hat{u}_j$ vs. $\hat{f}_j$. This type of plot is illustrated in figure 5.5-2. Under the given assumptions, the plot should display a roughly horizontal band of residuals having no apparent trend as sketched in figure 5.5-3. In this and succeeding figures 5.5-4 through 5.5-6, the dashed line outlines the limits of the data. A standard runs test (Draper and Smith, 1981,
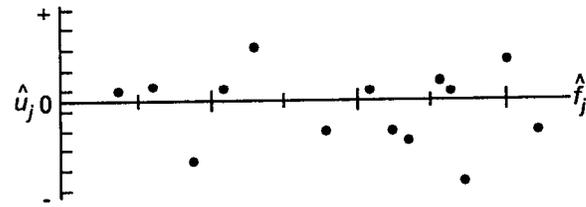
Figure 5.5-2
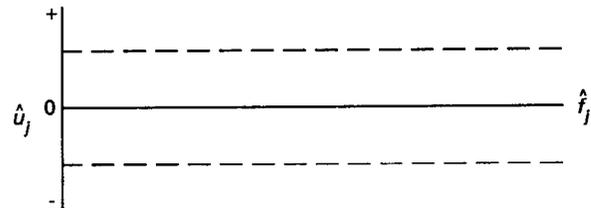
Figure 5.5-3

p. 157-162) could be used to test for randomness of signs of $\hat{u}_j$ along the $\hat{f}_j$ axis.

Three principal types of abnormalities in the plot of $\hat{u}_j$ vs. $\hat{f}_j$ are often apparent:

a. Unequal band width (figure 5.5-4). This type of trend (or one opposite to it) generally indicates that the variance of $\hat{u}_j$ is not constant. In figure 5.5-4 the observations would appear to be less reliable as $\hat{f}_j$ increases.

If the abnormal plot resulted from a least squares analysis where $\underline{\omega}=\underline{I}$ had been assumed, then a diagonal form of $\underline{\omega}\neq\underline{I}$ might be indicated. In the illustration, ω_j should decrease with $\hat{f}_j$. However, if some form of $\underline{\omega}\neq\underline{I}$ had originally been assumed, then an abnormal plot, of the form shown in figure 5.5-4, involving $\hat{u}_j$ would suggest the $\hat{u}_j$ are not of equal reliability and that $\underline{\omega}$ is not correct. Hence, $\underline{\omega}$ should be modified, and the regression performed again. If a full form of $\underline{\omega}$ is required by the true model,
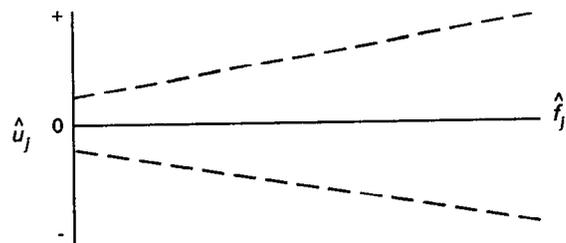
Figure 5.5-4

this form cannot generally be discovered by this type of analysis. Before a weighted least squares is performed, the underlying cause of the variable reliability problem should be investigated so that the diagonal weights can be added according to a rational criterion.

b. Sloping band (figure 5.5-5). This type of problem often indicates model error. Typically, it is caused by omitting an intercept from the model. For a linear model having an intercept, $\Sigma \hat{u}_j \hat{f}_j = 0$ always. Hence in this case an overall slope such as depicted in figure 5.5-5 cannot result from correlation and unequal variance in $(\underline{I} - \underline{R})s^2$. Also, if the model is linear (or effectively so) and $E(\underline{\epsilon}) = \underline{0}$, then $\text{Cov}(\underline{\hat{u}}, \underline{\hat{f}}) = \underline{0}$ even if there is no intercept. Thus, it might be expected that a plot such as shown in figure 5.5-5 would usually not result from correlation and unequal variance if the model is correct and if the observed data are adequate.

c. Curved or irregularly shaped band (figure 5.5-6). This is another indication of model error. The model is inadequate because it does not account for all of the sources of variability in the observed data. More, or different, terms should be added to the regression equation. For the plot shown, it is possible that a quadratic term should be added.

2. Plots of $\hat{u}_j$ vs. independent variables (ξ_i). These plots are interpreted in much the same way as the first type of plot; they simply provide a different viewpoint.

3. For trend surface types of regression: plot of $\hat{u}_j$ vs. Cartesian coordinates of point j (for one or two-dimensional systems). Three-dimensional systems can sometimes be reduced to two dimensions by using cross sections. Systematic highs and lows in the residual pattern usually suggest that the model does not fit the data well. In general, the easier it is to contour the residual map, the more nonrandom the residuals are likely to be.

Problem 5.5-1

To the program for computing T and S using the Theis equation (problem 3.3-1), add the code necessary to compute s^2 using equation 5.4-5 and $\widehat{\text{Var}}(\underline{\hat{b}})$ using equation 5.4-13. You will have to unscale the entries of $(\underline{Z}^T \underline{Z})^{-1}$ to obtain $\widehat{\text{Var}}(\underline{\hat{b}})$. Examination of the calculation
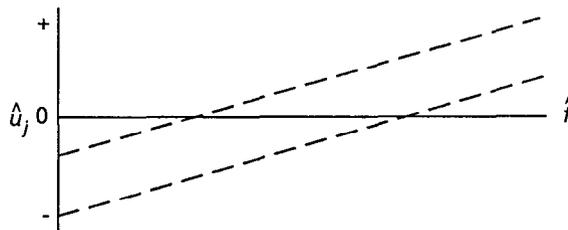
Figure 5.5-5
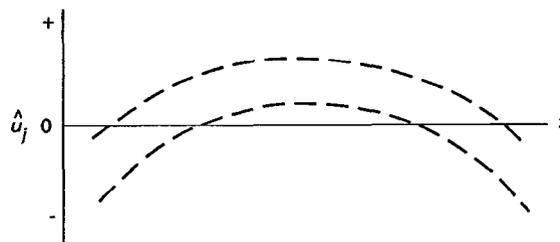
Figure 5.5-6

procedures for obtaining $\underline{Z}^T \underline{Z}$ and $(\underline{Z}^T \underline{Z})^{-1}$ will indicate how the unscaling should be accomplished.

Conduct a graphical analysis of residuals $\underline{\hat{e}}$ resulting from the Theis equation problem (problem 3.3-1). Compute five sets of random normal deviates $\underline{d}$ and simulated residuals $\underline{g}$ using the residuals analysis program (appendix 5.8.1). Plot $\underline{d}$ and $\underline{g}$ on normal probability paper. Does the plot of $\underline{\hat{e}}$ differ significantly from the plots of $\underline{g}$? What can you conclude about the distribution of $\underline{\hat{e}}$? Does the Theis model appear to be adequate?

The code given in appendix 5.8.1 is designed to read COV(I,J) (which is $(\underline{X}^T \underline{\omega} \underline{X})^{-1} s^2$), W(I) (which is $\underline{\omega}_s$) and X(I,J) (which is $\underline{X}_s$) in unformatted form from a file labeled ITB. Normally, these data would be read to file ITB from the numerical nonlinear regression program of appendix 4.3.4. However, the Theis data are not obtained from the numerical nonlinear regression program. The easiest way to read the Theis data is to modify the residuals analysis program by replacing READ(ITB) in statements reading data sets B, C, and D with READ(IIN,2). Data for the program should then be coded as explained in appendix 5.8.1, except that the data for data sets B, C, and D will now be coded in format 8F10.0.

Problem 5.5–2

Use the output from the two-dimensional flow model to analyze the model results of problem 4.2–2.

a. What is the value of s^2? R_y? Compute $s/\Delta Y_s$.

b. Is there any evidence of ill-conditioning in the results? If there is, what is the problem? Which parameters are well-determined and which are not? Why?

c. Conduct a graphical analysis of residuals. Develop four sets of random normal deviates d and simulated residuals g, computed by using the residuals analysis program (appendix 5.8.1). Plot d and g on normal probability paper using equation 5.5–11. Are correlation effects evident? Plot $\hat{u}$ (why $\hat{u}$ instead of $\hat{e}$?) on normal probability paper. Does the plot differ significantly from the plots of g? Plot $\hat{u}_j$ versus $\hat{f}_j$. Is there an abnormal pattern? Plot $\hat{e}_j$ versus Cartesian coordinates of point j (omitting the prior information). Again, is there a pattern to the residuals? What do you conclude about the adequacy of the model?

5.6 Investigation of Alternative Parameter Sets

5.6.1 Generalized W Statistic

Suppose we want to test the null hypothesis that some subset β_2 of parameter set β cannot be distinguished from some corresponding given subset $\tilde{\beta}_2$. That is, test

$$H_0{:}\beta_2=\tilde{\beta}_2 \text{ versus } H_1{:}\beta_2\neq\tilde{\beta}_2.$$

The linearized model assumed is

$$\underline{Y}=f_0+\underline{X}(\beta-\underline{b}_0)+\underline{\epsilon}$$

$$=f_0+\underline{X}_1(\beta_1-\underline{b}_{01})+\underline{X}_2(\beta_2-\underline{b}_{02})+\underline{\epsilon} \quad (5.6–1)$$

where $\underline{X}$, β, and $\underline{b}_0$ are conformably partitioned as follows:

$$\underline{X}=[\underline{X}_1,\underline{X}_2] \qquad (5.6–2)$$

$$\beta=\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$\underline{b}_0=\begin{bmatrix} \underline{b}_{01} \\ \underline{b}_{02} \end{bmatrix} \qquad \left.\vphantom{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}\right\} \quad (5.6–3)$$

Based on equation 5.6–1, we may state a predictive model of the form

$$\tilde{f}=f_0+\underline{X}_1(\tilde{b}_1-\underline{b}_{01})+\underline{X}_2(\tilde{\beta}_2-\underline{b}_{02}) \qquad (5.6–4)$$

where $\tilde{b}_1$ is an estimate of β_1 and, under the null hypothesis, β_2 is assumed to be given by $\tilde{\beta}_2$.

The W statistic, which is stated explicitly later on, is based on a comparison of the restricted sum of squares $(\underline{Y}-\tilde{f})^T\underline{\omega}(\underline{Y}-\tilde{f})$ and the unrestricted sum of squares $(\underline{Y}-\hat{f})^T\underline{\omega}(\underline{Y}-\hat{f})$. The unrestricted sum of squares is obtained from the standard least squares analysis. The restricted sum of squares is obtained by minimizing

$$S_0(\tilde{b}_1)=(\underline{Y}-\tilde{f})^T\underline{\omega}(\underline{Y}-\tilde{f}) \qquad (5.6–5)$$

with respect to $\tilde{b}_1$ while holding $\tilde{\beta}_2$ constant. For a linear model this results in the normal equations:

$$\tilde{b}_1-\underline{b}_{01}=(\underline{X}_1^T\underline{\omega}\underline{X}_1)^{-1}\underline{X}_1^T\underline{\omega}(\underline{Y}-f_0-\underline{X}_2(\tilde{\beta}_2-\underline{b}_{02})) \,(5.6–6)$$

where $\underline{b}_{01}$ and $\underline{b}_{02}$ can be set to zero if desired. If the model is nonlinear, then equation 5.6–5 is minimized with respect to $\tilde{b}_1$ by using the standard procedures, except that $\tilde{\beta}_2$ is held constant.

If many repeat experiments using linear normal equation 5.6–6 were performed, and if assumptions given by equations 5.2–5, 5.2–6, and 5.2–10 held true, then it would be found that

$$W=\frac{[(\underline{Y}-\tilde{f})^T\underline{\omega}(\underline{Y}-\tilde{f})-(\underline{Y}-\hat{f})^T\underline{\omega}(\underline{Y}-\hat{f})]/q}{(\underline{Y}-\hat{f})^T\underline{\omega}(\underline{Y}-\hat{f})/(n-p)}$$

$$=\frac{[(\underline{Y}-\tilde{f})^T\underline{\omega}(\underline{Y}-\tilde{f})-(\underline{Y}-\hat{f})^T\underline{\omega}(\underline{Y}-\hat{f})]/q}{s^2}$$

$$\sim F(q,n-p) \qquad (5.6–7)$$

where q is the order of $\tilde{\underline{\beta}}_2$, which is the number of restrictions in H_0, and equation 5.4-5 was used. The symbol W stands for a random variable. For each experiment, a value w of the random variable W could be computed. According to equation 5.6-7, by repeating the experiment many times, the probability of the ratio W having a value of w or a smaller value would be found to be given by the cumulative density function F.

Because equation 5.6-7 is proportional to the difference between the restricted and unrestricted sums of squares divided by the unrestricted sum of squares, one might suspect the null hypothesis $H_0{:}\underline{\beta}_2=\tilde{\underline{\beta}}_2$ to be true if w is small. However, if w is large, then one might suspect that H_0 is incorrect. The rejection region for the hypothesis test is determined by the probability statement $P(W{>}F_\alpha(q,n{-}p)){=}\alpha$, where α is the significance level of the test and $F_\alpha(q,n{-}p)$ is the upper $100\alpha\%$ point of the F distribution with q and $n{-}p$ degrees of freedom. If the ratio w is greater than $F_\alpha(q,n{-}p)$, as found in any table of critical values for the F distribution, then the null hypothesis is rejected because values of w greater than $F_\alpha(q,n{-}p)$ form the rejection region.

An alternative form for the numerator of equation 5.6-7 may be derived by manipulating the linearized models. The result, after extensive algebra, is

$$(\underline{Y}{-}\tilde{\underline{f}})^T\underline{\omega}(\underline{Y}{-}\tilde{\underline{f}}){-}(\underline{Y}{-}\hat{\underline{f}})^T\underline{\omega}(\underline{Y}{-}\hat{\underline{f}})$$

$$=(\tilde{\underline{\beta}}_2{-}\hat{\underline{b}}_2)^T[\underline{H}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{H}^T]^{-1}(\tilde{\underline{\beta}}_2{-}\hat{\underline{b}}_2) \quad (5.6\text{-}8)$$

where

$$\underline{H}=[\underline{0},\ \underline{I}_q]_{(q\times p)}, \quad\quad\quad\quad (5.6\text{-}9)$$

$\underline{I}_q=$identity matrix of order q, and

$\hat{\underline{b}}_2=$the partition corresponding to $\underline{\beta}_2$ found from the standard (unrestricted) least squares analysis. $\quad\quad (5.6\text{-}10)$

Thus, the alternative form for equation 5.6-7 is

$$W=\frac{(\tilde{\underline{\beta}}_2{-}\hat{\underline{b}}_2)^T[\underline{H}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{H}^T]^{-1}(\tilde{\underline{\beta}}_2{-}\hat{\underline{b}}_2)/q}{s^2}.$$
$$(5.6\text{-}11)$$

If the model is linear, either equation 5.6-7 or equation 5.6-11 may be used to compute w. Both equations require an unrestricted regression to obtain $\hat{\underline{b}}$ and $\hat{f}$, but equation 5.6-7 requires, in addition, a restricted regression using equation 5.6-6 to obtain $\tilde{\underline{b}}_1$ and $\tilde{f}$. Hence, for a linear model, equation 5.6-11 is often more efficient to use than equation 5.6-7 for practical computations. If equation 5.6-1 is a linearized equation system, derived from a nonlinear model, then neither equation 5.6-7 nor equations 5.6-8 nor 5.6-11 is exact. However, if equation 5.6-1 behaves in a way that is close enough to being linear, then equations 5.6-7, 5.6-8, and 5.6-11 are good approximations. When working with a nonlinear model, w should be computed using both equations 5.6-7 and 5.6-11. If the conclusions reached by using the two different expressions for w are different, then the model may be too nonlinear for investigations using the W statistic. Further investigation of model nonlinearity may be performed by employing the modified Beale's measure, which is discussed in section 6.2.

In summary, the procedure for testing

$$H_0{:}\underline{\beta}_2=\tilde{\underline{\beta}}_2 \text{ vs. } H_1{:}\underline{\beta}_2{\neq}\tilde{\underline{\beta}}_2$$

in the model $\underline{Y}{=}\underline{f}_0{+}\underline{X}(\beta{-}\underline{b}_0){+}\underline{\epsilon}$ is:
1. Carry out a regression without any restriction to find $\hat{\underline{b}}$ for the full model.
2. When using equation 5.6-7, fix $\tilde{\underline{\beta}}_2$ and use equation 5.6-6 to find $\tilde{\underline{b}}_1$ for the restricted model. When using equation 5.6-11, skip this step.
3. Form the ratio w using equation 5.6-7 or equation 5.6-11.
4. Compare w with the appropriate value of $F_\alpha(q,n{-}p)$.

5.6.2   Joint Confidence Region for $\underline{\beta}_2$

Equations 5.6-7 or 5.6-11 also may be used to obtain a <u>joint confidence region</u> on $\underline{\beta}_2$. The confidence region interpretation is based on fixing a probability of occurrence, $P(W{<}F_\alpha(q,n{-}p))$ $={1}{-}\alpha$, then finding those vectors $\tilde{\underline{\beta}}_2$ that would yield the specified F or a smaller value. The joint confidence region may be written as

$$(\underline{Y}{-}\tilde{\underline{f}})^T\underline{\omega}(\underline{Y}{-}\tilde{\underline{f}}){-}(\underline{Y}{-}\hat{\underline{f}})^T\underline{\omega}(\underline{Y}{-}\hat{\underline{f}}){\leq}s^2qF_\alpha(q,n{-}p)(5.6\text{-}12)$$

or

$$(\underline{\beta}_2-\hat{\underline{b}}_2)^T[\underline{H}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{H}^T]^{-1}(\underline{\beta}_2-\hat{\underline{b}}_2)$$

$$\leq s^2qF_\alpha(q,n-p) . \qquad (5.6\text{-}13)$$

Equation 5.6-13 plots as a family of q dimensional ellipsoids (which for $q=1$ is a line segment bounded by two values for $\underline{\beta}_2$) in parameter space, and these ellipsoids are centered on $\hat{\underline{b}}$. All ellipsoids corresponding to probability levels smaller than $1-\alpha$ lie within the outermost ellipsoid, which is defined by strict equality in equation 5.6-13. Hence, the specified probability is the probability that $\underline{\beta}_2$ lies within the ellipsoid. An equivalent statement is that, if many experiments were conducted, then $(1-\alpha)100\%$ of the ellipsoids would contain the true parameter set $\underline{\beta}_2$. Hence, the outermost ellipsoid may be considered to be a joint confidence region on $\underline{\beta}_2$.

Points (in parameter space) on the edge of the confidence region corresponding to the maximum and minimum (or extreme) values that some parameter $\tilde{\beta}_{2i}$ may attain and remain in the confidence region are given by

$$\tilde{\underline{b}}=\hat{\underline{b}}\pm\frac{\sqrt{qF_\alpha(q,n-p)}}{s_{bi}}\underline{V}_{bi} \qquad (5.6\text{-}14)$$

where $\tilde{\underline{b}}^T=[\tilde{\underline{b}}_1^T,\tilde{\underline{\beta}}_2^T]$ and $\underline{V}_{bi}$ is the ith column of $\underline{V}_b=(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2$. Note that s_{bi} is the square root of v_{bii}.

Equation i in 5.6-14 gives the extreme values of $\tilde{\beta}_{2i}$. The parameter vector $\tilde{\underline{b}}$ computed using equation 5.6-14 is the same vector that would result if (1) $\tilde{\beta}_{2i}$ were computed using equation i in 5.6-14, (2) then the remaining values in $\tilde{\underline{\beta}}_2$ were computed to satisfy equation 5.6-13 (with strict equality applying to give points on the edge of the confidence region), and (3) finally, partition $\tilde{\underline{b}}_1$ were computed using equation 5.6-6.

Parameter sets computed using equation 5.6-14 will exactly satisfy the relationship $s^2qF_\alpha(q,n-p)=(\underline{Y}-\tilde{\underline{f}})^T\underline{\omega}(\underline{Y}-\tilde{\underline{f}})-(\underline{Y}-\hat{\underline{f}})^T\underline{\omega}(\underline{Y}-\hat{\underline{f}})=(\tilde{\underline{\beta}}_2-\hat{\underline{b}}_2)^T[\underline{H}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{H}^T]^{-1}(\tilde{\underline{\beta}}_2-\hat{\underline{b}}_2)$ if the model is linear. However, as discussed in section 5.6.1, this relationship is not exact for a nonlinear model. Therefore, if the regression model is nonlinear, the parameter sets computed using equation 5.6-14 should be substituted into the

nonlinear model and $(\underline{Y}-\tilde{\underline{f}})^T\underline{\omega}(\underline{Y}-\tilde{\underline{f}})-(\underline{Y}-\hat{\underline{f}})^T\underline{\omega}(\underline{Y}-\hat{\underline{f}})$ should be computed. If this value is different enough from $s^2qF_\alpha(q,n-p)$ to change any conclusions, then the model is too nonlinear to use to generate linearized confidence regions. The modified Beale's measure discussed in section 6.2 also can be used to gauge nonlinearity.

Two end-member cases involving the W statistic are often considered separately. In one case $q=p$ so that $\underline{\beta}_2=\underline{\beta}$. All parameters are thus included in any test of H_0, and the confidence region is on all parameters simultaneously. This confidence region is called a joint confidence region on all parameters. An example for two parameters is diagrammed in figure 5.6-1. In the other case, $q=1$ and $\underline{\beta}_2=\beta_p$. Thus, only one parameter is considered in any test of H_0, and the confidence region is on only one parameter. The confidence region for this case is termed an individual confidence interval on parameter β_p.

Problem 5.6-1

a. A method for estimating recharge in Nevada is known as the Maxey-Eakin method. Using this method, recharge rate W in the vicinity of Lake Ohpupu (problem 3.2-1) was estimated to be 0.0003 ft/day. Also, by using specific capacity estimates from well-log analysis, T was found to be 10 ft²/day. Using these estimates, test the null hypothesis that there is no significant difference at $\alpha=0.05$ between W/T as estimated above
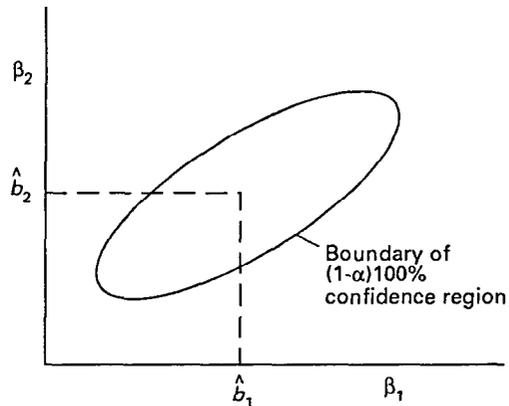

Boundary of $(1-\alpha)100\%$ confidence region

Figure 5.6-1

and the regression estimate of W/T. (Hint: Use equation 5.6–11.) Based on the result of the hypothesis test, would you consider using the prior estimate of W/T as prior information in the regression model? What other information would you need if you did use it?

b. In addition to asking whether or not regression estimate W/T is significantly different from another, independent, estimate, one might ask whether or not W/T is even a significant variable in the regression model. Set up and conduct a test to answer this question, then interpret the result.

c. Using equation 5.6–14, find the bounds of the confidence interval on W/T for $\alpha = 0.05$. (Hint: You need only consider the equation corresponding to parameter W/T in the system implied by equation 5.6–14.)

Problem 5.6–2

Use your Theis equation program (problems 3.3–1 and 5.5–1) and equation 5.6–14 to find the sets of parameters corresponding to extreme values of T and to extreme values of S, assuming $q = 2$.

Problem 5.6–3

Using equation 5.6–14 and the results of problem 4.2–2, find the sets of parameters corresponding to extreme values of T_3 and to extreme values of q_{B1} with $q = 2$.

5.7 Investigation of Predictive Reliability

5.7.1 The Variance-Covariance Matrix for $\hat{\underline{f}}$

Equation 5.1–8 is used to obtain

$$\mathrm{Var}(\hat{\underline{f}}) = \mathrm{Var}[\underline{\underline{X}}(\hat{\underline{b}} - \underline{b}_0) + \underline{f}_0]$$

$$= \underline{\underline{X}}\mathrm{Var}(\hat{\underline{b}})\underline{\underline{X}}^T \ . \qquad (5.7\text{–}1)$$

An analogous measure for weighted values of $\hat{\underline{f}}$ is

$$\mathrm{Var}(\underline{\omega}^{1/2}\hat{\underline{f}}) = \mathrm{Var}[\underline{\omega}^{1/2}\underline{\underline{X}}(\hat{\underline{b}} - \underline{b}_0) + \underline{\omega}^{1/2}\underline{f}_0]$$

$$= \underline{\omega}^{1/2}\underline{\underline{X}}\mathrm{Var}(\hat{\underline{b}})\underline{\underline{X}}^T\underline{\omega}^{1/2} \ . \qquad (5.7\text{–}2)$$

By using equations 5.3–6 and 5.4–12, equation 5.7–2 can be written

$$\mathrm{Var}(\underline{\omega}^{1/2}\hat{\underline{f}}) = \underline{\underline{R}}\sigma^2 \ . \qquad (5.7\text{–}3)$$

Estimates corresponding to equations 5.7–1, 5.7–2, and 5.7–3 are

$$\widehat{\mathrm{Var}(\hat{\underline{f}})} = \underline{\underline{X}}\widehat{\mathrm{Var}(\hat{\underline{b}})}\underline{\underline{X}}^T \qquad (5.7\text{–}4)$$

$$\widehat{\mathrm{Var}(\underline{\omega}^{1/2}\hat{\underline{f}})} = \underline{\omega}^{1/2}\underline{\underline{X}}\widehat{\mathrm{Var}(\hat{\underline{b}})}\underline{\underline{X}}^T\underline{\omega}^{1/2} \qquad (5.7\text{–}5)$$

$$= \underline{\underline{R}}s^2 \ . \qquad (5.7\text{–}6)$$

The standard error of $\hat{f}_i$ is given by the square root of the ith diagonal entry of $\widehat{\mathrm{Var}(\hat{\underline{f}})}$. This estimate gives a measure of potential variability in $\hat{f}_i$ resulting from $\widehat{\mathrm{Var}(\hat{\underline{b}})}$.

It is important to note that equations 5.7–1, 5.7–2, 5.7–4, and 5.7–5 are valid for prediction vectors $\hat{\underline{f}}$ having entries that are not necessarily at observation points. This fact may be understood by observing that prediction equation 5.1–8 used to derive $\mathrm{Var}(\hat{\underline{f}})$ or $\mathrm{Var}(\underline{\omega}^{1/2}\hat{\underline{f}})$ (or their estimates) is valid for any set of points, not just observation points. However, entries in $\underline{\underline{X}}$ and, if either equation 5.7–2 or 5.7–5 is used, ω must be available for all points in $\hat{\underline{f}}$. Matrix $\widehat{(\underline{\underline{X}}^T\underline{\omega}\underline{\underline{X}})^{-1}}s^2$ used for $\widehat{\mathrm{Var}(\hat{\underline{b}})}$ is, of course, the standard one based on entries in $\underline{\underline{X}}$ only at the observation points.

5.7.2 Confidence Interval for $f_{\beta j}$

If all parameters are allowed to vary over the confidence region given by equation 5.6–12 or equation 5.6–13 with $q = p$, then the maximum and minimum values produced for $\tilde{f}_j$ form a corresponding confidence interval for $\tilde{f}_j$. In this case, because $q = p$, $\underline{\beta}_2 = \underline{\beta}$ and $\tilde{f}_j = f_{\beta j}$. The resulting confidence interval for $f_{\beta j}$ is

$$f_{\beta j} = \hat{f}_j \pm \sqrt{pF_\alpha(p, n-p)}\, s_{yj} \qquad (5.7\text{–}7)$$

where

$$s_{yj} = s\sqrt{\underline{X}_j(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}_j^T} \qquad (5.7\text{-}8)$$

and $\underline{X}_j$ is a row of a sensitivity matrix corresponding to the point j. The point j need not correspond to an observation point.

Equation 5.7-8 gives a simultaneous confidence interval. That is, the probability is $1-\alpha$ that $f_{\beta j}$ lies within the interval indicated by equation 5.7-7, and that f_β for all other possible points lies within similar intervals simultaneously. If intervals on a number of consecutive points are computed and plotted, the result is called a <u>confidence band</u>.

5.7.3 Prediction Interval for Predicted Observation Y_j^{pred}

Equation 5.7-7 gives a confidence interval on a computed value f or, in other words, the mean of Y, which is a fixed, nonrandom quantity. In some instances, a corresponding interval on a predicted observation, which is a random quantity, is desired, and this interval is termed a <u>prediction interval</u>. Prediction intervals on k predicted values of Y simultaneously can be readily computed if $\underline{\underline{\omega}}=\underline{\underline{V}}^{-1}$ is diagonal, and they are given by Lieberman (1961):

$$Y_j^{pred} = \hat{f}_j \pm \sqrt{kF_\alpha(k, n-p)}\sqrt{s^2/\omega_j + s_{yj}^2},$$
$$j = 1, 2, \dots, k \qquad (5.7\text{-}9)$$

As in equation 5.7-7, point j need not (and, in general, probably would not) correspond to an observation point. However, ω_j for the prediction point has to be known.

The term $s^2/\omega_j + s_{yj}^2$ is the total variance in predicted observation Y_j^{pred}. This form results because $Y_j^{pred} = (Y_j^{pred} - \hat{f}_j) + \hat{f}_j$, where $Y_j^{pred} - \hat{f}_j$ is statistically independent of $\hat{f}_j$, so that, as an estimate, $\text{Var}(Y_j^{pred}) = \text{Var}(Y_j^{pred} - \hat{f}_j) + \text{Var}(\hat{f}_j)$ or $\text{Var}(Y_j^{pred}) = s^2/\omega_j + s_{yj}^2$.

Equation 5.7-9 does not give prediction intervals on all $\underline{Y}_{pred}$ simultaneously. Furthermore, as k increases, the prediction interval increases without bound. This result is because the normal distribution, which the errors in Y are assumed to follow, has infinite tails. Hence, even though the probability of an error that is large in magnitude is small, as the number of

values of Y considered simultaneously increases, the probability of an arbitrarily large error in at least one of them increases also. Usually the prediction interval is computed using $k=1$.

Problem 5.7-1

a. Write out explicitly the form for general entry (ij) of $\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T$, which is used in equation 5.7-1. For this exercise let $(\underline{X}^T\underline{\omega}\underline{X})^{-1}=\underline{A}$ and write the result in terms of $\underline{A}$ to simplify the expression. Select and compute a diagonal entry of this matrix at an observation point used for the linear regression solution of problem 3.2-1. Note that you can replace $\underline{X}$ by $\underline{S}$ in $\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T$, and that this replacement leaves the result unaltered. Can you show this?

b. Using equation 5.7-4 and the diagonal entry computed in part a, determine $\text{Var}(\hat{f}_j)$, where j is the selected diagonal entry.

c. Using equation 5.7-7 and the results of b, find the confidence interval on $f_{\beta j}$.

5.8 Appendix

5.8.1 Documentation of Program to Compute Vectors $\underline{d}$ and $\underline{g}$ of Section 5.5.2.

This program computes vector $\underline{d}$ of random normal deviates, vector $\underline{g}$ of correlated normal deviates as defined by equation 5.5-10, and other useful information related to the distributions of $\hat{\underline{e}}$ and $\hat{\underline{u}}$. Sample and direct prior information are assumed to be given in the form of equation 3.4-12. The sensitivity matrix for the direct prior information is theoretically of the form $\underline{X}_p = [\underline{I}_p, \underline{0}]$. However, matrix $[\underline{I}_p, \underline{0}]$ may be rearranged to conform with any parameter ordering.

The program was developed using the Microsoft Fortran Compiler, Version 3.3, with the DOS 2.0 operating system on an IBM PC/XT computer with the IBM 8088 Math Coprocessor and 256 KB memory. Except for the OPEN statements near the beginning of the

code, Fortran 66 was used throughout to make the code as machine independent as possible. The source code is contained in file RESAN.FOR in the diskette accompanying this report. A random (0,1) number generator is employed as a function subroutine. This routine assumes an integer computer word length equal to at least 1,077,109,141.

As coded, the contents of data sets B, C, and D are assumed to be stored in unformatted form on file ITB=8. This is so that these sets do not have to be input manually. The listing appended contains the code (file RESINS.FOR on the diskette) and instructions for insertion into the program of appendix 4.3.4 so that the required data are stored in the proper form and order for use in the present program. The user will have to supply the job control language necessary to store the data and retrieve them for use.

Two variables, NVD and NTD defined near the beginning of the program, must be redefined each time the dimensions of the program are changed. NVD must be set equal to dimensions of COV and the first dimension of S, all three of which are at least NVAR, and NTD must be set equal to the dimensions of R, which are at least NTOT=NOBS+NPRIR.

Input Data.—Data Set A.
Problem size information; one line (format 5I5, F10.0).

Line columns	Variable	Definition
1-5	NVAR	Number of parameters, p.
6-10	NOBS	Number of sample observations, n_s.
11-15	NPRIR	Number of regression parameters having direct prior information, n_p.
16-20	NSETS	Number of sets of $\underline{d}$ and g vectors to be computed.
21-25	NRAN	Seed for random number generator: any odd number between 1 and 1,048,575.
26-35	VAR	Error variance, s^2.

Data Set B.
Covariance matrix, $(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2$ (unformatted; stored in file ITB).

Variable	Definition
COV(1,1) COV(2,1) $\vdots$ COV(NVAR,1) COV(2,2) $\vdots$ COV(NVAR,2) $\vdots$ COV(NVAR,NVAR)	Covariance matrix, entered sequentially from the diagonal element through NVAR for each new regression parameter number. Each new diagonal element begins a new record.

Data Set C.
Weight matrix for sample information, $\underline{V}_s^{-1}$ (unformatted; stored in file ITB).

Variable	Definition
W(1) W(2) $\vdots$ W(NOBS)	Diagonal weight matrix for sample information, entered sequentially from 1 through NOBS.

Data Set D.
Sensitivity matrix for sample information, $\underline{X}_s$ (unformatted; stored in file ITB).

Variable	Definition
X(1,1) X(2,1) $\vdots$ X(NVAR,1) X(1,2) $\vdots$ X(NVAR,2) $\vdots$ X(NVAR,NOBS)	Sensitivity matrix for sample information, entered sequentially 1 through NVAR for each observation. Each new observation begins a new record, for a total of NOBS observations.

Data Set E.
Estimated error variance used with prior information of known reliability; one line (format F10.0).

Line columns	Variable	Definition
1-10	EV	The initial estimate of s^2 used in conjunction with prior information.

Omit data set if NPRIR=0.

Data Set F.
Parameter numbers having prior information
(format 16I5).

Line columns	Variable	Definition
1-5	IPR(1)	Array subscript numbers
6-10	IPR(2)	for regression parameters
⋮	⋮	in $\hat{b}$ having prior infor-
	IPR(NPRIR)	mation. For use with
		the regression ground-
		water program, the
		array subscript numbers
		must be the subscript
		numbers in the parameter
		vector computed by that
		program.

Omit data set if NPRIR=0.

Data Set G.
Standard deviation matrix for prior informa-
tion $\underline{U}^{1/2}$ (format 8F10.0).

Line columns	Variable	Definition
1-10	F(1)	Diagonal standard devia-
11-20	F(2)	tion matrix for prior in-
⋮	⋮	formation, entered in the
	F(NPRIR)	same order as IPR(I)
		from 1 through NPRIR.

Omit data set if NPRIR=0.

Output.—Output is all clearly labeled. It is
ordered as follows:

1. Data sets A through G.
2. Set number of vectors $\underline{d}$ and g. Data for
   numbers 2 through 5 below are printed
   sequentially for each set.
3. Vector $\underline{d}$. This vector is ordered from
   smallest to largest entry, and each en-
   try is paired with its theoretical frequen-
   cy as computed by using equation
   5.5-11.
4. Vector g. Each entry is printed in its
   natural position corresponding to its
   position in a row or column of $\underline{R}$. Rows
   and columns of $\underline{R}$ are ordered by first
   sample observation numbers followed
   by prior information numbers, which are
   the subscripts I of IPR(I).
5. Vector g. This vector is ordered and
   paired with its theoretical frequency in
   the same way as $\underline{d}$ is.
6. Covariance matrix $(\underline{I}-\underline{R})s^2$. This matrix,
   which is an estimate of the one defined
   by equation 5.5-3, is for weighted resid-
   uals $\hat{\underline{u}}$ composed of both sample and
   prior information, with the prior infor-
   mation occupying the last n_p rows and
   columns.
7. Correlation matrix for $\hat{\underline{u}}$. This matrix is
   derived from $(\underline{I}-\underline{R})s^2$.

Program Listing

```
C          RESIDUALS ANALYSIS PROGRAM BY R. L. COOLEY, USGS, DENVER, COLO.
$LARGE: R
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        DIMENSION X(20,70),COV(20,20),W(70),WP(20),IPR(20),R(90,90),D(90)
       1,G(90),F(90)
        COMMON/ITP/IIN,IOUT
        COMMON/FLT/X,R
        EQUIVALENCE (X(1,1),D(1)),(W(1),F(1),WP(1)),(COV(1,1),G(1))
        OPEN (5,FILE='RESAN.DAT',STATUS='OLD',ACCESS='SEQUENTIAL'
       1,FORM='FORMATTED')
        OPEN (6,FILE='RESAN.OUT',STATUS='NEW',ACCESS='SEQUENTIAL'
       1,FORM='FORMATTED')
        OPEN (8,FILE='RESAN.IN',STATUS='OLD',ACCESS='SEQUENTIAL'
       1,FORM='UNFORMATTED')
C**FORMAT LIST
      1 FORMAT (5I5,F10.0)
      2 FORMAT (8F10.0)
      3 FORMAT (9H1NVAR  = ,I4/9H NOBS  = ,I4/9H NPRIR = ,I4
       1/9H NSETS = ,I4/9H NRAN  = ,I4/9H VAR   = ,G11.5)
      4 FORMAT (1H0,14X,42HRELIABILITY WEIGHTS FOR SAMPLE INFORMATION
       1/1H ,3X,3(3HNO.,11X,1HW,9X))
      5 FORMAT (16I5)
      6 FORMAT (1H0,14X,43HNO.S OF PARAMETERS HAVING PRIOR INFORMATION
       1/1H ,3X,3(3HNO.,8X,3HIPR,10X))
      7 FORMAT (1H0,14X,40HSTANDARD DEVIATIONS OF PRIOR INFORMATION
       1/1H ,3X,3(3HNO.,10X,2HWP,9X))
      8 FORMAT (19H0 COVARIANCE MATRIX)
      9 FORMAT (38H0 SENSITIVITIES FOR OPTIMUM PARAMETERS)
     10 FORMAT (6H0EV = ,G11.5)
     11 FORMAT (1H0,20X,31HORDERED, RANDOM NORMAL DEVIATES/1H ,3X,2(3HNO.
       1,8X,1HD,14X,1HF,10X))
     12 FORMAT (1H0,18X,35HORDERED, CORRELATED NORMAL DEVIATES/1H ,3X
       1,2(3HNO.,8X,1HG,14X,1HF,10X))
     13 FORMAT (33H0 COVARIANCE MATRIX FOR RESIDUALS)
     14 FORMAT (43H0DATA GENERATED FROM RANDOM NUMBER SET NO. ,I3)
     15 FORMAT (34H0 CORRELATION MATRIX FOR RESIDUALS)
     16 FORMAT (1H0,22X,26HCORRELATED NORMAL DEVIATES/1H ,3X,3(3HNO.,11X
       1,1HG,9X))
C**DEFINE INPUT FILES, OUTPUT FILE, AND ARRAY DIMENSIONS FOR PRTOT
        IIN=5
        ITB=8
        IOUT=6
        NVD=20
        NTD=90
C**READ AND PRINT INPUT DATA THEN CONVERT IT INTO FORMS NEEDED
C  FOR CALCULATIONS
C**NOTE: NRAN MUST BE ODD AND MUST LIE BETWEEN 1 AND 1048575
        READ(IIN,1) NVAR,NOBS,NPRIR,NSETS,NRAN,VAR              SET A
        WRITE(IOUT,3) NVAR,NOBS,NPRIR,NSETS,NRAN,VAR
        DO 25 J=1,NVAR
        READ(ITB) (COV(I,J),I=J,NVAR)                          SET B
        DO 20 I=J,NVAR
```

Program Listing—Continued

```
   20 COV(J,I)=COV(I,J)
   25 CONTINUE
      WRITE(IOUT,8)
      CALL PRTOT(COV,NVAR,NVAR,NVD)
      READ(ITB) (W(I),I=1,NOBS)                          SET C
      WRITE(IOUT,4)
      CALL PRTOT(W,NOBS,1,0)
      DO 35 J=1,NOBS
      READ(ITB) (X(I,J),I=1,NVAR)                        SET D
   35 CONTINUE
      WRITE(IOUT,9)
      CALL PRTOT(X,NVAR,NOBS,NVD)
      DO 45 J=1,NOBS
      WT=W(J)**.5
      DO 40 I=1,NVAR
   40 X(I,J)=X(I,J)*WT
   45 CONTINUE
      IF(NPRIR.LT.1) GO TO 55
      READ(IIN,2) EV                                     SET E
      WRITE(IOUT,10) EV
      READ(IIN,5) (IPR(I),I=1,NPRIR)                     SET F
      WRITE(IOUT,6)
      CALL PRTOTC(IPR,NPRIR)
      READ(IIN,2) (WP(I),I=1,NPRIR)                      SET G
      WRITE(IOUT,7)
      CALL PRTOT(WP,NPRIR,1,0)
      SIGMA=EV**.5
      DO 50 I=1,NPRIR
   50 WP(I)=SIGMA/WP(I)
C**COMPUTE (I-R)*VAR MATRIX
   55 DO 80 K=1,NOBS
      DO 70 J=1,NVAR
      SUM=0.
      DO 60 I=1,NVAR
   60 SUM=SUM+X(I,K)*COV(I,J)
   70 R(J,K)=SUM
   80 CONTINUE
      IF(NPRIR.LT.1) GO TO 90
      DO 84 K=1,NOBS
      DO 82 I=1,NPRIR
      J=IPR(I)
   82 R(I+NOBS,K)=-WP(I)*R(J,K)
   84 CONTINUE
      DO 88 J=1,NPRIR
      L=IPR(J)
      DO 86 I=J,NPRIR
      K=IPR(I)
   86 R(I+NOBS,J+NOBS)=-WP(I)*COV(K,L)*WP(J)
   88 CONTINUE
   90 DO 110 K=1,NOBS
      DO 100 J=K,NOBS
      SUM=0.
      DO 95 I=1,NVAR
```

Program Listing—Continued

```
   95 SUM=SUM+X(I,K)*R(I,J)
  100 R(J,K)=-SUM
  110 CONTINUE
      NTOT=NOBS+NPRIR
      DO 130 J=1,NTOT
      DO 120 I=J,NTOT
  120 R(J,I)=R(I,J)
  130 R(J,J)=VAR+R(J,J)
C**COMPUTE THEORETICAL FREQUENCIES FOR DATA SETS
      TMP=NTOT+1
      DO 135 I=1,NTOT
      TEMP=I
  135 F(I)=TEMP/TMP
      SIGMA=VAR**.5
      DO 180 K=1,NSETS
      WRITE(IOUT,14) K
C**COMPUTE RANDOM NORMAL DEVIATES D AND CORRELATED NORMAL DEVIATES G
      DO 150 I=1,NTOT
      SUM=-6.
      DO 140 J=1,12
  140 SUM=SUM+RANUM(NRAN)
  150 D(I)=SIGMA*SUM
      DO 170 J=1,NTOT
      SUM=0.
      DO 160 I=1,NTOT
  160 SUM=SUM+R(I,J)*D(I)
  170 G(J)=SUM/VAR
C**ORDER AND PRINT RANDOM NORMAL DEVIATES AND CORRELATED NORMAL DEVIATES
      DO 174 I=1,NTOT
      DO 172 J=I,NTOT
      IF(D(J).GE.D(I)) GO TO 172
      TMP=D(I)
      D(I)=D(J)
      D(J)=TMP
  172 CONTINUE
  174 CONTINUE
      WRITE(IOUT,11)
      CALL PRTOTA(D,F,NTOT)
      WRITE(IOUT,16)
      CALL PRTOT(G,NTOT,1,0)
      DO 178 I=1,NTOT
      DO 176 J=I,NTOT
      IF(G(J).GE.G(I)) GO TO 176
      TMP=G(I)
      G(I)=G(J)
      G(J)=TMP
  176 CONTINUE
  178 CONTINUE
      WRITE(IOUT,12)
      CALL PRTOTA(G,F,NTOT)
  180 CONTINUE
C**PRINT COVARIANCE MATRIX (I-R)*VAR
      WRITE(IOUT,13)
```

Program Listing—Continued

```
      CALL PRTOT(R,NTOT,NTOT,NTD)
C**COMPUTE AND PRINT CORRELATION MATRIX
      DO 210 I=1,NTOT
  210 D(I)=R(I,I)**.5
      DO 230 J=1,NTOT
      TMP=D(J)
      DO 220 I=J,NTOT
      R(I,J)=R(I,J)/(TMP*D(I))
  220 R(J,I)=R(I,J)
  230 CONTINUE
      WRITE(IOUT,15)
      CALL PRTOT(R,NTOT,NTOT,NTD)
      STOP
      END
      SUBROUTINE PRTOT(C,NR,NC,NRD)
C**PRINT MATRICES AND VECTORS
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION C(1)
      COMMON/ITP/IIN,IOUT
      IF(NC.EQ.1) GO TO 25
      DO 20 L=1,NC,10
      J10=L+9
      IF(J10.GT.NC) J10=NC
      WRITE(IOUT,35) (J,J=L,J10)
      WRITE(IOUT,50)
      KBC=(L-1)*NRD
      KEC=(J10-1)*NRD
      DO 10 I=1,NR
      KB=KBC+I
      KE=KEC+I
   10 WRITE(IOUT,40) I,(C(K),K=KB,KE,NRD)
   20 CONTINUE
      RETURN
   25 N=NR/3
      IF((3*N).NE.NR) N=N+1
      DO 30 K=1,N
   30 WRITE(IOUT,80) (L,C(L),L=K,NR,N)
      RETURN
   35 FORMAT (1H0,10(9X,I3))
   40 FORMAT (1H ,I3,1X,10(1X,G11.5))
   50 FORMAT (1H )
   80 FORMAT (1H ,2X,3(I3,7X,G11.5,3X))
      END
      SUBROUTINE PRTOTA(VALA,VALB,NO)
C**PRINT VALUES IN TWO GROUPS OF THREE COLUMNS
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION VALA(NO),VALB(NO)
      COMMON/ITP/IIN,IOUT
      NR=NO/2
      IF(2*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
      WRITE(IOUT,20) (L,VALA(L),VALB(L),L=K,NO,NR)
   10 CONTINUE
```

Program Listing—Continued

```
          RETURN
     20   FORMAT (1H ,2X,2(I3,4X,G11.5,4X,G11.5,4X))
          END
          SUBROUTINE PRTOTC(IVAL,NO)
C**PRINT INTEGERS IN THREE GROUPS OF TWO COLUMNS
          DIMENSION IVAL(NO)
          COMMON/ITP/IIN,IOUT
          NR=NO/3
          IF(3*NR.NE.NO) NR=NR+1
          DO 10 K=1,NR
          WRITE(IOUT,20) (L,IVAL(L),L=K,NO,NR)
     10   CONTINUE
          RETURN
     20   FORMAT (1H ,2X,3(I3,8X,I4,9X))
          END
          FUNCTION RANUM(IRAN)
          IMPLICIT DOUBLE PRECISION (A-H,O-Z)
          DATA MODU,MULT,NADD/1048576,1027,221589/
          IRAN=MULT*IRAN+NADD
          IRAN=IRAN-(IRAN/MODU)*MODU
          RANUM=FLOAT(IRAN)/FLOAT(MODU)
          RETURN
          END
```

Listing of Inserts to the Regression Ground-Water Flow Program.

```
C
C**INSERT AFTER STATEMENT 480
          OPEN (8,FILE='RESAN.IN',STATUS='NEW',ACCESS='SEQUENTIAL'
        1,FORM='UNFORMATTED')
          ITB=8
          REWIND ITB
          DO 1000 J=1,NVAR
   1000   WRITE(ITB) (A(I,J),I=J,NVAR)
          DO 1100 I=1,NOBS
   1100   VL(I)=W(I)*W(I)
          WRITE(ITB) (VL(I),I=1,NOBS)
          DO 1200 J=1,NOBS
   1200   WRITE(ITB) (X(I,J),I=1,NVAR)
C
```

References Cited

Beck J.V., and Arnold K.J., 1977, Parameter estimation in engineering and science: New York, John Wiley, 501 p.

Draper, N.R., and Smith, H., 1981, Applied regression analysis [2d ed]: New York, John Wiley, 709 p.

Lieberman, G.J., 1961, Prediction regions for several predictions from a single regression line: Technometrics, v. 3, p. 21–27.

Additional Reading

Beck, J.V., and Arnold, K.J., 1977, Parameter estimation in engineering and science: New York, John Wiley, 501 p.

Graybill, F. A., 1976, Theory and application of the linear model: North Scituate, Massachusetts, Duxbury, 704 p.

6 Some Advanced Topics

6.1 Advanced Models

6.1.1 Regression When the Dependent Variable is Implicit

In all cases considered in previous sections, the assumption has been made that the deterministic part of the model equation can be solved explicitly for the dependent variable. However, this may not be true in some cases. An example of such a model written in terms of the true value for the dependent variable $f=f(\xi,\beta_1,\beta_2)$ is

$$\frac{\beta_1 k}{3(\beta_1+\beta_2)}\left\{\frac{1}{2}\ln\left|\frac{(k+f)^2}{k^2-kf+f^2}\right|+\sqrt{3}\left[\tan^{-1}\left(\frac{2f-k}{k\sqrt{3}}\right)\right.\right.$$

$$\left.\left.-\tan^{-1}\left(\frac{-1}{\sqrt{3}}\right)\right]\right\}-\xi=0 \qquad (6.1-1)$$

where $k=\sqrt[3]{\beta_1+\beta_2}$ and $f=f(\xi,\beta_1,\beta_2)$ is the exact solution of equation 6.1-1. As can be seen, f is implicit in the model equation and cannot be directly solved for. A general deterministic form for an exact model (that is, a model that does not contain ϵ) where the dependent variable is implicit is

$$g[f(\xi,\beta),\xi,\beta]=0 \qquad (6.1-2)$$

where ξ and β are defined as usual.

Based on equation 6.1-2, a true regression model can be written in terms of observation vector $\underline{Y}$ and disturbances $\underline{\epsilon}$ in the usual form

$$\underline{Y}=\underline{f}(\underline{\xi},\beta)+\underline{\epsilon} \qquad (6.1-3)$$

where $\underline{f}(\underline{\xi},\beta)$ is the vector of order n that is computed from

$$g[f(\xi,\beta),\underline{\xi},\beta]=\underline{0} \ . \qquad (6.1-4)$$

In equation 6.1-4 vector g of order n represents n equations, each of which has the form of equation 6.1-2 written for an observation point. As an example, equation 6.1-1 would be written in the form

$$g_i=$$

$$\frac{\beta_1 k}{3(\beta_1+\beta_2)}\left\{\frac{1}{2}\ln\left|\frac{(k+f_i)^2}{k^2-kf_i+f_i^2}\right|+\sqrt{3}\left[\tan^{-1}\left(\frac{2f_i-k}{k\sqrt{3}}\right)\right.\right.$$

$$\left.\left.-\tan^{-1}\left(\frac{-1}{\sqrt{3}}\right)\right]\right\}-\xi_i=0, \ i=1,2,...,n \ . \qquad (6.1-5)$$

Note that in equation 6.1-5 only f evaluated at point i (that is, f_i) appears in the equation to compute g_i. However, in general this equation could contain values of f evaluated at any number of the possible points $j=1,2, ..., n$. An example of this type of model is the numerical model discussed in the next section.

The estimated regression model derived from equations 6.1-3 and 6.1-4 is

$$\underline{Y}=\underline{f}(\underline{\xi},\underline{b})+\underline{e} \qquad (6.1-6)$$

and

$$g[f(\xi,\underline{b}),\underline{\xi},\underline{b}]=\underline{0} \qquad (6.1-7)$$

where $\underline{b}$ and $\underline{e}$ are, as usual, general estimates of β and ϵ, respectively.

The general approach to solving the implicit-variable problem is very similar to that followed in section 3.3.1. First, the dependent variable values are written using a Taylor series expansion about an initial set of parameters. Then, from this, the linearized regression problem is set up and solved recursively to give the final solution to the nonlinear problem.

Taylor series expansion of $\underline{f}$ about an arbitrary initial set of parameters $\underline{b}_0$ can be written in the form of equation 3.3-5,

$$\underline{f}(\underline{\xi},\underline{b})\cong\underline{f}_0+\underline{X}_0(\underline{b}-\underline{b}_0) \qquad (6.1-8)$$

where

$$\underline{f}_0=\underline{f}(\underline{\xi},\underline{b}_0) \qquad (6.1-9)$$

$$\underline{X}_0=\left\{\left.\frac{\partial f_i}{\partial b_j}\right|_{\underline{b}=\underline{b}_0}\right\}. \qquad (6.1-10)$$

By using equation 6.1-6, equation 6.1-8 can be modified to give the estimated linearized regression model

$$\underline{Y}-\underline{f}_0=\underline{X}_0(\underline{b}-\underline{b}_0)+\underline{e} \qquad (6.1\text{-}11)$$

which is exactly the same as the model used for the standard nonlinear case discussed in section 3.3.1.

To solve the linearized regression model based on equation 6.1-11, $\underline{f}_0$ can be computed from equation 6.1-7 with $\underline{b}=\underline{b}_0$, and $\underline{X}_0$ can be computed by implicitly differentiating equation 6.1-7 with respect to $b_j (j=1,2,...,p)$, setting $\underline{b}=\underline{b}_0$, then solving for $\underline{X}_0$. To accomplish this computation of $\underline{X}_0$, note that for any differential change $d\underline{b}$ in parameter vector $\underline{b}$ to produce a new solution $\underline{f}$ of $\underline{g}=\underline{0}$, the total differential $d\underline{g}$ must equal zero because $\underline{g}$ is always zero. Hence, by employing the chain rule of calculus, there results

$$d\underline{g}=\left(\frac{\partial \underline{g}}{\partial b_j}+\underline{M}\,\frac{\partial \underline{f}}{\partial b_j}\right)db_j=\underline{0},\, j=1,2,...,p \quad (6.1\text{-}12)$$

where $\underline{M}=\{M_{ij}\}=\{\partial g_i/\partial f_j\}$. Note that if g_i contains only f_i, then $\underline{M}$ is diagonal. Equation 6.1-12 can be evaluated using $\underline{b}=\underline{b}_0$ and $\underline{f}_0=\underline{f}(\underline{\xi},\underline{b}_0)$ to give

$$\left(\frac{\partial \underline{f}}{\partial b_j}\right)_0=\underline{X}_j^0=-\underline{M}_0^{-1}\left(\frac{\partial \underline{g}}{\partial b_j}\right)_0,\, j=1,2,...,p \quad (6.1\text{-}13)$$

where

$$\underline{M}_0=\left\{\frac{\partial g_i}{\partial f_j}\bigg|_{f=f_0}\right\}_{(n\times n)} \qquad (6.1\text{-}14)$$

and subscript (or superscript) 0 means that the quantity is evaluated using $\underline{b}=\underline{b}_0$ and $\underline{f}=\underline{f}_0$. Solution of equation 6.1-7 for $\underline{f}_0$ (using, for example, Newton iteration) followed by solution of equation 6.1-13 for $\underline{X}_0$ provides a convenient method of obtaining initial values $\underline{f}_0$ and $\underline{X}_0$ from initial parameter estimates $\underline{b}_0$. However, for subsequent iterations this method can be time consuming because it involves solving $\underline{g}=\underline{0}$ each time a new vector $\underline{f}$ and a new set of sensitivity vectors $\underline{X}_j$ are to be computed from an updated parameter set.

A good method for computing good approximate values of $\underline{f}$ and $\underline{X}$ corresponding to some arbitrary parameter set $\underline{b}$ that is close to $\underline{b}_0$ involves approximating $\underline{g}$ with another Taylor series expansion. If $\underline{b}$ is close enough to $\underline{b}_0$ to allow dropping all terms except linear terms, then Taylor series expansion of equation 6.1-7 about initial set of dependent variable values $\underline{f}_0$ can be written as

$$\underline{g}[\underline{f}(\underline{\xi},\underline{b}),\underline{\xi},\underline{b}]=\underline{0}$$

$$\cong \underline{g}_0(\underline{b})+\underline{M}_0(\underline{b})(\underline{f}-\underline{f}_0) \qquad (6.1\text{-}15)$$

where $\underline{g}_0(\underline{b})=\underline{g}(\underline{f}_0,\underline{\xi},\underline{b})$, $\underline{M}_0(\underline{b})=\{\partial g_i/\partial f_j|_{f=f_0}\}$ and $\underline{f}\cong\underline{f}(\underline{\xi},\underline{b})$. By knowing $\underline{b}$, equation 6.1-15 may be solved for $\underline{f}$. Corresponding values of $\underline{X}$ are obtained as follows. If equation 6.1-15 is implicitly differentiated with respect to $b_j (j=1,2,...,p)$, there results

$$\frac{\partial \underline{g}_0(\underline{b})}{\partial b_j}=-\underline{M}_0(\underline{b})\,\frac{\partial \underline{f}}{\partial b_j}-\frac{\partial \underline{M}_0(\underline{b})}{\partial b_j}(\underline{f}-\underline{f}_0) \quad (6.1\text{-}16)$$

or,

$$\frac{\partial \underline{f}}{\partial b_j}=\underline{X}_j=-\underline{M}_0^{-1}(\underline{b})\left(\frac{\partial \underline{g}_0(\underline{b})}{\partial b_j}+\frac{\partial \underline{M}_0(\underline{b})}{\partial b_j}(\underline{f}-\underline{f}_0)\right),$$

$$j=1,2,...,p . \qquad (6.1\text{-}17)$$

By using the above results, solution of the nonlinear regression problem is obtained by a procedure analogous to the procedure followed for the standard nonlinear problem. As indicated previously, to begin the first iteration assume an initial set of parameters $\underline{b}_0$ and solve equation 6.1-7 for $\underline{f}_0=\underline{f}(\underline{\xi},\underline{b}_0)$. Then solve equation 6.1-13 for $\underline{X}_0$. Next, form and solve normal equation 3.3-10 by minimizing $S(\underline{b})$ (given by equation 3.3-4) with respect to $\underline{b}$, then scaling the resulting equations with $\underline{C}_0$. That is, form and solve for $\underline{\delta}_1$

$$\underline{S}_0^T\underline{\omega}\underline{S}_0\underline{\delta}_1=\underline{S}_0^T\underline{\omega}(\underline{Y}-\underline{f}_0) \qquad (6.1\text{-}18)$$

where

$$\underline{S}_0=\underline{X}_0\underline{C}_0 \qquad (6.1\text{-}19)$$

$$\underline{\delta}_1=\underline{C}_0^{-1}(\underline{b}_1-\underline{b}_0). \qquad (6.1\text{-}20)$$

For the second iteration, solve equation 6.1-15, written using $\underline{b}=\underline{b}_1$, for $f\text{-}f_0=f_1\text{-}f_0$. The result can be stated as

$$\underline{u}_1=-\underline{M}_1^{-1}\underline{g}_1 \qquad (6.1\text{-}21)$$

where subscript 1 on $\underline{M}$ and g indicates evaluation using the most recent values available for f and $\underline{b}$ (that is, f_0 and $\underline{b}_1$), and

$$\underline{u}_1=f_1\text{-}f_0 \qquad (6.1\text{-}22)$$

Next, solve equation 6.1-17 for $\underline{X}_1$:

$$\underline{X}_j^1=-\underline{M}_1^{-1}\left[\left(\frac{\partial\underline{g}_0(\underline{b})}{\partial b_j}\right)_1+\left(\frac{\partial\underline{M}_0(\underline{b})}{\partial b_j}\right)_1\underline{u}_1\right],$$

$$j=1,2,...,p \qquad (6.1\text{-}23)$$

where $(\cdot)_1$ indicates evaluation using $\underline{b}=\underline{b}_1$. Finally, form and solve the normal equations, written in terms of $\underline{S}_1$ and f_1, for $\underline{\delta}_2$.

For general iteration r, the equations to solve are

$$\underline{u}_r=-\underline{M}_r^{-1}\underline{g}_r \qquad (6.1\text{-}24)$$

$$\underline{X}_j^r=-\underline{M}_r^{-1}\left[\left(\frac{\partial\underline{g}_{r-1}(\underline{b})}{\partial b_j}\right)_r+\left(\frac{\partial\underline{M}_{r-1}(\underline{b})}{\partial b_j}\right)_r\underline{u}_r\right],$$

$$j=1,2,...,p \qquad (6.1\text{-}25)$$

$$f_r=\underline{u}_r+f_{r-1} \qquad (6.1\text{-}26)$$

$$\underline{S}_r^T\underline{\omega}\underline{S}_r\underline{\delta}_{r+1}=\underline{S}_r^T\underline{\omega}(\underline{Y}\text{-}f_r) \qquad (6.1\text{-}27)$$

$$\underline{b}_{r+1}=\underline{C}_r\underline{\delta}_{r+1}+\underline{b}_r \qquad (6.1\text{-}28)$$

where

$$\underline{S}_r=\underline{X}_r\underline{C}_r \ , \qquad (6.1\text{-}29)$$

$f_0=f(\underline{\xi},\underline{b}_0)$ so that $g_0=\underline{0}$, and $f_{-1}=f_0$.

At convergence of the solution g_r, $\underline{u}_r$, and $\underline{\delta}_{r+1}$ all tend to zero so that $f_r=f(\underline{\xi},\underline{b})$ where $\underline{b}=\underline{b}_{r+1}\cong\underline{b}_r$. At this point $S(\underline{b})=(\underline{Y}\text{-}f(\underline{\xi},\underline{b}))^T\underline{\omega}$ $\cdot(\underline{Y}\text{-}f(\underline{\xi},\underline{b}))$ is at a minimum, and the nonlinear regression problem has been solved.

The solution procedure given by equations 6.1-24 through 6.1-29 can actually be considered to be a generalization of the Gauss-Newton procedure discussed in section 3.3.1, because if the standard nonlinear model is stated in the form

$$g=f(\underline{\xi},\underline{b})\text{-}f=\underline{0} \qquad (6.1\text{-}30)$$

then

$$g_r=f(\underline{\xi},\underline{b}_r)\text{-}f_{r-1} \qquad (6.1\text{-}31)$$

$$\underline{M}_r=-\underline{I} \qquad (6.1\text{-}32)$$

$$\left(\frac{\partial\underline{g}_{r-1}(\underline{b})}{\partial b_j}\right)_r=\left(\frac{\partial f}{\partial b_j}\right)_r=\underline{X}_j^r \qquad (6.1\text{-}33)$$

and equations 6.1-24 through 6.1-29 become the standard Gauss-Newton algorithm.

Iteration parameters ρ and μ should be applied to the present method in the same manner as they are for the Gauss-Newton method. Use of ρ to modify step size $\underline{\delta}_{r+1}$ leads to equation 3.3-19 $(\underline{b}_{r+1}=\rho\underline{C}_r\underline{\delta}_{r+1}+\underline{b}_r)$ to compute $\underline{b}_{r+1}$. To employ μ, equation 6.1-27 is transformed to

$$(\underline{S}_r^T\underline{\omega}\underline{S}_r+\mu\underline{I})\underline{\delta}_{r+1}=\underline{S}_r^T\underline{\omega}(\underline{Y}\text{-}f_r) \ . \qquad (6.1\text{-}34)$$

The method for solution of the implicit-variable model given here requires the same three conditions to guarantee convergence to a global minimum as discussed for the modified Gauss-Newton method in section 3.3.3. However, in addition, the method requires that $\underline{M}$ and $\partial g/\partial b_j$ $(j=1,2, ..., p)$ be continuous and unique for all $\underline{b}$ belonging to region R (see equation 3.3-25).

Solution Algorithm.
1. Before the first iteration, solve equation 6.1-7 for f_0 using an initial estimate $\underline{b}_0$ for $\underline{b}$, and set $f_{-1}=f_0$.
2. Solve equation 6.1-24 and equation 6.1-25 for $\underline{u}_r$ and $\underline{X}_j^r$ $(j=1,2,...,p)$.
3. Solve equation 6.1-26 for f_r.
4. Solve equation 6.1-34 for $\underline{\delta}_{r+1}$.
5. Solve equation 3.3-17 for $\underline{d}_{r+1}$.
6. Solve equation 3.3-19 for $\underline{b}_{r+1}$.

7. Test to determine if $|d_i^{r+1}/c| > \epsilon$ $(i=1,2,...,p)$. (See step 5, Gauss-Newton algorithm.)

8. If $|d_i^{r+1}/c| > \epsilon$, increment r by one and return to 2. If not, then the process has converged.

6.1.2 Regression When the Implicit-Variable Model is Numerical

If the numerical model assumes the general form of equation 3.3-21, which for convenience is restated here as

$$\underline{D}(\underline{h},\underline{\xi},\underline{\beta})\underline{h}=\underline{q}(\underline{h},\underline{\xi},\underline{\beta}), \qquad (6.1-35)$$

then the method derived in the previous section can be applied. The solution can be conceptually developed in two stages, first making the assumption that numerical solution points coincide exactly with observation points, which implies $m=n$, then, second, relaxing the assumption by following either of the two procedures described to obtain f from $\underline{h}$ for the Gauss-Newton method in section 3.3.2.

To develop the first stage of the solution, first note that because $m=n$, equation 6.1-35 may be written in the form $\underline{g}=\underline{0}$ analogous to equation 6.1-7:

$$\underline{g}=\underline{q}-\underline{D}\underline{h}=\underline{0} \qquad (6.1-36)$$

where $\underline{h}=\underline{f}$. Next, expand equation 6.1-36 in a Taylor series to give equations exactly analogous to equations 6.1-24 and 6.1-25. Pertinent quantities in these equations are given by

$$\underline{M}_j^r=\left(\frac{\partial \underline{q}}{\partial h_i}\right)_r-\left(\frac{\partial \underline{D}}{\partial h_i}\right)_r \underline{h}_{r-1}-\underline{D}_j^r, \; i=1,2,...,m \, (6.1-37)$$

$$\left(\frac{\partial \underline{g}_{r-1}(\underline{b})}{\partial b_j}\right)_r=\left(\frac{\partial \underline{q}}{\partial b_j}\right)_r-\left(\frac{\partial \underline{D}}{\partial b_j}\right)_r \underline{h}_{r-1} \qquad (6.1-38)$$

$$\left(\frac{\partial \underline{M}_i^{r-1}(\underline{b})}{\partial b_j}\right)_r=\left(\frac{\partial^2 \underline{q}}{\partial b_j \partial h_i}\right)_r-\left(\frac{\partial^2 \underline{D}}{\partial b_j \partial h_i}\right)_r \underline{h}_{r-1}$$

$$-\left(\frac{\partial \underline{D}_i}{\partial b_j}\right)_r \qquad (6.1-39)$$

where subscript i on a matrix denotes column i of the matrix. By using equations 6.1-38 and 6.1-39, an equation analogous to 6.1-25 can be written

$$\left(\frac{\partial \underline{h}}{\partial b_j}\right)_r = -\underline{M}_r^{-1}\left\{\left[\left(\frac{\partial \underline{q}}{\partial b_j}\right)_r-\left(\frac{\partial \underline{D}}{\partial b_j}\right)_r \underline{h}_{r-1}\right]+ \right.$$

$$\sum_{i=1}^{m}\left[\left(\frac{\partial^2 \underline{q}}{\partial b_j \partial h_i}\right)_r-\left(\frac{\partial^2 \underline{D}}{\partial b_j \partial h_i}\right)_r \underline{h}_{r-1}-\left(\frac{\partial \underline{D}_i}{\partial b_j}\right)_r\right]$$

$$\left. \cdot u_i^r\right\},\, j=1,2,...,p \; . \qquad (6.1-40)$$

The second stage results from using one of the two procedures for obtaining f from $\underline{h}$ described in section 3.3.2 to obtain f from $\underline{h}$ and $\underline{X}_j$ from $\partial \underline{h}/\partial b_j$ in the present case. With f and $\underline{X}$ defined, the solution algorithm of section 6.1.1 can be applied directly.

6.2 Modified Beale's Measure of Nonlinearity

Most of the methodology discussed here to analyze regression models is based on the assumption that the model is linear in the parameters. In the case that the model is nonlinear, Beale (1960, p. 54-55) developed an empirical measure of degree of nonlinearity with respect to the confidence regions on parameters. However, Guttman and Meeter (1965, p. 635) showed that if the degree of nonlinearity is high, Beale's measure seriously underestimates it. To correct for this underestimation problem, Linssen (1975) modified Beale's measure. More recently, Bates and Watts (1980) developed measures of nonlinearity based on the concepts of differential geometry. Although these measures are based on a much more extensive theory than Beale's (1960) measure or Linssen's (1975) modification, they also require extensive computation. Thus, here Beale's empirical measure as modified by Linssen is extended to give an approximate indication of the degree of average model nonlinearity.

To develop the measure, consider a linearized model of the form of equation 5.1-1, where, for convenience, general estimate $\underline{b}_r$ replaces $\underline{\beta}$, and $\underline{b}_0$ is set equal to $\hat{\underline{b}}$ to result in

$$f_\ell^0 = \hat{f} + \underline{X}(\underline{b}_\ell - \hat{\underline{b}}) \qquad (6.2\text{-}1)$$

where $\hat{f} = f(\underline{\xi}, \hat{\underline{b}})$. To emphasize that f_ℓ^0 is a linearized estimate of $f(\underline{\xi}, \underline{b}_\ell) = f_\ell$, a superscript o is appended to f_ℓ. If f_ℓ^0 is calculated from equation 6.2-1 for m sets of parameter vectors $\underline{b}_\ell$ ($\ell = 1, 2, \ldots, m$), then a measure c^2 of model nonlinearity in the region covered by the varied parameter sets is

$$c^2 = \sum_{\ell=1}^{m} (f_\ell - f_\ell^0)^T \underline{\omega} (f_\ell - f_\ell^0) . \qquad (6.2\text{-}2)$$

Equation 6.2-2 is the sum of squared distances (that is, squared lengths of vectors) between points $\omega^{1/2} f_\ell^0$ and $\omega^{1/2} f_\ell$ in observation space. (Recall that the distance between two points is the length of the vector joining the points, and that the squared length of a vector is given by the sum of squared lengths of its components.)

As explained further on, the most useful measure of nonlinearity is obtained by multiplying equation 6.2-2 by the quantity

$$q s^2 / \sum_{\ell=1}^{m} [(f_\ell^0 - \hat{f})^T \underline{\omega} (f_\ell^0 - \hat{f})]^2$$

to obtain

$$\hat{N}_b = q s^2 \frac{\sum_{\ell=1}^{m} (f_\ell - f_\ell^0)^T \underline{\omega} (f_\ell - f_\ell^0)}{\sum_{\ell=1}^{m} [(f_\ell^0 - \hat{f})^T \underline{\omega} (f_\ell^0 - \hat{f})]^2} \qquad (6.2\text{-}3)$$

which is an extension for $q \leq p$ of Linssen's (1975) modification of Beale's measure of nonlinearity (Beale, 1960, p. 54-55).

Equation 6.2-3 can be justified as follows (see also Guttman and Meeter, 1965, and Linssen, 1975). The weighted distance between f_ℓ^0 and f_ℓ is designated ϵd so that the geometric relationships among weighted vectors $\hat{f}$, f_ℓ, and f_ℓ^0 can be diagrammed in observation space as shown in figure 6.2-1. Now,

$$\epsilon^2 = \frac{(\epsilon d)^2}{d^2} = \frac{d^2 (\epsilon d)^2}{(d^2)^2}$$

$$= d^2 \frac{(f_\ell - f_\ell^0)^T \underline{\omega} (f_\ell - f_\ell^0)}{[(f_\ell^0 - \hat{f})^T \underline{\omega} (f_\ell^0 - \hat{f})]^2} \qquad (6.2\text{-}4)$$
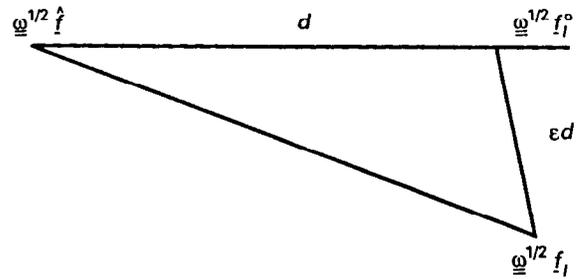
Figure 6.2-1

where, by definition,

$$d^2 = (f_\ell^0 - \hat{f})^T \underline{\omega} (f_\ell^0 - \hat{f})$$

$$= [\underline{X}(\underline{b}_\ell - \hat{\underline{b}})]^T \underline{\omega} [\underline{X}(\underline{b}_\ell - \hat{\underline{b}})]$$

$$= (\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega} \underline{X} (\underline{b}_\ell - \hat{\underline{b}}) . \qquad (6.2\text{-}5)$$

To obtain a more convenient form for d^2, note that

$$(\underline{Y} - f_\ell^0)^T \underline{\omega} (\underline{Y} - f_\ell^0) - (\underline{Y} - \hat{f})^T \underline{\omega} (\underline{Y} - \hat{f})$$

$$= (\underline{Y} - \hat{f} - \underline{X}(\underline{b}_\ell - \hat{\underline{b}}))^T \underline{\omega} (\underline{Y} - \hat{f} - \underline{X}(\underline{b}_\ell - \hat{\underline{b}}))$$

$$- (\underline{Y} - \hat{f})^T \underline{\omega} (\underline{Y} - \hat{f})$$

$$= -2(\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega} (\underline{Y} - \hat{f}) + (\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega} \underline{X} (\underline{b}_\ell - \hat{\underline{b}})$$

$$= (\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega} \underline{X} (\underline{b}_\ell - \hat{\underline{b}}) \qquad (6.2\text{-}6)$$

where equation 5.1-9 was used. The combination of equations 6.2-5 and 6.2-6 shows that

$$d^2 = (\underline{Y} - f_\ell^0)^T \underline{\omega} (\underline{Y} - f_\ell^0) - (\underline{Y} - \hat{f})^T \underline{\omega} (\underline{Y} - \hat{\underline{b}}). \quad (6.2\text{-}7)$$

If f_ℓ^0 is assumed to lie on the edge of the confidence region given by equation 5.6-12 so that $\underline{b}_\ell^T = [\tilde{b}_{1\ell}^T, \ \tilde{\beta}_{2\ell}^T]$, then from equations 5.6-12 and 6.2-7 it can be seen that

$$d^2 = q s^2 F_\alpha(q, n-p) . \qquad (6.2\text{-}8)$$

Hence, if both numerator and denominator of equation 6.2-4 are averaged over m sets of parameters, there results

$$E(\epsilon^2) \cong \hat{N}_b F_\alpha(q, n-p) . \qquad (6.2\text{-}9)$$

Based on equation 6.2–9, Beale (1960, p. 60) ranked the degrees of nonlinearity as follows: The model is highly nonlinear if

$$\hat{N}_b > 1/F_\alpha(q, n-p) \qquad (6.2\text{–}10)$$

because in this case $E(\epsilon^2) > 1$, and the discrepancy is actually greater than d. If

$$\hat{N}_b < 0.01/F_\alpha(q, n-p), \qquad (6.2\text{–}11)$$

then the model is classed as being effectively linear because $E(\epsilon^2) < 0.01$. For points in between, Beale (1960, p. 60) stated that the linear theory is adequate to give a rough idea of significance but may not bring out full implications of the analysis. However, Guttman and Meeter (1965, p. 636) noted that equation 6.2–11 may be overly conservative to define a maximum value of $\hat{N}_b$ for an approximately linear model. Experiments conducted by the authors indicate that, if

$$\hat{N}_b < 0.09/F_\alpha(q, n-p), \qquad (6.2\text{–}12)$$

then confidence intervals given by linear theory are fairly good approximations of the exact ones as given by Vecchia and Cooley (1987). Thus, equation 6.2–12 is used to define the maximum value of $\hat{N}_b$ to consider the model to be roughly linear.

Because equation 6.2–9 is justified by assuming that the points f_ℓ^o lie on the edge of the confidence region, a reasonable way to obtain the points is to choose them from equation 5.6–14, although, as noted by Beale (1960, p. 55), the points do not have to lie on the edge of the confidence region. Thus, one could use $m \leq 2q$ sets of parameters $b_\ell^T = [\underline{b}_{1\ell}^T, \tilde{\beta}_2^T]$. Note that whether or not the model is linear, $\tilde{b}_{1\ell}$ and $\tilde{\beta}_2$ corresponding to the partition of $\underline{b}$ given in equation 5.6–3 are properly chosen without the necessity of performing additional least squares solutions to obtain each set $\tilde{b}_{1\ell}$. This fact is true because subset $\tilde{b}_{1\ell}$ is required to lie on the edge of the linearized confidence region.

Rigorous use of equations 6.2–10 through 6.2–12 theoretically requires that disturbances be distributed normally. However, it would be convenient to be able to gauge the degree of nonlinearity of the model irrespective of the

properties of $\underline{\epsilon}$. If the confidence region in equation 5.6–12 were large enough to encompass virtually all physically plausible sets of parameters, then model nonlinearity as assessed using equations 6.2–10 through 6.2–12 would be meaningful. Based upon past experience, F values generated using $\alpha = 0.05$ have been found to yield such a confidence region and thus to be adequate to gauge nonlinearity.

Problem 6.2–1

Four sets of parameters that correspond to four points on the edge of the linearized confidence region in equation 5.6–13 result from problem 5.6–2. These four sets of parameters can be subdivided into two groups of two. Pick two different parameter sets from the two groups and compute two corresponding sets of drawdowns at the observation points using the nonlinear (Theis) model. Then, using the modified Beale's measure program (appendix 6.4.1), find the modified Beale's measure. Is the model nearly linear?

Problem 6.2–2

Use the four parameter sets resulting from problem 5.6–3 in the nonlinear regression flow program of appendix 4.3.4, as augmented by the inserts of appendix 6.4.1, to compute the modified Beale's measure. Are the various statistical measures obtained from the linearized model approximately valid (at least as determined from the four parameter sets employed)?

6.3 Compatibility of Prior and Regression Estimates of Parameters

If the regression model contains prior information on the parameters, an important part of the analysis to determine whether or not the model is correct is to test the null hypothesis that the prior and sample information are in agreement; in other words,

$$H_0: E(\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b}_0)) = \underline{X}_p(\underline{\beta} - \underline{b}_0) \text{ versus}$$

$$H_1: E(\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b}_0)) \neq \underline{X}_p(\underline{\beta} - \underline{b}_0).$$

As indicated in section 5.5, graphical analysis of residuals can usually detect an incompatibility between sample and prior information. However, in some cases an additional test might be desired. Theil (1963) showed that the test statistic

$$\Gamma = (\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b}_0) - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0))^T$$

$$[\sigma^2 \underline{X}_p(\underline{X}_s^T \underline{\omega}_s \underline{X}_s)^{-1} \underline{X}_p^T + \sigma^2 \underline{\omega}_p^{-1}]^{-1}$$

$$\cdot (\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b}_0) - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0)) \qquad (6.3-1)$$

where vector $\hat{\underline{b}}^*$ is the ordinary least squares estimate of vector $\underline{\beta}$, is Chi square distributed with n_p degrees of freedom $(\chi^2(n_p))$ provided that all of the assumptions given by equations 5.2-1 through 5.2-3, 5.2-6, and 5.2-10 hold true, σ^2 is known, and $\underline{\omega}$ is of the form

$$\underline{\omega} = \begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix} \qquad (6.3-2)$$

where $\underline{\omega}_s$ and $\underline{\omega}_p$ are known and symmetric positive definite of order n_s and n_p, respectively. If $\underline{\omega}$ is of the form

$$\underline{\omega} = \begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{U}^{-1}\sigma^2 \end{bmatrix} \qquad (6.3-3)$$

and σ^2 is unknown, then the test statistic

$$\hat{\Gamma} = (\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b}_0) - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0))^T$$

$$\cdot [s^2 \underline{X}_p(\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U}]^{-1}$$

$$\cdot (\underline{Y}_p - \underline{f}_p(\underline{\xi}, \underline{b}_0) - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0)) \qquad (6.3-4)$$

is asymptotically $\chi^2(n_p)$ distributed. If the computed value of $\hat{\Gamma}$, γ, is greater than $\chi_\alpha^2(n_p)$, where α indicates significance level, then the null hypothesis is rejected.

Problem 6.3-1

Using equation 6.3-4, test the compatibility of the prior estimate of the boundary head, f_{p2} and the pure regression estimate, $\hat{b}_2^*$, of problem 3.2-1. To conduct this test you will have to do an ordinary least-squares regression. The model of appendix 4.3.4 may be employed for this in the same manner as for problem 4.2-1. Use the model output to obtain the necessary quantities in equation 6.3-4.

6.4 Appendix

6.4.1 Documentation of Program to Compute the Modified Beale's Measure

This program performs a straightforward computation of the modified Beale's measure, equation 6.2-3. Vectors $\underline{f}_t$, $\underline{f}_t^o$, and $\hat{\underline{f}}_t$ are assumed to be composed of sample information and direct prior information on some or all parameters. The weight matrix for sample and prior information is assumed to be given in the form of equation 3.4-12, and the sensitivity matrix for the prior information is assumed to be given by equation 4.1-6.

There are two versions of the program. One is for general use, and all variables needed for the calculation must be read in. The other version is designed to be an integral part of the regression ground-water model documented in appendix 4.3.4 and requires only q and the extra sets of parameters needed for the Beale's measure calculation as input in addition to input already required for the regression solution.

The programs were developed using the Microsoft Fortran Compiler, Version 3.3, with the DOS 2.0 operating system on an IBM PC/XT computer with the IBM 8088 Math Coprocessor and 256 KB memory. Except for the OPEN statements near the beginning of the general code, Fortran 66 was used throughout to make the codes as machine independent as possible. The general source code is contained in file BEALE.FOR, and the version designed to be inserted into the regression code is contained in file BLEINS.FOR, both of which are in the diskette accompanying this report.

Input data for General Version.—Data Set A.
Problem size information; one card (format 4I5, F10.0).

Line Columns	Variable	Definition
1-5	NVAR	Number of parameters, p.
6-10	NRES	Number of restrictions, q.
11-15	NOBS	Number of sample observations, n_s.
16-20	NPRIR	Number of regression parameters having direct prior information, n_p.
21-25	NPTS	Number of data sets to compute the modified Beale's measure, m.
26-35	VAR	Error variance, s^2.

Data Set B.
Estimated regression parameters, $\hat{\underline{b}}$ (format 8F10.0).

Line columns	Variable	Definition
1-10	BOPT(1)	Estimated regression parameters, entered sequentially
11-20	BOPT(2)	from 1 through NVAR.
⋮	⋮	
	BOPT(NVAR)	

Data Set C.
Dependent variable vector for sample information, $\hat{\underline{f}}_s$, computed using $\hat{\underline{b}}$ (format 8F10.0).

Line columns	Variable	Definition
1-10	FOPT(1)	Computed dependent variable values, entered sequen-
11-20	FOPT(2)	tially from 1 through NOBS.
⋮	⋮	
	FOPT(NOBS)	

Data Set D.
Weight matrix for sample information, $\underline{V}_s^{-1}$ (format 8F10.0).

Line columns	Variable	Definition
1-10	W(1)	Diagonal weight matrix for sample information,
11-20	W(2)	entered sequentially from 1 through NOBS.
⋮	⋮	
	W(NOBS)	

Data Set E.
Sensitivity matrix for sample information, $\underline{X}_s$ (format 8F10.0)

Line columns	Variable	Definition
1-10	X(1,1)	Sensitivity matrix for sample information, entered se-
11-20	X(2,1)	quentially 1 through NVAR for each observation.
⋮	⋮	Each new observation begins a new line, for a total
	X(NVAR,1)	of NOBS observations.
1-10	X(1,2)	
⋮	⋮	
	X(NVAR,2)	
	⋮	
	X(NVAR,NOBS)	

Data Set F.
Parameter numbers having prior information (format 16I5).

Line columns	Variable	Definition
1–5 6–10 ⋮	IPR(1) IPR(2) ⋮ IPR(NPRIR)	Array subscript numbers for regression parameters in BOPT(1) having prior information, entered in any order from 1 through NPRIR.

Omit data set of NPRIR=0.

Data Set G.
Standard deviation matrix for prior information, $\underline{U}^{\frac{1}{2}}$ (format 8F10.0).

Line columns	Variable	Definition
1–10 ⋮	WP(1) ⋮ WP(NPRIR)	Diagonal standard deviation matrix for prior information, entered in the same order as IPR(I) from 1 through NPRIR.

Omit data set if NPRIR=0.

Data Set H.
Alternate parameters sets, $\underline{b}_\ell$ (format 8F10.0). This data set and the next one are read in sequence (H, I, H, I, ...) a total of NPTS times.

Line columns	Variable	Definition
1–10 11–20 ⋮	B(1) B(2) ⋮ B(NVAR)	Alternate parameter sets, entered sequentially 1 through NVAR. Order must be the same as for BOPT(1).

Data Set I.
Alternate dependent variable vectors for sample information, $\underline{f}_{s\ell}$, computed using $\underline{b}_\ell$ (format 8F10.0).

Line columns	Variable	Definition
1–10 11–20 ⋮	FC(1) FC(2) ⋮ FC(NOBS)	Alternate sample dependent variable values computed using the nonlinear model, entered sequentially 1 through NOBS. Order must be the same as for FOPT(I).

Output for General Version.—Output is all clearly labeled; it is ordered as follows:

1. Data sets A through G.
2. Data sets H and I. Data for numbers 2 through 4 below are printed sequentially for each data set ℓ ($\ell=1,2,...,m$).
3. Dependent variable vector, $f_{s\ell}^o$, for sample information, computed using the linearized model.
4. Total sums of squared differences $(f_\ell-\hat{f})^T$ $\cdot\underline{\omega}(f_\ell-\hat{f})$ and $(f_\ell^o-\hat{f})^T\underline{\omega}(f_\ell^o-\hat{f})$, where

$$f_\ell^T=[f_{s\ell}^T, f_{p\ell}^T], \quad f_\ell^{oT}=[f_{s\ell}^{oT}, f_{p\ell}^T], \quad \hat{f}^T=[\hat{f}_s^T, \hat{f}_p^T], \text{ and}$$

$$\underline{\omega}=\begin{vmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{U}^{-1}s^2 \end{vmatrix}.$$

5. Beale's measure, $\hat{N}_b=BN$.

Use of Version Integral with the Regression Ground-Water Program.—This version consists of sets of statements to be inserted into the program of appendix 4.3.4, as indicated on the appended listing. Input is the same as if a regression solution were to be obtained, except that the initial set of parameters must be the optimum set $\hat{b}$, and extra data relating to the modified Beale's measure is required. After entering data set T, use data sets U and V to enter the data for $\ell=1,2,...,m$ alternate solutions. Follow these data with a final line to input q and s^2 with format I5,F10.0. A complete regression solution is not obtained; only computations through the calculation of sensitivities on the first iteration are completed before proceeding to calculate the modified Beale's measure. Thus, output consists of regression output through number 19 (see "Output" in appendix 4.3.4) plus output analogous to numbers 2 through 4 of the general version of the modified Beale's measure program.

Program Listing for General Version.

```
C        MODIFIED BEALE'S MEASURE PROGRAM BY R. L. COOLEY, USGS, DENVER,
C        COLO.
         DIMENSION BOPT(20),FOPT(70),B(20),FC(70),FL(70),X(20,70)
        1,W(70),IPR(20),WP(20)
         COMMON/ITP/IIN,IOUT
         COMMON/FLT/X
         OPEN (5,FILE='BEALE.DAT',STATUS='OLD',ACCESS='SEQUENTIAL'
        1,FORM='FORMATTED')
         OPEN (6,FILE='BEALE.OUT',STATUS='NEW',ACCESS='SEQUENTIAL'
        1,FORM='FORMATTED')
C**FORMAT LIST
    1 FORMAT (5I5,F10.0)
    2 FORMAT (8F10.0)
    3 FORMAT (9H1NVAR  - ,I4/9H NRES  - ,I4/9H NOBS  = ,I4
       1/9H NPRIR - ,I4/9H NPTS  - ,I4/9H VAR   - ,G11.5)
    4 FORMAT (1H0,26X,18HOPTIMUM PARAMETERS
       1/1H ,3X,3(3HNO.,9X,4HBOPT,8X))
    5 FORMAT (1H0,9X,52HDEPENDENT VARIABLES COMPUTED WITH OPTIMUM PARAME
       1TERS/1H ,3X,3(3HNO.,9X,4HFOPT,8X))
    6 FORMAT (1H0,21X,26HPARAMETERS FOR SAMPLE NO. ,I3
       1/1H ,3X,3(3HNO.,11X,1HB,9X))
    7 FORMAT (1H0,12X,44HDEPENDENT VARIABLES COMPUTED FOR SAMPLE NO. ,I3
       1/1H ,3X,3(3HNO.,10X,2HFC,9X))
    8 FORMAT (38H0 SENSITIVITIES FOR OPTIMUM PARAMETERS)
    9 FORMAT (1H0,6X,55HLINEARIZED DEPENDENT VARIABLES COMPUTED FOR SAMP
       1LE NO. ,I3/1H ,3X,3(3HNO.,10X,2HFL,9X))
   10 FORMAT (1H0,5HBN - ,G11.5)
   11 FORMAT (23H0SS((FC-FOPT)*W**.5) - ,G11.5
       1/23H SS((FL-FOPT)*W**.5) - ,G11.5)
   12 FORMAT (1H0,14X,42HRELIABILITY WEIGHTS FOR SAMPLE INFORMATION
       1/1H ,3X,3(3HNO.,10X,1HW,10X))
   13 FORMAT (16I5)
   14 FORMAT (1H0,12X,43HNO.S OF PARAMETERS HAVING PRIOR INFORMATION
       1/1H ,3X,3(3HNO.,8X,3HIPR,10X))
   15 FORMAT (1H0,14X,40HSTANDARD DEVIATIONS OF PRIOR INFORMATION
       1/1H ,3X,3(3HNO.,10X,2HWP,9X))
   16 FORMAT (6H0EV - ,G11.5)
C**DEFINE INPUT FILE, OUTPUT FILE, AND ARRAY DIMENSION
         IIN-5
         IOUT=6
         NVD-20
C**READ BASE DATA
         READ(IIN,1) NVAR,NRES,NOBS,NPRIR,NPTS,VAR              SET A
         WRITE(IOUT,3) NVAR,NRES,NOBS,NPRIR,NPTS,VAR
         READ(IIN,2) (BOPT(J),J=1,NVAR)                         SET B
         WRITE(IOUT,4)
         CALL PRTOTB(BOPT,NVAR)
         READ(IIN,2) (FOPT(I),I-1,NOBS)                         SET C
         WRITE(IOUT,5)
         CALL PRTOTB(FOPT,NOBS)
         READ(IIN,2) (W(I),I-1,NOBS)                            SET D
         WRITE(IOUT,12)
         CALL PRTOTB(W,NOBS)
```

Program Listing for General Version—Continued

```
      DO 20 J=1,NOBS
      READ(IIN,2) (X(I,J),I=1,NVAR)                        SET E
   20 CONTINUE
      WRITE(IOUT,8)
      CALL PRTOT(X,NVAR,NOBS,NVD)
      IF(NPRIR.LT.1) GO TO 45
      READ(IIN,2) EV                                       SET F
      WRITE(IOUT,16) EV
      READ(IIN,13) (IPR(I),I=1,NPRIR)                      SET G
      WRITE(IOUT,14)
      CALL PRTOTC(IPR,NPRIR)
      READ(IIN,2) (WP(I),I=1,NPRIR)                        SET H
      WRITE(IOUT,15)
      CALL PRTOTB(WP,NPRIR)
      DO 40 I=1,NPRIR
   40 WP(I)=EV/(WP(I)*WP(I))
C**READ DATA FOR EACH SAMPLE AND COMPUTE MODIFIED BEALE'S MEASURE, BN
   45 SUMA=0.
      SUMB=0.
      DO 80 M=1,NPTS
      READ(IIN,2) (B(J),J=1,NVAR)                          SET I
      WRITE(IOUT,6) M
      CALL PRTOTB(B,NVAR)
      READ(IIN,2) (FC(I),I=1,NOBS)                         SET J
      WRITE(IOUT,7) M
      CALL PRTOTB(FC,NOBS)
      SUMC=0.
      SUMD=0.
      DO 60 J=1,NOBS
      SUM=FOPT(J)
      DO 50 I=1,NVAR
   50 SUM=SUM+X(I,J)*(B(I)-BOPT(I))
      FL(J)=SUM
      TMP=FC(J)-SUM
      SUMA=SUMA+TMP*W(J)*TMP
      TMP=FC(J)-FOPT(J)
      SUMC=SUMC+TMP*W(J)*TMP
      TMP=SUM-FOPT(J)
      SUMD=SUMD+TMP*W(J)*TMP
   60 CONTINUE
      IF(NPRIR.LT.1) GO TO 75
      DO 70 J=1,NPRIR
      I=IPR(J)
      TMP=B(I)-BOPT(I)
      TMP=TMP*WP(J)*TMP
      SUMC=SUMC+TMP
   70 SUMD=SUMD+TMP
   75 WRITE(IOUT,9) M
      CALL PRTOTB(FL,NOBS)
      WRITE(IOUT,11) SUMC,SUMD
   80 SUMB=SUMB+SUMD*SUMD
      TMP=NRES
      BN=TMP*VAR*SUMA/SUMB
```

Program Listing for General Version—Continued

```
      WRITE(IOUT,10) BN
      STOP
      END
      SUBROUTINE PRTOTB(VAL,NO)
C**PRINT VALUES IN THREE GROUPS OF TWO COLUMNS
      DIMENSION VAL(NO)
      COMMON/ITP/IIN,IOUT
      NR=NO/3
      IF(3*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
      WRITE(IOUT,20) (L,VAL(L),L=K,NO,NR)
   10 CONTINUE
      RETURN
   20 FORMAT (1H ,2X,3(I3,7X,G11.5,3X))
      END
      SUBROUTINE PRTOTC(IVAL,NO)
C**PRINT INTEGERS IN THREE GROUPS OF TWO COLUMNS
      DIMENSION IVAL(NO)
      COMMON/ITP/IIN,IOUT
      NR=NO/3
      IF(3*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
      WRITE(IOUT,20) (L,IVAL(L),L=K,NO,NR)
   10 CONTINUE
      RETURN
   20 FORMAT (1H ,2X,3(I3,8X,I4,9X))
      END
      SUBROUTINE PRTOT(C,NR,NC,NRD)
C**PRINT MATRICES DIVIDED VERTICALLY INTO TEN-COLUMN BLOCKS
      DIMENSION C(NRD,NC)
      COMMON/ITP/IIN,IOUT
      DO 60 K=1,NC,10
      J10=K+9
      IF(J10.GT.NC) J10=NC
      WRITE(IOUT,70) (J,J=K,J10)
      WRITE(IOUT,90)
      DO 30 I=1,NR
   30 WRITE(IOUT,80) I,(C(I,J),J=K,J10)
   60 CONTINUE
   70 FORMAT(1H0,10(9X,I3))
   80 FORMAT (1H ,I3,1X,10(1X,G11.5))
   90 FORMAT (1H )
      RETURN
      END
```

Listing of Inserts to the Regression Ground-Water Flow Program.

```
C
C**INSERT JUST BEFORE EQUIVALENCE STATEMENT FOR MODIFIED BEALE'S MEASURE
      DIMENSION BOPT(20),HOPT(70)
      EQUIVALENCE (P(1),BOPT(1)),(HO(1),HOPT(1))
C
C
C**INSERT AFTER STATEMENT LABEL 260 FOR MODIFIED BEALE'S MEASURE
      DO 1000 J=1,NVAR
 1000 BOPT(J)=B(J)
      DO 1100 I=1,NOBS
      K=KOBS(I)
 1100 HOPT(I)=BK(I)*HC(K)+BL(I)*HC(K+1)+BM(I)*HC(K+ID)+BN(I)*HC(K+ID+1)
      SUMA=0.
      SUMB=0.
      GO TO 640
C
C
C**INSERT JUST BEFORE STATEMENT LABEL 690 FOR MODIFIED BEALE'S MEASURE
      SUMC=0.
      SUMD=0.
      WRITE(IOUT,2000)
 2000 FORMAT (1H0,3X,28HCOMPUTED AND LINEARIZED HEADS/1H ,3X,3HNO.,7X
     1,2HHC,13X,2HHL)
      DO 2200 J=1,NOBS
      K=KOBS(J)
      HCJ=BK(J)*HC(K)+BL(J)*HC(K+1)+BM(J)*HC(K+ID)+BN(J)*HC(K+ID+1)
      HL=HOPT(J)
      DO 2100 I=1,NVAR
 2100 HL=HL+X(I,J)*(B(I)-BOPT(I))
      TMP=HCJ-HL
      SUMA=SUMA+TMP*W(J)*TMP
      TMP=HCJ-HOPT(J)
      SUMC=SUMC+TMP*W(J)*TMP
      TMP=HL-HOPT(J)
      SUMD=SUMD+TMP*W(J)*TMP
      WRITE(IOUT,856) J,HCJ,HL
 2200 CONTINUE
      IF(NPRIR.LT.1) GO TO 2400
      DO 2300 J=1,NVAR
      IF(WP(J).LT.1.E-10) GO TO 2300
      TMP=B(J)-BOPT(J)
      TMP=TMP*WP(J)*TMP
      SUMC=SUMC+TMP
      SUMD=SUMD+TMP
 2300 CONTINUE
 2400 WRITE(IOUT,2500) SUMC,SUMD
 2500 FORMAT (23H0SS((HC-HOPT)*W**.5) = ,G11.5
     1/23H SS((HL-HOPT)*W**.5) = ,G11.5)
      SUMB=SUMB+SUMD*SUMD
C
C
C**INSERT AFTER STATEMENT LABEL 690 FOR MODIFIED BEALE'S MEASURE
```

Listing of Inserts to the Regression Ground-Water Flow Program—Continued

```
      READ(IIN,812) NRES,VAR
      WRITE(IOUT,2600) NRES,VAR
 2600 FORMAT (8H0NRES = ,I4/7H VAR = ,G11.5)
      TMP=NRES
      BLN=TMP*VAR*SUMA/SUMB
      WRITE(IOUT,2700) BLN
 2700 FORMAT (1H0,5HBN = ,G11.5)
C
```

References Cited

Bates, D.M., and Watts, D.G., 1980, Relative curvature measures of nonlinearity: Journal of the Royal Statistical Society, Series B, v. 42, no. 1, p. 1–25.

Beale,E.M.L., 1960, Confidence regions in nonlinear estimation: Journal of the Royal Statistical Society, Series B, v. 22, no. 1, p. 41–76.

Guttman, I., and Meeter, D.A., 1965, On Beale's measure of nonlinearity: Technometrics, v. 7, no. 4, p. 623–637.

Linssen, H.N., 1975, Nonlinearity measures: a case study: Statistica Neerlandica, v. 29, p. 93–99.

Theil, H., 1963, On the use of incomplete prior information in regression analysis: American Statistical Association Journal, v. 58, no. 302, p. 401–414.

Vecchia, A.V., and Cooley, R.L., 1987, Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater flow model: Water Resources Research, v. 23, no. 7, p. 1237–1250.

Additional Reading

Cooley, R.L., 1977, A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 1—Theory and numerical properties: Water Resources Research, v. 13, no. 2, p. 318–324.

_____1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1—Theory: Water Resources Research, v. 18, no. 4, p. 965–976.

7 Answers to Exercises

In this section answers to the exercises from the first six sections are given. For brevity, the exercises are not restated.

Problem 2.2-1

a. (RR),(RW),(RB),(WW),(WR),(WB),(BB), (BW),(BR).

b.

	R	W	B1	B2
R	RR	RW	RB1	RB2
W	WR	WW	WB1	WB2
B1	B1R	B1W	B1B1	B1B2
B2	B2R	B2W	B2B1	B2B2

$P(RR)=1/16$ $P(WW)=1/16$ $P(BB)=1/4$
$P(RW)=1/16$ $P(WR)=1/16$ $P(BW)=1/8$
$P(RB)=1/8$ $P(WB)=1/8$ $P(BR)=1/8$

c.
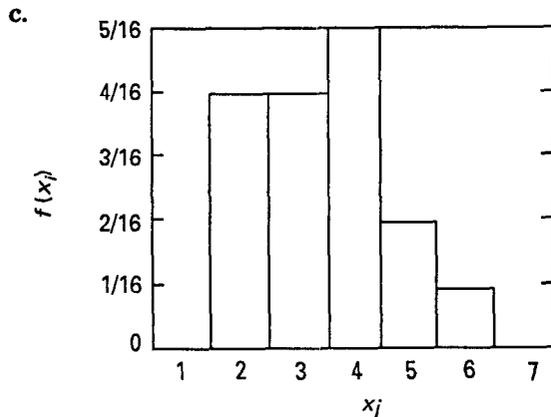
d. $P(X=4)=5/16$
    (RR),(WB),(BW).

Problem 2.2-2

a. The effect of increasing the class interval size is to increase the frequency value for each class. This increase tends to make the histogram appear more peaked for larger intervals. Increasing the class interval also tends to smooth out irregularities.
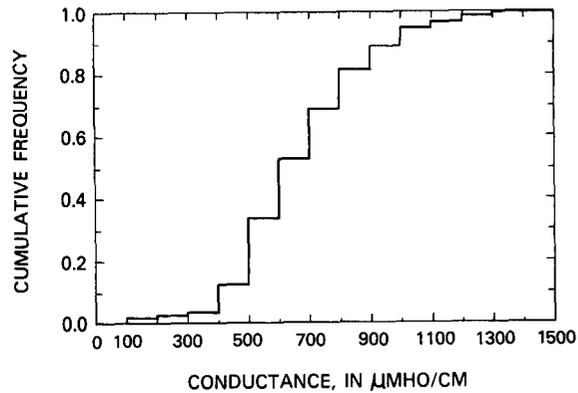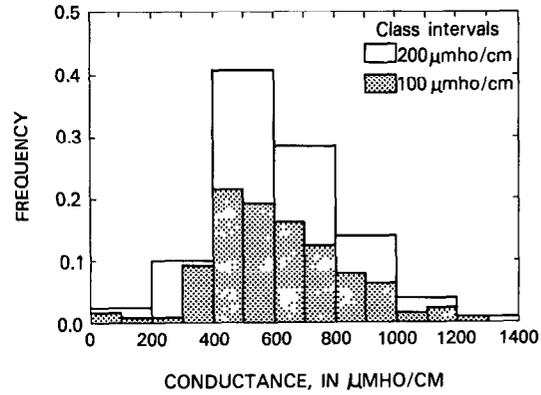
b. $P(X \leq 600)=0.53$
    $P(X>400)=0.88$
    $P(400<X \leq 600)=0.41$
    $P(X \leq 1300)=1.0.$

Problem 2.3-1

a. $\bar{x}= \sum_{(all\ i)} \bar{x}_i f_i^{*} =(2\times50+1\times150+1\times250+12$
    $\times350+28\times450+25\times550+21\times650+16$
    $\times750+10\times850+8\times950+2\times1050+3$
    $\times1150+1\times1250)/130$
    $=612.$

b. By replacing the integration in equation 2.3-3 with a summation, it is seen that the population mean μ_X for this random variable is represented by

$$\mu_X= \sum_{i=2}^{6} i f_i$$

$$=(2\times1/4+3\times1/4+4\times5/16+5\times1/8+6\times1/16)$$

$$=3\frac{1}{2}.$$

Problem 2.3-2

a. Triangular density problem.

(*i*)
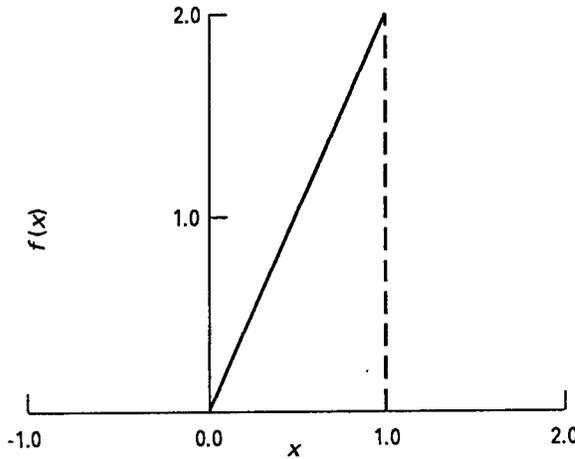
(*ii*)

$$F(x)=\int_{-\infty}^{x}f(s)ds=\begin{cases}1 & ,x>1\\ 2\int_{0}^{x}sds=x^2 & ,0\leq x\leq 1\\ 0 & ,x<0\end{cases}$$

(*iii*)
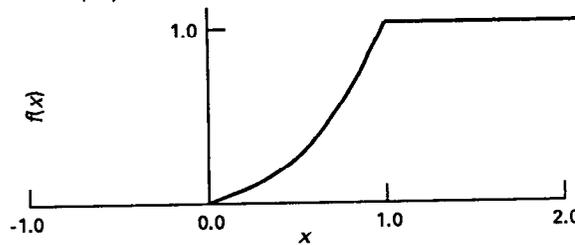
$$E[X]=\int_{-\infty}^{\infty}sf(s)ds=2\int_{0}^{1}s^2ds=\frac{2}{3}$$

$$E[X^2]=\int_{-\infty}^{\infty}s^2f(s)ds=2\int_{0}^{1}s^3ds=\frac{1}{2}$$

$$\sigma_X^2=E[X^2]-(E[X])^2=\frac{1}{2}-\frac{4}{9}=\frac{1}{18}.$$

b. $\bar{x}=1.25\times0.051+1.75\times0.103+2.25\times0.103$
$+2.75\times0.154+3.25\times0.154+3.75\times0.128$
$+4.25\times0.128+4.75\times0.090+5.25\times0.026$
$+5.75\times0.064=3.36.$

$$s_x^{*2}=\sum_{i=1}^{10}(\bar{x}_i-\bar{x})^2f_i^*$$

$$=(-2.11)^2\times0.051+(-1.61)^2\times0.103$$
$$+(-1.11)^2\times0.103+(-0.61)^2\times0.154$$
$$+(-0.11)^2\times0.154+(0.39)^2\times0.128$$
$$+(0.89)^2\times0.128+(1.39)^2\times0.090$$
$$+(1.89)^2\times0.026+(2.39)^2\times0.064=1.43.$$

Problem 2.4-1

a. $P(X=3 \text{ and } Y=2)=P(X=3)P(Y=2)=1/36.$
b. $P(X+Y=5)=1/9.$
c. $P(Y=2|X=3)=1/6.$
d. $P(X+Y=5|X=3)=1/6.$
e. $P(X+Y\leq5|X=3)=1/3.$

Problem 2.4-2

a. $E[Y]=a_1\mu_{X_1}+a_2\mu_{X_2}+a_3\mu_{X_3}$

$\text{Var}[Y]=E\{[Y-E(Y)]^2\}=E\{[a_1(X_1-\mu_{X_1})$
$+a_2(X_2-\mu_{X_2})+a_3(X_3-\mu_{X_3})]^2\}$
$=E[a_1^2(X_1-\mu_{X_1})^2+a_2^2(X_2-\mu_{X_2})^2$
$+a_3^2(X_3-\mu_{X_3})^2$
$+2a_1a_2(X_1-\mu_{X_1})(X_2-\mu_{X_2})$
$+2a_1a_3(X_1-\mu_{X_1})(X_3-\mu_{X_3})$
$+2a_2a_3(X_2-\mu_{X_2})(X_3-\mu_{X_3})]$
$=a_1^2\sigma_{X_1}^2+a_2^2\sigma_{X_2}^2+a_3^2\sigma_{X_3}^2$
$+2a_1a_2\sigma_{X_1X_2}+2a_1a_3\sigma_{X_1X_3}$
$+2a_2a_3\sigma_{X_2X_3}$

because $\sigma_{X_i}^2=E[(X_i-\mu_{X_i})^2]$ and $\sigma_{X_iX_j}=E[(X_i-\mu_{X_i})\cdot(X_j-\mu_{X_j})]$.

b. $\underline{a}\text{Var}[\underline{X}]\,\underline{a}^T=\underline{a}E[(\underline{X}-E[\underline{X}])(\underline{X}-E[\underline{X}])^T]\underline{a}^T$
$=\underline{a}E[\underline{X}\underline{X}^T-\underline{X}E[\underline{X}]^T-E[\underline{X}]\underline{X}^T$
$+E[\underline{X}]E[\underline{X}]^T]\underline{a}^T$
$=\underline{a}(E[\underline{X}\underline{X}^T]-E[\underline{X}]E[\underline{X}]^T)\underline{a}^T$
$=E[\underline{a}\underline{X}\underline{X}^T\underline{a}^T]-E[\underline{a}\underline{X}]E[\underline{X}^T\underline{a}^T]$
$=\text{Var}[\underline{a}\underline{X}]$

c. $\text{Var}[\underline{Y}]=\begin{vmatrix}\sigma_{Y_1}^2 & \sigma_{Y_2Y_2} & \cdots & \sigma_{Y_1Y_p}\\ & \sigma_{Y_2}^2 & \cdots & \sigma_{Y_2Y_p}\\ \text{symmetry} & & & \vdots\\ & & & \sigma_{Y_p}^2\end{vmatrix}$

$$\sigma^2_{Y_i} = \mathrm{Var}[Y_i] = \mathrm{Var}[\underline{a_i}\underline{X}] = \underline{a_i}\mathrm{Var}[\underline{X}]\underline{a_i}^T$$

$$\sigma_{Y_iY_j} = E[Y_iY_j] - E[Y_i]E[Y_j]$$
$$= E[\underline{a_i}\underline{X}\underline{X}^T\underline{a_j}^T] - E[\underline{a_i}\underline{X}]E[\underline{X}^T\underline{a_j}^T]$$
$$= \underline{a_i}(E[\underline{X}\underline{X}^T] - E[\underline{X}]E[\underline{X}^T])\underline{a_j}^T$$
$$= \underline{a_i}\mathrm{Var}[\underline{X}]\underline{a_j}^T$$

$$\mathrm{Var}[\underline{Y}] = \begin{vmatrix} \underline{a_1} \\ \underline{a_2} \\ \vdots \\ \underline{a_p} \end{vmatrix} \mathrm{Var}[\underline{X}][\underline{a_1}^T,\underline{a_2}^T,\ldots,\underline{a_p}^T]$$

$$= \underline{A}\,\mathrm{Var}[\underline{X}]\underline{A}^T$$

Problem 2.5–1

a. $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n}x_i = 610$.

The results differ very little, indicating the validity of approximation in equation 2.3–1. The repeat values probably result from measurement errors, indicating that a value of the random variable could not be determined more precisely than one μmho/cm.

b. $\bar{x} = 3.36$ $\quad s^2_X = 1.43$
$\bar{t}_g = 10^{\bar{x}} = 2{,}290$ gal/d/ft
$d_g = 10^{(\bar{x} \pm s_x)} = \bar{t}_g \times 10^{\pm s_x}$
$\quad = (146,\ 35900)$ gal/d/ft.

The large dispersion strongly suggests that $\bar{t}_g$ may not represent the true geometric mean, since there is significant scatter in T in the vicinity of $\bar{t}_g$.

Problem 2.5–2

Let X be the specific conductance random variable and Y be the dissolved solids random variable. Then

$\sum x_i = 14{,}999$	$\sum y_i = 9{,}140$
$\sum x_i^2 = 11{,}743{,}379$	$\sum y_i^2 = 4{,}316{,}966$
$\sum x_iy_i = 7{,}095{,}973$	$n = 23$
$\bar{x} = 652$	$\bar{y} = 397$
$s^2_X = 89{,}200$	$s^2_Y = 31{,}300$
$s_X = 299$	$s_Y = 176$

$$r_{XY} = 0.98 \ .$$

Problem 2.6–1

$$\bar{x} = 3 \times 10^{-4} \qquad s^2_X = 6.6 \times 10^{-5}$$

$$P(\bar{X}^2/(S^2_X/n) \le F_\alpha(1,n-1))$$
$$= P(-\sqrt{F_\alpha(1,n-1)} \le \bar{X}/(S_X/\sqrt{n}) \le \sqrt{F_\alpha(1,n-1)})$$
$$= P(-2.26 \le \bar{X}/(S_X/\sqrt{n}) \le 2.26) = 0.95$$

$$\bar{x}/(s_X/\sqrt{n}) = 0.12 \ .$$

The value of the statistic is well within the interval $(-2.26 < 0.12 < 2.26)$. This result is expected, if the titration experiment is valid, as 95 percent of all values of the statistic $\bar{X}/(S_X/\sqrt{n})$, calculated from repeated random sampling, would be expected to fall in this interval. If the value had fallen outside the interval, one should feel uneasy because this should occur only 5 percent of the time. One would then be obligated to question whether the assumption $\mu_X = 0$ inherent to the titration test is valid.

Problem 2.8–1

a. $\bar{x} = 10$ $\quad s^2_X = 0.08$ $\quad n = 7$

$$P\left(\frac{(\bar{X} - \mu_X)^2}{S^2_X/n} \le F_\alpha(1,6)\right)$$

$$= P\left(-\sqrt{F_\alpha(1,6)} \le \frac{\bar{X} - \mu_X}{S_X/\sqrt{n}} \le \sqrt{F_\alpha(1,6)}\right) \ .$$

Therefore, an interval can be constructed from

$$-\sqrt{F_\alpha(1,6)} \le \frac{\bar{x} - \mu_X}{s_X/\sqrt{n}} \le \sqrt{F_\alpha(1,6)}$$

where $F_\alpha(1,6) = 5.99$. Thus

$$-s_X\sqrt{\frac{F_\alpha(1,6)}{n}} \le (\bar{x} - \mu_X) \le s_X\sqrt{\frac{F_\alpha(1,6)}{n}}$$

or

$$\bar{x} + s_X\sqrt{\frac{F_\alpha(1,6)}{n}} \ge \mu_X \ge \bar{x} - s_X\sqrt{\frac{F_\alpha(1,6)}{n}}$$

and

$$10.26 \ge \mu_X \ge 9.74 \ .$$

b. As 95 percent of all intervals so constructed will contain μ_X, there is a 0.95 probability that this interval contains μ_X.

Problem 2.9-1

a. $H_0 : \mu_X = 0$

$H_1 : \mu_X \neq 0$.

b. From equation 2.6-21, the statement $P(\text{reject } H_0 / H_0 \text{ true}) = \alpha$ becomes

$$P\left(\frac{S_1/\sigma_1^2}{S_2^2/\sigma_2^2} > a \,\bigg|\, \frac{\sigma_1^2}{\sigma_2^2} = 1\right) = P(S_1^2/S_2^2 > c) = 0.05$$

where S_1^2/S_2^2 is an $F(24,15)$ random variable. Because the critical region is defined by values of $F(24,15)$ greater than or equal to c, c must be equal to $F_{0.05}(24,15)$ (see equation 2.6-15). Thus, $c = 2.29$ and, because $s_1^2/s_2^2 = 1.31$, we accept H_0 at 0.05 significance level.

c. Hypothesis to be tested for rejection:

$$H_0 : \mu_X = 9.8 \ .$$

Alternate hypothesis:

$$H_1 : \mu_X \neq 9.8 \ .$$

From equation 2.9-7,

$$P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} < -\sqrt{F_\alpha(\nu_1, \nu_2)}\right)$$

$$+ P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} > \sqrt{F_\alpha(\nu_1, \nu_2)}\right)$$

$$= P\left(\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}}\right)^2 > F_\alpha(1, n-1)\right) = \alpha \ .$$

Thus, the critical value for the statistic $(\bar{X} - \mu_0)^2/(S_X^2/n)$ is $F_\alpha(1, n-1)$, which, at the 0.05 significance level, has a value of 5.99. Because the statistic for the random sample in question takes on the value 3.50, we are forced to accept the possibility that $\mu_X = 0$: We cannot safely reject H_0 at the 0.05 significance level.

Problem 3.1-1

Upon substitution of equation 4 into equation 3, one obtains

$$h = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ .$$

For n observations

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\epsilon}$$

where $\underline{Y} - \underline{h} = \underline{\epsilon}$,

$$\underline{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_{n_s} \end{bmatrix} \quad \underline{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_{n_s} \end{bmatrix} \quad \underline{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_3 \end{bmatrix} \quad \underline{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_s} \end{bmatrix}$$

and

$$\underline{X} = \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ \vdots & \vdots & \vdots \\ X_{n_s 1} & X_{n_s 2} & X_{n_s 3} \end{bmatrix} \ .$$

In the regression model:

$\underline{Y}$ = observed dependent variable vector;
$\underline{h}$ = computed dependent variable vector;
s = independent variable (distance along stream tube);
$\underline{X}$ = sensitivities; and
$\underline{\beta}$ = parameters.

a. $S(\underline{b}) = [\underline{Y} - \underline{X}\underline{b}]^T [\underline{Y} - \underline{X}\underline{b}]$

where

$$\underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

That is,

$$S(\underline{b}) = \left[Y_1 - (b_1 X_{11} + b_2 X_{12} + b_3 X_{13}) \right]^2$$

$$+ \left[Y_2 - (b_1 X_{21} + b_2 X_{22} + b_3 X_{23}) \right]^2$$

$$+ \dots$$

$$+ \left[Y_{n_s} - (b_1 X_{n_s 1} + b_2 X_{n_s 2} + b_3 X_{n_s 3}) \right]^2 .$$

b. $S(\underline{b})=[\underline{Y}-\underline{Xb}]^T\underline{\omega}[\underline{Y}-\underline{Xb}]$

where $\underline{\omega}=\underline{V}^{-1}$ and

$$\underline{V}=\begin{bmatrix} \sigma_1^2/\sigma^2 & & & \underline{0} \\ & \sigma_2^2/\sigma^2 & \cdot & \\ & & \cdot & \\ \underline{0} & & & \sigma_{n_s}^2/\sigma^2 \end{bmatrix}$$

so that

$$\underline{V}^{-1}=\begin{bmatrix} \sigma_1^{-2} & & & \underline{0} \\ & \sigma_2^{-2} & \cdot & \\ & & \cdot & \\ \underline{0} & & & \sigma_{n_s}^{-2} \end{bmatrix}\sigma^2 \ .$$

That is,

$$S(\underline{b})=\sigma^2\left\{ \left[\frac{1}{\sigma_1}(Y_1-(b_1X_{11}+b_2X_{12}+b_3X_{13}))\right]^2 \right.$$

$$+\left[\frac{1}{\sigma_2}(Y_2-(b_1X_{21}+b_2X_{22}+b_3X_{23}))\right]^2$$

$$+ \ . \ . \ .$$

$$\left. +\left[\frac{1}{\sigma_{n_s}}(Y_{n_s}-(b_1X_{n_s1}+b_2X_{n_s2}+b_3X_{n_s3}))\right]^2\right\}.$$

c. $S(\underline{b})=[\underline{Y}-\underline{Xb}]^T\underline{\omega}[\underline{Y}-\underline{Xb}]$

where

$$\underline{Y}-\underline{Xb}=\begin{bmatrix} Y_1-(b_1X_{11} & +b_2X_{12} & +b_3X_{13}) \\ \vdots \\ Y_{n_s}-(b_1X_{n_s1} & +b_2X_{n_s2}+b_3X_{n_s3}) \\ h_b & -b_2 \end{bmatrix}$$

and $\underline{\omega}=\underline{V}^{-1}$, so that

$$\underline{V}=\begin{bmatrix} 1 & & & \underline{0} \\ & \ddots & 1 & \\ \underline{0} & & & \sigma_{h_b}^2/\sigma^2 \end{bmatrix}$$

and

$$\underline{V}^{-1}=\begin{bmatrix} 1 & & & \underline{0} \\ & \ddots & 1 & \\ \underline{0} & & & \sigma^2/\sigma_{h_b}^2 \end{bmatrix}$$

That is,

$$S(\underline{b})=\left[Y_1-(b_1X_{11}+b_2X_{12}+b_3X_{13}) \right]^2$$

$$+ \ . \ . \ .$$

$$+\left[Y_{n_s}-(b_1X_{n_s1}+b_2X_{n_s2}+b_3X_{n_s3}) \right]^2$$

$$+\sigma^2\left[h_b-b_2 \right]^2/\sigma_{h_b}^2 \ .$$

Problem 3.2-1

a. Note that if $\underline{\omega}$ is diagonal

$$\underline{X}^T\underline{\omega}\underline{X}=\left\{ \sum_{i=1}^{n_s} X_{ik}\omega_i X_{ij}+\delta_{k2}\delta_{2j}\sigma^2/\sigma_{h_b}^2 \right\}$$

$$=\left\{ \sum_{i=1}^{n_s+1} X_{ik}\omega_i X_{ij} \right\}$$

and

$$\underline{X}^T\underline{\omega}\underline{Y}=\left\{ \sum_{i=1}^{n_s} X_{ik}\omega_i Y_i+\delta_{k2}h_b\sigma^2/\sigma_{h_b}^2 \right\}$$

$$=\left\{ \sum_{i=1}^{n_s+1} X_{ik}\omega_i Y_i \right\}$$

where

$$\delta_{ij}=\begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}$$

$$\underline{X}^T=\begin{bmatrix} X_{11} & X_{21} & & X_{n_s1} & 0 \\ X_{12} & X_{22} & \cdots & X_{n_s2} & 1 \\ X_{13} & X_{23} & & X_{n_s3} & 0 \end{bmatrix}$$

$$=\begin{bmatrix} X_{11} & X_{21} & X_{n_s1} & X_{n_s+1,1} \\ X_{12} & X_{22} \cdots X_{n_s2} & X_{n_s+1,2} \\ X_{13} & X_{23} & X_{n_s3} & X_{n_s+1,3} \end{bmatrix}$$

$$\underline{\omega}=\begin{bmatrix} 1 & & & \underline{0} \\ & \ddots & 1 & \\ \underline{0} & & & \sigma^2/\sigma_{h_b}^2 \end{bmatrix}$$

$$=\begin{bmatrix} \omega_1 & & & \underline{0} \\ & \ddots & & \\ & & \omega_{n_s} & \\ \underline{0} & & & \omega_{n_s+1} \end{bmatrix}$$

and

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_{n_s} \\ h_b \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_{n_s} \\ Y_{n_{s+1}} \end{bmatrix} .$$

Thus, the normal equations

$$\underline{X}^T \underline{\omega} \underline{X} \hat{\underline{b}} = \underline{X}^T \underline{\omega} \underline{Y}$$

can be written

$$\sum_{j=1}^{p} \sum_{i=1}^{n_s+1} X_{ik}\omega_i X_{ij}\, \hat{b}_j = \sum_{i=1}^{n_s+1} X_{ik}\omega_i Y_i, \quad k=1,2,...,p$$

or

$$\sum_{i=1}^{n_s+1} \left(\sum_{j=1}^{p} X_{ik}\omega_i X_{ij}\, \hat{b}_j \right) = \sum_{i=1}^{n_s+1} X_{ik}\omega_i Y_i, \quad k=1,2,...,p.$$

As a comparison, the normal equations may be derived without using matrix techniques. Let $n=n_s+1$ so that

$$S(\underline{b}) = \sum_{i=1}^{n} e_i^2 \omega_i .$$

The model equations are given as

$$Y_i = \sum_{j=1}^{p} X_{ij} b_j + e_i, \quad i=1,2,...,n .$$

Therefore

$$S(\underline{b}) = \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right)^2 \omega_i .$$

Take the derivative with respect to any parameter b_k $(k=1,2,...,p)$:

$$\frac{\partial S}{\partial b_k} = \frac{\partial}{\partial b_k} \left[\sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right)^2 \omega_i \right] .$$

Because the derivative of a sum is the sum of derivatives, look at one term:

$$\frac{\partial}{\partial b_k} \left[\left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right)^2 \omega_i \right]$$

$$= -2 \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right) \omega_i X_{ik} .$$

Thus,

$$\frac{\partial S}{\partial b_k} = -2 \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right) \omega_i X_{ik}$$

Set $\dfrac{\partial S}{\partial b_k} = 0$ to find the minimum so that

$$\sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} \hat{b}_j \right) \omega_i X_{ik} = 0, \quad k=1,2,...,p$$

or

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{p} X_{ik}\omega_i X_{ij}\, \hat{b}_j \right) = \sum_{i=1}^{n} X_{ik}\omega_i Y_i, \quad k=1,2,...,p.$$

b.   Consider matrix $\underline{A}$ such that

$$\underline{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} .$$

Assume $\underline{A}$ is symmetric; its cofactor matrix $\underline{A}_c$ is

$$\underline{A}_c = \begin{bmatrix} (a_{22}a_{33}-a_{23})^2 & (a_{13}a_{23}-a_{12}a_{33}) & (a_{12}a_{23}-a_{13}a_{22}) \\ (a_{13}a_{23}-a_{12}a_{33}) & (a_{11}a_{33}-a_{13})^2 & (a_{12}a_{13}-a_{11}a_{23}) \\ (a_{12}a_{23}-a_{13}a_{22}) & (a_{12}a_{13}-a_{11}a_{23}) & (a_{11}a_{22}-a_{12})^2 \end{bmatrix} .$$

The determinant of $\underline{A}$ is

$$|\underline{A}|=a_{11}(a_{22}a_{33}-a_{23}^2)+a_{12}(a_{13}a_{23}-a_{12}a_{33})$$
$$+a_{13}(a_{12}a_{23}-a_{13}a_{22}) \ .$$

The inverse of $\underline{A}$ is

$$\underline{A}^{-1}=\underline{A}_c^T/|\underline{A}|$$

Let $\underline{A}=\underline{X}^T\underline{\omega}\underline{X}$ where

$$\underline{\omega}=\begin{bmatrix} 1 & & & \\ & \ddots & 1 & \underline{0} \\ \underline{0} & & & \sigma^2/\sigma_{h_b}^2 \end{bmatrix} \ .$$

Thus by adding the prior information to the information in table 3, there results:

	Data Set 1	Data Set 2
$\sum_j X_{j2}^2 \omega_j$	3.53161157	3.12700831
$\sum_j X_{j2}\omega_j Y_j$	127.0992273	122.5185789

and, for data set 1,

$$\underline{X}^T\underline{\omega}\underline{X}=\begin{bmatrix} 3.3250 & 1.6750 & 418{,}750 \\ & 3.53161157 & 418{,}750 \\ \text{symmetric} & & 83{,}340{,}625{,}000 \end{bmatrix}$$

or, for data set 2,

$$\underline{X}^T\underline{\omega}\underline{X}=\begin{bmatrix} 2.8500 & 1.6500 & 412{,}500 \\ & 3.12700831 & 412{,}500 \\ \text{symmetric} & & 8.3325\times10^{10} \end{bmatrix} \ .$$

Then, for data set 1,

$$\underline{A}_c=\begin{bmatrix} 1.18975153\times10^{11} & 3.575601563\times10^{10} & -777{,}456.0949 \\ & 1.017560156\times10^{11} & -690{,}937.5 \\ \text{symmetric} & & 8.93698347 \end{bmatrix}$$

and

$$|\underline{X}^T\underline{\omega}\underline{X}|=1.299239702\times10^{11} \ .$$

Therefore

$$[\underline{X}^T\underline{\omega}\underline{X}]^{-1}=\begin{bmatrix} 0.9157290438 & 0.2752072275 & -5.983931169\times10^{-6} \\ & 0.7831966299 & -5.318014058\times10^{-6} \\ \text{symmetric} & & 6.878625596\times10^{-11} \end{bmatrix} \ .$$

Similarly, for data set 2

$$\underline{A}_c=\begin{bmatrix} 9.040171743\times10^{10} & 3.267\times10^{10} & -609{,}265.9279 \\ & 6.732\times10^{10} & -495{,}000 \\ \text{symmetric} & & 6.189473684 \end{bmatrix}$$

and

$$|\underline{X}^T\underline{\omega}\underline{X}|=6.022819942\times10^{10} \ .$$

Therefore

$$[\underline{X}^T\underline{\omega}\underline{X}]^{-1}=\begin{bmatrix} 1.500986553 & 0.5424369368 & -1.011595787\times10^{-5} \\ & 1.117748839 & -8.218741466\times10^{-6} \\ \text{symmetric} & & 1.027670384\times10^{-10} \end{bmatrix} \ .$$

c. By adding the prior information to the information in table 3, we obtain for $\underline{X}^T \underline{\omega} \underline{Y}$:

	Data Set 1	Data Set 2
$\sum_j X_{j1} \omega_j Y_j$	192.18350	168.2030
$\sum_j X_{j2} \omega_j Y_j$	127.0992273	122.5185789
$\sum_j X_{j3} \omega_j Y_j$	26,879,687.5	26,583,550

By evaluating $(\underline{X}^T \underline{\omega} \underline{X})^{-1} \underline{X}^T \underline{\omega} \underline{Y}$, estimates $\hat{\underline{b}}$ are obtained as:

	Data Set 1	Data Set 2
$\hat{b}_1$	50.12043881	50.01097198
$\hat{b}_2$	9.487418691	9.701194703
$\hat{b}_3$	0.00002302475114	0.00002342971729

Problem 3.3–1

The answers to parts a, b, and c are found in the section at the end of the problem where aids in debugging the computer code are given. The authors' computer code and output are:

```
      DIMENSION T(10),S(10),F(10),Z(2,10),C(2,2),G(2),D(2)
   10 FORMAT (I5,6F10.0,I5)
   20 FORMAT (1H1,35HNO. OF OBSERVATIONS (N) -------- = ,I7
     $/1H ,35HPUMPING RATE (Q) ---------------- = ,G11.5
     $/1H ,35HDISTANCE FROM WELL CENTER (R) -- = ,G11.5
     $/1H ,35HINITIAL TRANSMISSIVITY (TO) ---- = ,G11.5
     $/1H ,35HINITIAL STORAGE COEFFICIENT (SO) = ,G11.5
     $/1H ,35HDAMPING PARAMETER (AP) --------- = ,G11.5
     $/1H ,35HCLOSURE CRITERION (ER) --------- = ,G11.5
     $/1H ,35HMAXIMUM NO. OF ITERATIONS (ITMX) = ,I7)
   30 FORMAT (8F10.0)
   32 FORMAT (10HOTIMES (T))
   34 FORMAT (23HOOBSERVED DRAWDOWNS (S))
   36 FORMAT (15HOITERATION NO. ,I4
     $/36H CURRENT ESTIMATES OF PARAMETERS (B))
   38 FORMAT (28HOSOLUTION FAILED TO CONVERGE)
   40 FORMAT ((1H ,10(G11.5,2X)))
   42 FORMAT (19HOSOLUTION CONVERGED)
   44 FORMAT (29HOFINAL COMPUTED DRAWDOWNS (F))
   46 FORMAT (16HORESIDUALS (F-S))
   48 FORMAT (18HOERROR VARIANCE = ,G11.5)
   50 FORMAT (12HOVAR(T)   = ,G11.5/12H COV(T,S) = ,G11.5
     $/12H VAR(S)   = ,G11.5)
   52 FORMAT (26HOSCALED SENSITIVITIES (Z):/5H TO T)
   54 FORMAT (5H TO S)
C         READ AND PRINT INPUT DATA
      READ(5,10) N,Q,R,TO,SO,AP,ER,ITMX
      WRITE(6,20) N,Q,R,TO,SO,AP,ER,ITMX
      READ(5,30) (T(I),I=1,N)
      WRITE(6,32)
      WRITE(6,40) (T(I),I=1,N)
      READ(5,30) (S(I),I=1,N)
      WRITE(6,34)
      WRITE(6,40) (S(I),I=1,N)
      FU=R*R/4.
      FW=Q/12.5664
      DMAX=ER+1.
      DO 140 KNT=1,ITMX
      UTMP=FU*SO/TO
      WTMP=FW/TO
      DO 80 I=1,N
C        COMPUTE NEW DRAWDOWN (F(I))
C          COMPUTE U AND W(U) FIRST
      U=UTMP/T(I)
      WU=W(U)
C          THEN F(I)
      F(I)=WTMP*WU
C        COMPUTE SCALED SENSITIVITIES (Z(I,J))
      TMP=EXP(-U)
      Z(1,I)=WTMP*(TMP-WU)
   80 Z(2,I)=-WTMP*TMP
C        CHECK FOR CONVERGENCE
      IF(DMAX.LT.ER) GO TO 150
C        ASSEMBLE COEFFICIENT MATRIX (C(I,J)) AND
C        GRADIENT VECTOR (G(J))
      DO 86 J=1,2
      DO 84 I=J,2
      C(I,J)=0.
   84 C(J,I)=0.
   86 G(J)=0.
      DO 110 K=1,N
      TMP=S(K)-F(K)
      DO 100 J=1,2
      DO 90 I=J,2
      C(I,J)=Z(I,K)*Z(J,K)+C(I,J)
```

```
   90 C(J,I)=C(I,J)
  100 G(J)=Z(J,K)*TMP+G(J)
  110 CONTINUE
C          INVERT COEFFICIENT MATRIX
      DET=C(1,1)*C(2,2)-C(1,2)*C(2,1)
      TMP=C(1,1)
      C(1,1)=C(2,2)/DET
      C(2,2)=TMP/DET
      C(1,2)=-C(1,2)/DET
      C(2,1)=C(1,2)
C          COMPUTE PARAMETER DISPLACEMENTS (D(I)), MAX. DISPLACEMENT
C          (DMAX), AND PARAMETERS (TO AND SO)
      DMAX=0.
      DO 130 J=1,2
      D(J)=C(1,J)*G(1)+C(2,J)*G(2)
      TMP=ABS(D(J))
      IF(TMP.GT.DMAX) DMAX=TMP
  130 CONTINUE
      TO=TO*(1.+AP*D(1))
      SO=SO*(1.+AP*D(2))
C          PRINT PARAMETERS
      WRITE(6,36) KNT
      WRITE(6,40) TO,SO
  140 CONTINUE
      WRITE(6,38)
      GO TO 160
  150 WRITE(6,42)
C          PRINT DRAWDOWNS, RESIDUALS (F(I)-S(I)), AND SCALED
C          SENSITIVITIES
  160 WRITE(6,44)
      WRITE(6,40) (F(I),I=1,N)
      DO 170 I=1,N
  170 F(I)=F(I)-S(I)
      WRITE(6,46)
      WRITE(6,40) (F(I),I=1,N)
      WRITE(6,52)
      WRITE(6,40) (Z(1,I),I=1,N)
      WRITE(6,54)
      WRITE(6,40) (Z(2,I),I=1,N)
C          COMPUTE AND PRINT ERROR VARIANCE (VAR) AND
C          COVARIANCE MATRIX FOR PARAMETERS
      VAR=0.
      DO 180 I=1,N
  180 VAR=VAR+F(I)*F(I)
      VAR=VAR/(N-2.)
      WRITE(6,48) VAR
      C(1,1)=TO*C(1,1)*TO*VAR
      C(1,2)=TO*C(1,2)*SO*VAR
      C(2,2)=SO*C(2,2)*SO*VAR
      WRITE(6,50) C(1,1),C(1,2),C(2,2)
      STOP
      END
      FUNCTION W(X)
C          COMPUTE THE WELL FUNCTION OF X
      W=0.
      IF(X.GT.10.) GO TO 20
      W=-0.577216-ALOG(X)+X
      TERM=X
      DO 10 J=2,36
      RJ=J
      TERM=-TERM*X/RJ
      TMP=TERM/RJ
      W=W+TMP
      IF(ABS(TMP).LT.1.E-7) GO TO 20
   10 CONTINUE
   20 RETURN
      END
```

```
1NO. OF OBSERVATIONS (N) -------- =        7
 PUMPING RATE (Q) --------------- =   1.1600
 DISTANCE FROM WELL CENTER (R) -- =   175.00
 INITIAL TRANSMISSIVITY (TO) ---- =   .10000
 INITIAL STORAGE COEFFICIENT (SO) =   .50000E-03
 DAMPING PARAMETER (AP) --------- =   1.0000
 CLOSURE CRITERION (ER) --------- =   .10000E-01
 MAXIMUM NO. OF ITERATIONS (ITMX) =       10
OTIMES (T)
    480.00        1020.0        1500.0        2040.0        2700.0        3720.0        4920.0
OOBSERVED DRAWDOWNS (S)
    1.7100        2.2300        2.5400        2.7700        3.0400        3.2500        3.5600
OITERATION NO.    1
 CURRENT ESTIMATES OF PARAMETERS (B)
    .11188        .54748E-03
OITERATION NO.    2
 CURRENT ESTIMATES OF PARAMETERS (B)
    .11347        .55219E-03
OITERATION NO.    3
 CURRENT ESTIMATES OF PARAMETERS (B)
    .11349        .55221E-03
OSOLUTION CONVERGED
OFINAL COMPUTED DRAWDOWNS (F)
    1.6715        2.2521        2.5564        2.8012        3.0256        3.2832        3.5086
ORESIDUALS (F-S)
  -.38538E-01    .22079E-01    .16407E-01    .31217E-01   -.14393E-01    .33213E-01   -.51356E-01
OSCALED SENSITIVITIES (Z):
 TO T
  -.91882       -1.4679       -1.7630       -2.0026       -2.2234       -2.4779       -2.7014
 TO S
  -.75264       -.78421       -.79343       -.79866       -.80223       -.80527       -.80724
OERROR VARIANCE =  .14328E-02
OVAR(T)    =  .95030E-05
 COV(T,S)  = -.11369E-06
 VAR(S)    =  .14595E-08
```

Problem 3.3-2

b.

$$
\underline{D}=\begin{bmatrix}
2T_1 & -T_1 & & -T_1 & & & & & \\
-T_1 & 2(T_1+T_2) & -T_2 & & -(T_1+T_2) & & & & \\
 & -T_2 & 2T_2 & & & -T_2 & & & \\
-T_1 & & & 4T_1 & -2T_1 & & & & \\
 & -(T_1+T_2) & & -2T_1 & 4(T_1+T_2) & -2T_2 & -(T_1+T_2 & & \\
 & & -T_2 & & -2T_2 & 4T_2 & & & -T_2 \\
 & & & & -(T_1+T_2) & & 2(T_1+T_2) & -T_2 \\
 & & & & & -T_2 & -T_2 & 2T_2
\end{bmatrix}
$$

$$
\underline{h}=\begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix}
\qquad
\underline{q}=\begin{bmatrix}
\tfrac{1}{2}a^2 W_1 \\
\tfrac{1}{2}a^2(W_1+W_2) \\
\tfrac{1}{2}a^2 W_2 \\
a^2 W_1+T_1 h_{B1} \\
a^2(W_1+W_2) \\
a^2 W_2+a q_{B1} \\
h_{B1} \\
\tfrac{1}{2}a^2(W_1+W_2)+T_1 h_{B1} \\
\tfrac{1}{2}a^2 W_2+a q_{B1}
\end{bmatrix}
$$

c.

$$J_1 = \begin{vmatrix} h_2+h_4-2h_1 \\ h_1+h_5-2h_2 \\ 0 \\ h_1+2h_5-4h_4+h_{B1} \\ h_2+2h_4+h_8-4h_5 \\ 0 \\ 0 \\ h_5-2h_8+h_{B1} \\ 0 \end{vmatrix}$$

$$J_2 = \begin{vmatrix} 0 \\ h_3+h_5-2h_2 \\ h_2+h_6-2h_3 \\ 0 \\ h_2+2h_6+h_8-4h_5 \\ h_3+2h_5+h_9-4h_6 \\ 0 \\ h_5+h_9-2h_8 \\ h_6+h_8-2h_9 \end{vmatrix}$$

$$J_3 = \begin{vmatrix} \tfrac{1}{2}a^2 \\ \tfrac{1}{2}a^2 \\ 0 \\ a^2 \\ a^2 \\ 0 \\ 0 \\ \tfrac{1}{2}a^2 \\ 0 \end{vmatrix} \qquad J_4 = \begin{vmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ a \\ 0 \\ 0 \\ a \end{vmatrix}$$

d.

1. Let $r=0$.
2. Compute $\underline{D}_r$ and $\underline{q}_r$.
3. $\underline{h}_r = \underline{D}_r^{-1}\,\underline{q}_r$.
4. Obtain $\underline{f}(\underline{\xi},\underline{b}_r)$ from $\underline{h}_r$ by deleting node 7.
5.

$$\left.\left(\frac{\partial \underline{h}}{\partial b_j}\right)\right|_r = \underline{D}_r^{-1}\left[\left(\frac{\partial \underline{q}}{\partial b_j}\right)_r - \left(\frac{\partial \underline{D}}{\partial b_j}\right)\underline{h}_r\right]\Bigg\} \; j=1,2,...,p$$

Obtain $\underline{X}_j^r$ from $\left(\dfrac{\partial \underline{h}}{\partial b_j}\right)_r$ by deleting node 7.

6. Compute $\underline{X}_r^T\underline{\omega}\underline{X}_r$ and $\underline{X}_r^T\underline{\omega}(\underline{Y}-\underline{f}(\underline{\xi},\underline{b}_r))$.
7. Define $\underline{C}_r = \{C_{ii}^r\} = \{(\underline{X}_r^T\underline{\omega}\underline{X}_r)_{ii}^{-\frac{1}{2}}\}$.
8. Compute $\underline{S}_r^T\underline{\omega}\underline{S}_r$ and $\underline{S}_r^T\underline{\omega}(\underline{Y}-\underline{f}(\underline{\xi},\underline{b}_r))$
   where $\underline{S}_r = \underline{X}_r\underline{C}_r$.
9. $\underline{\delta}_{r+1} = (\underline{S}_r^T\underline{\omega}\underline{S}_r + \mu\underline{I})^{-1}\underline{S}_r^T\underline{\omega}(\underline{Y}-\underline{f}(\underline{\xi},\underline{b}_r))$
10. $\underline{d}_{r+1} = \underline{C}_r\underline{\delta}_{r+1}$.
11. $\underline{b}_{r+1} = \rho\underline{d}_{r+1} + \underline{b}_r$.
12. If $|d_i^{r+1}/c| > \epsilon$ (where $c = b_i^r$ for $b_i^r \ne 0$, and $c=1$ for $b_i^r=0$) for any $i=1,2,...,p$, then increment r by one and return to 2. If not then:
13. $\underline{h}_{\text{final}} = \underline{D}_{r+1}^{-1}\underline{q}_{r+1}$.

Values of μ and ρ can be computed at each iteration by using algorithms defined by equations 3.3–28 through 3.3–30 if desired.

Problem 4.2-1

Data Set 1

Nodes in columns 2 through 11 have observations. Nodes in columns 1 and 12 form specified head boundaries. Spacing: Cell row 1, 1 ft; cell columns 1 and 11, 50 ft; cell columns 2 through 10, 100 ft (figure 1).

Data Set 2.

Nodes in columns 2 through 10 have observations. Nodes in columns 1 and 11 form specified head boundaries. Spacing: Cell row 1, 1 ft; cell columns 1 through 10, 100 ft (figure 2). The input data to the regression program for the two data sets are shown in figures 3 and 4.

If both T and W were estimated, the problem would be singular because the only unique parameter is W/T.
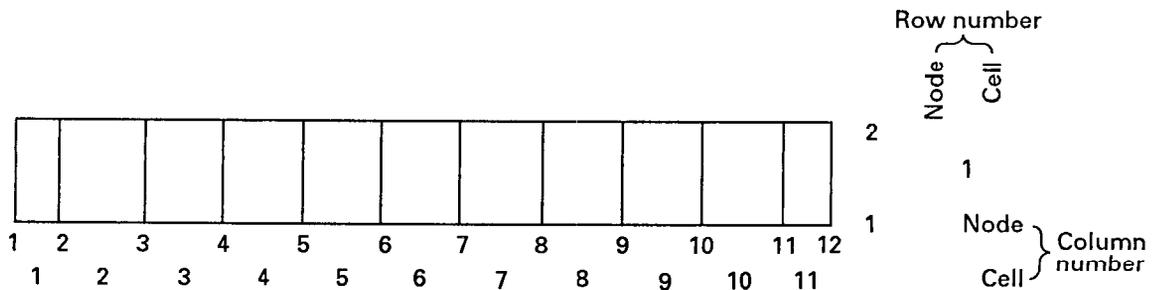

Figure 1

Row number

Node } Cell

```
┌───┬───┬───┬───┬───┬───┬───┬───┬───┬───┐   2
│   │   │   │   │   │   │   │   │   │   │
│   │   │   │   │   │   │   │   │   │   │      1
└───┴───┴───┴───┴───┴───┴───┴───┴───┴───┘
1   2   3   4   5   6   7   8   9   10  11    1
  1   2   3   4   5   6   7   8   9   10
```

Node } Column
Cell } number

Figure 2

LAKE OHPUPU, DATA SET 1.

12	2	1	10	1	3	0	0	2	2	1	1	0

50		0		0		0		.25

DX   1   0

1	11	1	1		1	1
50		100		100		100		100		100		100		100
100		100		50

DY   1   0

1	1	1	1		1

CX   1   0

1	11	1	1		1

CY   1   0

1	11	1	1		1

VL   1   0

1	11	1	1		0

HR   1   0

1	12	1	2		0

QR   1   0

1	11	1	1		1

HC   1   0

2	11	1	2		11

IZN   1   0

1	11	1	1	1

1	2	1	50	.5	48.33	1
2	3	1	150	.5	45.76	1
3	4	1	250	.5	42.08	1
4	5	1	350	.5	38.34	1
5	6	1	450	.5	35.30	1
6	7	1	550	.5	31.00	1
7	8	1	650	.5	25.85	1
8	9	1	750	.5	21.76	1
9	10	1	850	.5	16.11	1
10	11	1	950	.5	12.48	1

1	0	0	0	1
1		0
1		1		1		0		.001

IN   2   0

1	1	1	2	−1
12	12	1	2	−1
1	2	2	2		0		0
1	1		49
1	2		49
2	2	3	3		1.1		1.1
12	1		11
12	2		11

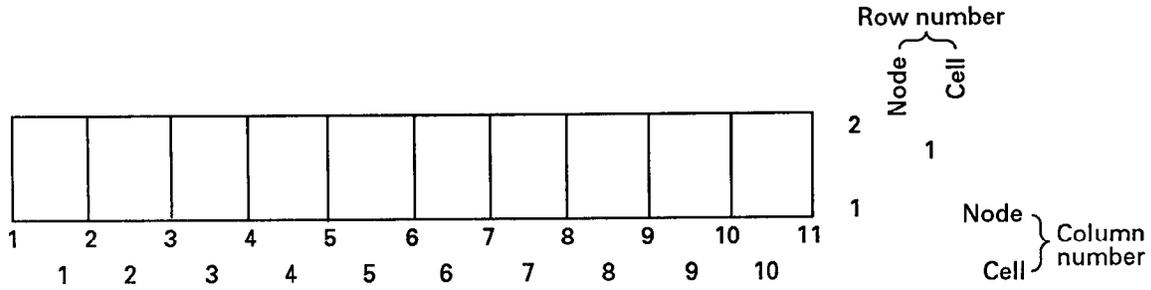**Figure 3**

LAKE OHPUPU, DATA SET 2.

```
      11    2    1    9    1    3    0    0    2    2    1    1    0
            50             0         0              0        .25
DX           1    0
       1    10   1    1         100
DY           1    0
       1     1   1    1           1
CX           1    0
       1    10   1    1           1
CY           1    0
       1    10   1    1           1
VL           1    0
       1    10   1    1           0
HR           1    0
       1    11   1    2           0
QR           1    0
       1    10   1    1           1
HC           1    0
       2    10   1    2          10
IZN          1    0
       1    10   1    1    1
       1     2   1         100      .5      47.13          1
       2     3   1         200      .5      44.14          1
       3     4   1         300      .5      39.89          1
       4     5   1         400      .5      36.36          1
       5     6   1         500      .5      32.48          1
       6     7   1         600      .5      29.70          1
       7     8   1         700      .5      24.33          1
       8     9   1         800      .5      19.10          1
       9    10   1         900      .5      14.96          1
       1     0   0    0    1
       1          0
       1          1         1         1       .001
IN           2    0
       1     1   1    2   -1
      11    11   1    2   -1
       1     2   2    2           0         0
       1     1       49
       1     2       49
       2     2   3    3         .95       .95
      11     1      9.5
      11     2      9.5
```

Figure 4

Problem 4.2-2

Because prior information is available, all parameters can be estimated. Data points are located in areas of relatively high sensitivity for all parameters. More data points in areas of highest sensitivity might improve results for parameters having low sensitivity. The input data to the regression program are shown in figure 1.

```
        CLASS  PROBLEM
        TWO-DIMENSIONAL  FLOW
        SEVERAL  ZONES
    15  16    4   32   7   14    0    4    2   10    1    1    0
         2        .08       0             0         1
DX       1    0
     1  14    1    1        1    1
        1000       1000       1000       1000       1000       1000       1000       400
        1000       1000       1000       1000       1000       1000
DY       1    0
     1  15    1    1        1    1
        1000       1000       1000       1000       1000        400       1000      1000
        1000       1000       1000       1000       1000       1000
CX       2    0
     6  14    1    8        1    0
     1  14    9   15        1    0
CY       2    0
     6  14    1    8        1    0
     1  14    9   15        1    0
VL       2    0
     8  14    6    6        1    0
     8   8    7   15        1    0
HR       2    0
     8  15    6    7      4.5    0
     8   9    7   16      4.5    0
QR       2    0
     6  14    1    8        1    0
     1  14    9   15        1    0
HC       1    0
     1  15    1   16        0    0
IZN      7    0
     6  14    1    8    2    0
     1  14    9   15    2    0
     1   6   12   15    1    0
     6   7    1    3    3    0
     8  14    1    4    3    0
     8  14    6    6    4    0
     8   8    7   15    4    0
     1   8    2      7000      1000    60.70        1
     2  14    2     12400      1000    75.64        1
     3  12    3     10400      2000    60.27        1
     4  10    4      8400      3000    29.67        1
     5   7    5      6000      4000     4.22        1
     6  11    5      9400      4000     4.37        1
     7  13    5     11400      4000     6.07        1
     8  14    5     13400      4000     5.81        1
     9  10    7      8400      5400     4.57        1
    10   8    8      7000      6400     5.21        1
    11  12    8     10400      6400   -44.89        1
    12  14    8     13400      6400     7.01        1
    13   7    9      6000      7400     6.95        1
    14   4   10      3000      8400    12.21        1
    15   9   10      7400      8400     4.04        1
```

Figure 1

```
     16  11  10      9400      8400    -89.36        1
     17   7  11      6000      9400      6.68        1
     18  13  11     11400      9400    -15.32        1
     19   3  12      2000     10400     16.88        1
     20   5  12      4000     10400     15.87        1
     21   9  12      7400     10400      4.48        1
     22  11  12      9400     10400    -18.34        1
     23  13  13     11400     11400     -2.47        1
     24  14  13     13400     11400      8.10        1
     25   3  14      2000     12400     54.12        1
     26   5  14      4000     12400     38.27        1
     27  10  14      8400     12400      .053        1
     28  12  14     10400     12400     -2.92        1
     29   7  15      6000     13400      8.30        1
     30  14  15     12400     13400      4.54        1
     31   2  15      1000     14400     85.82        1
     32  11  15      9400     14400      2.26        1
      1   8   8   0    9
      2  10  10   0   11
      3  12  12   0   13
      4  10  10  14    0
      8       0
      9   .00012
     10      84
     11       0
     12       0
     13  .000051
     14     .008
      1      70       70        0    .0004
      2     420      420        0   -.0002
      3      15       15        0   .00017
      4     420      420      .08        0
      1  16   7  16   1       45        0      .01
      7  16  15  16   2      420        0   .00056
     11  10  11  10   3   -97000     1940        1
     12   8  12   8   4   -51000     1020        1
IN    1   0
     15  15   5  16  -1    0
      1  10   5   6       1.04              .48
     15  16      10.4
     15  15
     15  14
     15  13
     15  12
     15  11
     15  10
     15   9
     15   8
     15   7       4.8
      2   2   6   7          .48              .54
     15   6       4.8
     15   5       5.4
```

Figure 1

Problem 5.4-1

a. Let $\underline{b}_0 = \underline{0}$. Then

$$s^2 = \frac{\underline{Y}^T \underline{\omega} \underline{Y} - \hat{\underline{b}}^T \underline{X}^T \underline{\omega} \underline{Y}}{n-p}$$

$$\underline{Y} = \begin{bmatrix} 48.33 \\ 45.76 \\ 42.08 \\ 38.34 \\ 35.30 \\ 31.00 \\ 25.85 \\ 21.76 \\ 16.11 \\ 12.48 \\ 11 \end{bmatrix} \text{ for data set 1, or } \begin{bmatrix} 47.13 \\ 44.14 \\ 39.89 \\ 36.36 \\ 32.48 \\ 29.70 \\ 24.33 \\ 19.10 \\ 14.96 \\ 9.5 \end{bmatrix} \text{ for data set 2}$$

$$\underline{\omega} = \begin{bmatrix} 1 \\ & 1 \\ & & \cdots \\ & & & 1 \\ & & & & 0.25 \\ & & & & & \frac{1}{(1.1)^2} \end{bmatrix} \text{ or } \begin{bmatrix} 1 \\ & 1 \\ & & \cdots \\ & & & 1 \\ & & & & 0.25 \\ & & & & & \frac{1}{(0.95)^2} \end{bmatrix}$$

$$\hat{\underline{b}} = \begin{bmatrix} 50.12043881 \\ 9.487418691 \\ 0.00002302475114 \end{bmatrix} \text{ or } \begin{bmatrix} 50.01097198 \\ 9.701194703 \\ 0.00002342971729 \end{bmatrix}$$

$$\underline{X}^T \underline{\omega} \underline{Y} = \begin{bmatrix} 192.18350 \\ 127.0992273 \\ 26{,}879{,}687.5 \end{bmatrix} \text{ or } \begin{bmatrix} 168.2030 \\ 122.5185789 \\ 26{,}583{,}550 \end{bmatrix}$$

$\underline{Y}^T \underline{\omega} \underline{Y} = 11{,}459.5411$ or $10{,}225.4391$

$\hat{\underline{b}}^T \underline{X}^T \underline{\omega} \underline{Y} = 11{,}457.06305$ or $10{,}223.41717$

$s^2 = 0.30975625$ or 0.2888471429

$s/\Delta Y_s = 0.5565574993/(48.33{-}12.48) = 0.0155$ or $0.5374450138/(47.13{-}14.96) = 0.0167$

The fit is fairly good.

b.

$$(\underline{X}^T \underline{\omega} \underline{X})^{-1} s^2 =$$

$$0.30975625 \begin{bmatrix} 0.9157290438 & 0.2752072275 & -5.983931169 \times 10^{-6} \\ & 0.7831966299 & -5.318014058 \times 10^{-6} \\ \text{symmetric} & & 6.878625596 \times 10^{-11} \end{bmatrix}$$

$$= \begin{bmatrix} 0.2836527946 & 0.08524715876 & -1.853560079 \times 10^{-6} \\ & 0.2426000511 & -1.647288092 \times 10^{-6} \\ \text{symmetric} & & 2.130697270 \times 10^{-11} \end{bmatrix}$$

or

$$0.2888471429 \begin{bmatrix} 1.500986553 & 0.5424369368 & -1.011595787 \times 10^{-5} \\ & 1.117748839 & -8.218741466 \times 10^{-6} \\ \text{symmetric} & & 1.027670384 \times 10^{-10} \end{bmatrix}$$

$$= \begin{bmatrix} 0.4335556774 & 0.1566813594 & -2.921965528 \times 10^{-6} \\ & 0.3228585586 & -2.373959991 \times 10^{-6} \\ \text{symmetric} & & 2.968396543 \times 10^{-11} \end{bmatrix} .$$

Because the standard errors (the square roots of the diagonal elements) are small, the parameters are well determined.

c.

$$r = \begin{bmatrix} 1 & 0.3249682531 & -0.7539668635 \\ & 1 & -0.7245414573 \\ \text{symmetric} & & 1 \end{bmatrix}$$

or

$$\begin{bmatrix} 1 & 0.4187827075 & -0.8145011290 \\ & 1 & -0.7668426954 \\ \text{symmetric} & & 1 \end{bmatrix}.$$

The problem is fairly well conditioned.

Problem 5.5-1

The plots of $\underline{d}$ and g (figures 1,2) suggest that correlation manifests itself in two ways: The variability of g from set to set is smaller than the variability of $\underline{d}$ from set to set, and the plots of g do not appear to be linear on normal probability paper. The plot of $\hat{\underline{e}}$ (figure 3) does not appear to differ very much from either the plots of $\underline{d}$ or g, although the plot of $\hat{\underline{e}}$ does show the same nonlinear trend as displayed by the plots of g. Therefore, one may say that the plot of $\hat{\underline{e}}$ does not appear to differ significantly from the plot of a $N(\underline{0},(\underline{I}-\underline{R})s^2)$ random variable. Furthermore, this distribution of residuals suggests that the Theis model is adequate to describe the observed-drawdown data set. Input data for the residuals analysis program are shown in figure 4.

Problem 5.5-2

a. $s^2 = 0.98677$

$R_y = 0.99964$

$s/\Delta Y_s = 0.99336297/175.18 = 0.00567$

b. The problem is not well conditioned. Entries (1,8), (1,9), (8,9), and (12,13) (and their symmetric counterparts) of the scaled least squares matrix have absolute values greater than 0.9, although these large values yielded large (absolute value >0.9) parameter correlations only for entries (1,8) and (12,13). The correlation between

parameters 12 and 13 is exceptionally large so that these parameters are behaving as one. Because parameters 12 and 13 are the transmissivity T_3 and recharge W_3 in zone 3, respectively, the physical interpretation is that the ratio W_3/T_3 is much more unique than either W_3 or T_3. The only poorly determined parameter is the flow across boundary zone 2, q_{B2}. Because of the large value of transmissivity in the aquifer zone adjacent to this boundary flow zone, the gradient near the boundary is low. Hence, the flow across the boundary is estimated poorly. Well discharges Q_1 and Q_2 and transmissivity T_2 are estimated quite well. The drawdown cones provide large gradients that serve to determine Q_1, Q_2, and T_2 precisely.

c. Effects of correlation within g are not large. They may be exhibited as a slight steepening of the curves for g (figure 2) compared to those for $\underline{d}$ (figure 1) on the normal probability plots. The plot of $\hat{\underline{u}}$ (figure 3) is very similar to those of g, which suggests that the distribution of $\hat{\underline{u}}$ does not differ significantly from a $N(\underline{0},(\underline{I}-\underline{R})s^2)$ distribution. Note that $\hat{\underline{e}}$ could not have been used instead of $\hat{\underline{u}}$ to make the comparison because the weight matrix $\underline{\omega}$ is not equal to $\underline{I}$. The plot of $\hat{u}_j$ versus $\hat{f}_j$ (figure 4) shows no pattern. However, the plot of $\hat{e}_j$ versus Cartesian coordinate (figure 5) shows a group of negative residuals in the upper center of the area. This sign pattern was inherited from the sign pattern of the original errors $\underline{\epsilon}$ that were generated (recall that this exercise is based on a hypothetical problem), and the original errors are random $N(0,1)$ deviates. Hence, the sign pattern occurred entirely by chance. The lesson is that apparently nonrandom patterns can, and often do, develop by chance, and the analyst must learn to distinguish true problem areas from apparent ones.

The input data for the residuals analysis program are shown in figure 6.

Ordered residual distribution: (The calculated values are prior information residuals $\hat{u}_j = \hat{e}_j \omega_j^{1/2}$, where $\omega_j^{1/2} = s/(\text{Var}(\epsilon_{pj}))^{1/2}$.)

No.	$\hat{u}_j$	No.	$\hat{u}_j$
1	-2.3900	21	-0.013826
2	-1.4247	22	$(-97000+97000)5.1204\times10^{-4}= \approx 0$
3	-1.3248	23	0.034151
4	-1.2986	24	0.036783
5	-1.1685	25	$(-50961+51000)9.7389\times10^{-4}=0.037982$
6	-0.95063	26	$(0.080716-0.08)124.17=0.088906$
7	$(0.00031149-0.0004)8278.0=-0.73269$	27	$(5.4730-5.4)1.8396=0.13429$
8	-0.72189	28	0.22868
9	-0.70275	29	0.24557
10	$(0.00013516-0.00017)19478=-0.67861$	30	0.33619
11	-0.50751	31	0.40574
12	-0.43737	32	0.41665
13	-0.43737	33	0.64095
14	-0.35588	34	0.65517
15	-0.33702	35	$(5.1211-4.8)2.0695=0.66452$
16	-0.26802	36	$(487.89-420)0.011826=0.80287$
17	$(10.198-10.4)0.95516=-0.19294$	37	0.90021
18	-0.18186	38	0.98513
19	-0.087697	39	1.1503
20	-0.080833	40	1.3703
		41	1.6158

Figure 1

Figure 2

Figure 3

```
    2     7     0     5     1    .0014328
.95030E-5-.11369E-6
.14595E-8
          1          1       1         1         1         1         1
 -8.09604   -1362.96
-12.9342    -1420.13
-15.5344    -1436.83
-17.6456    -1446.30
-19.5912    -1452.76
-21.8336    -1458.27
-23.8030    -1461.84
```
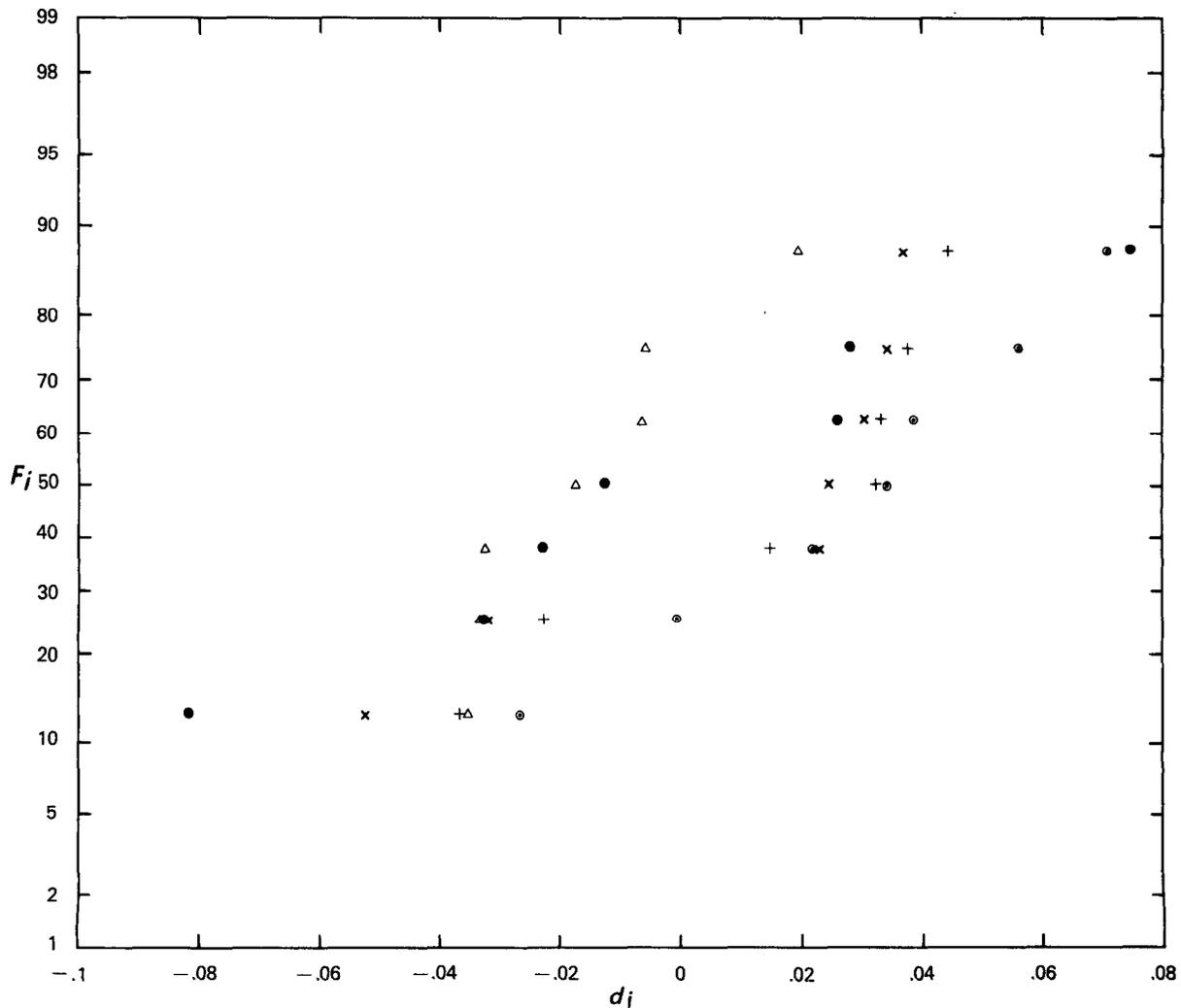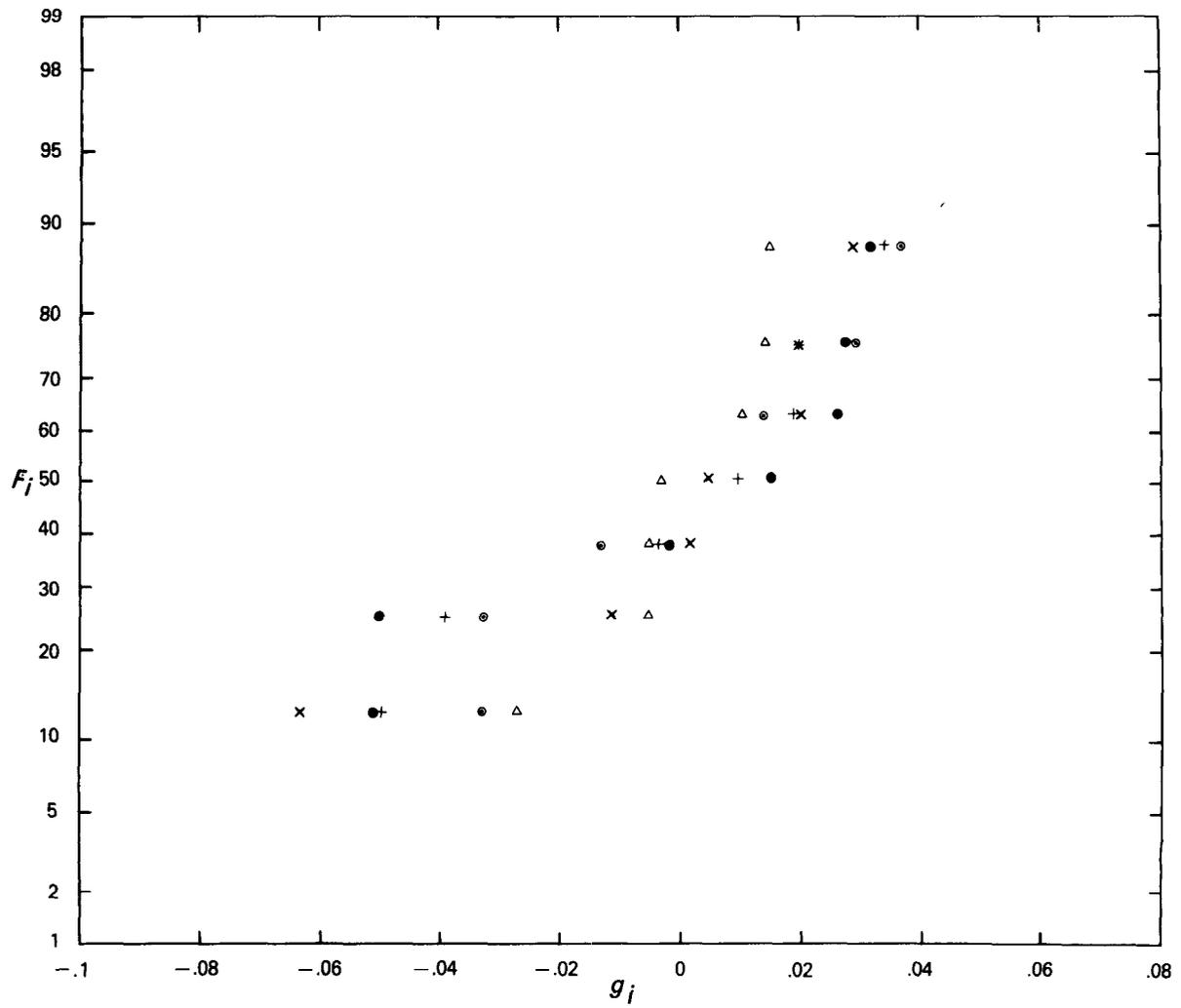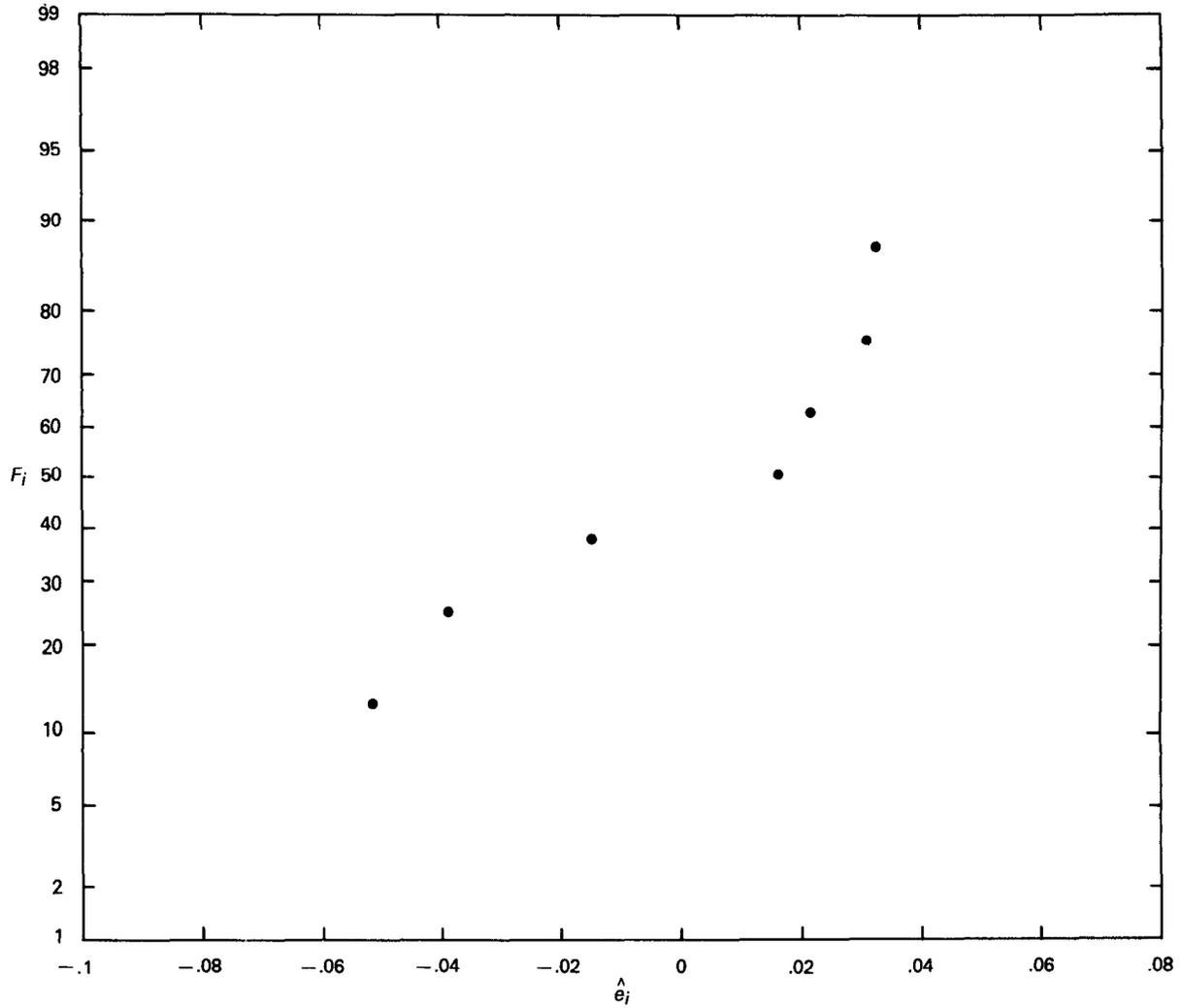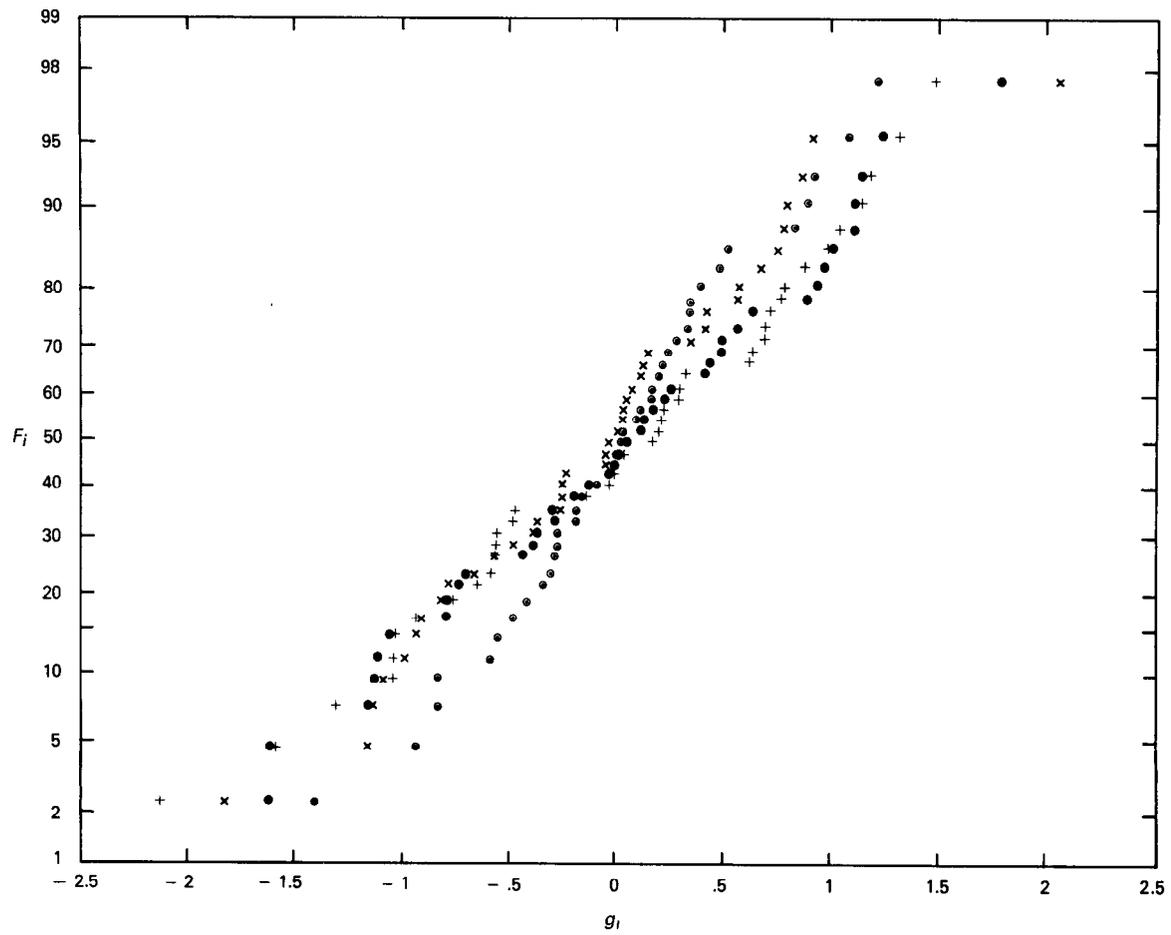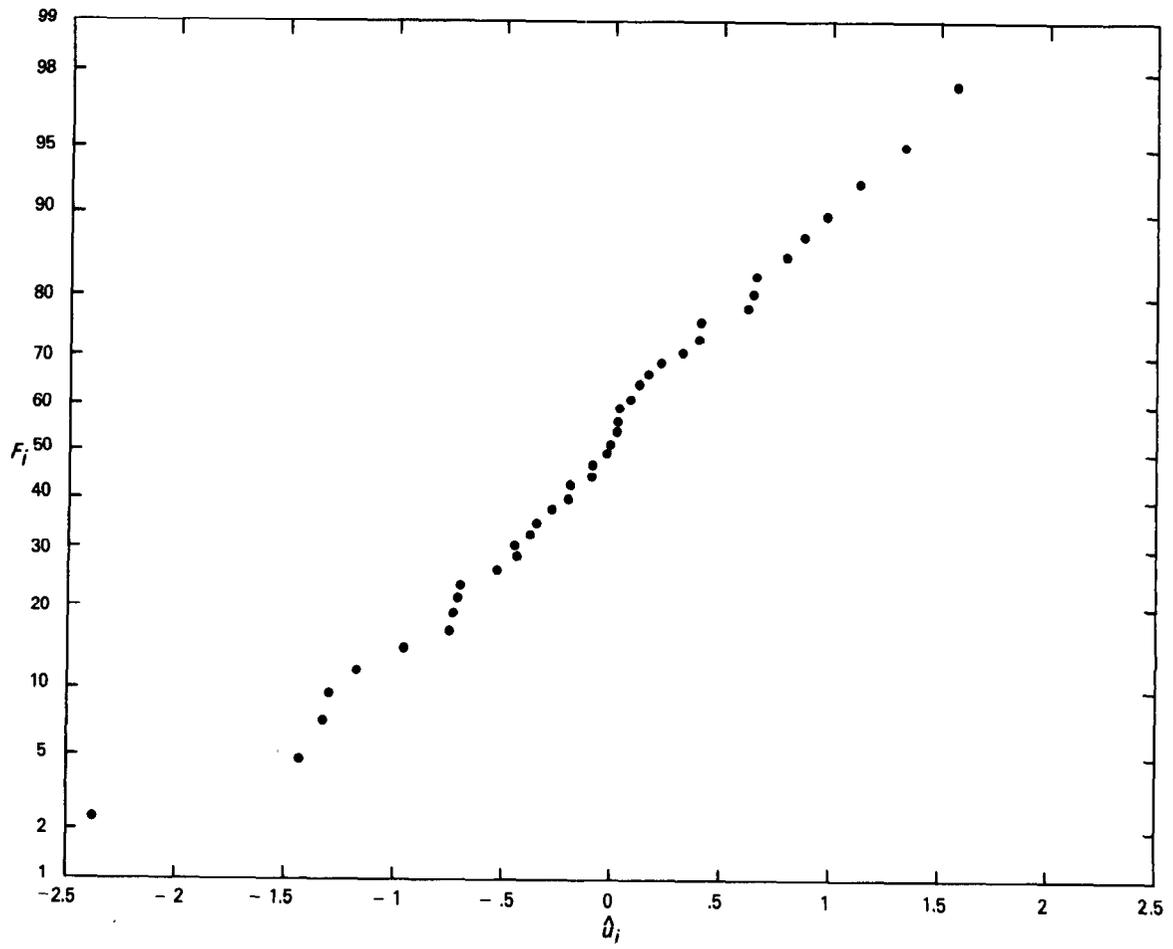
Figure 4

Figure 1

Figure 2

Figure 3
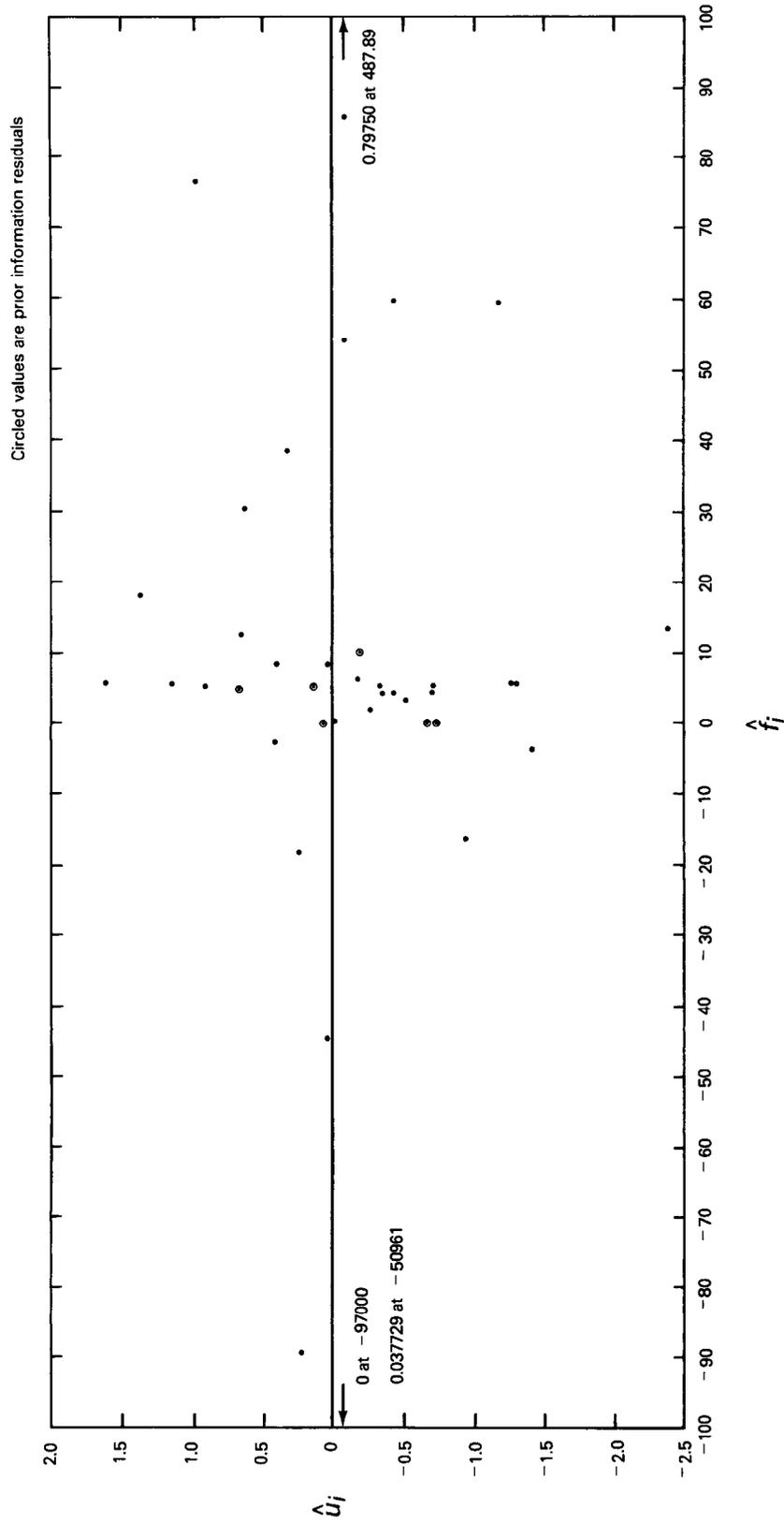
Figure 4

Figure 5

```
14    32    9    4   437      .98677
       1
 3     4    5    6    7    9   10   13   14
   1940      1020      1.04      .48      .54     .00012      84     .000051
   .008
```
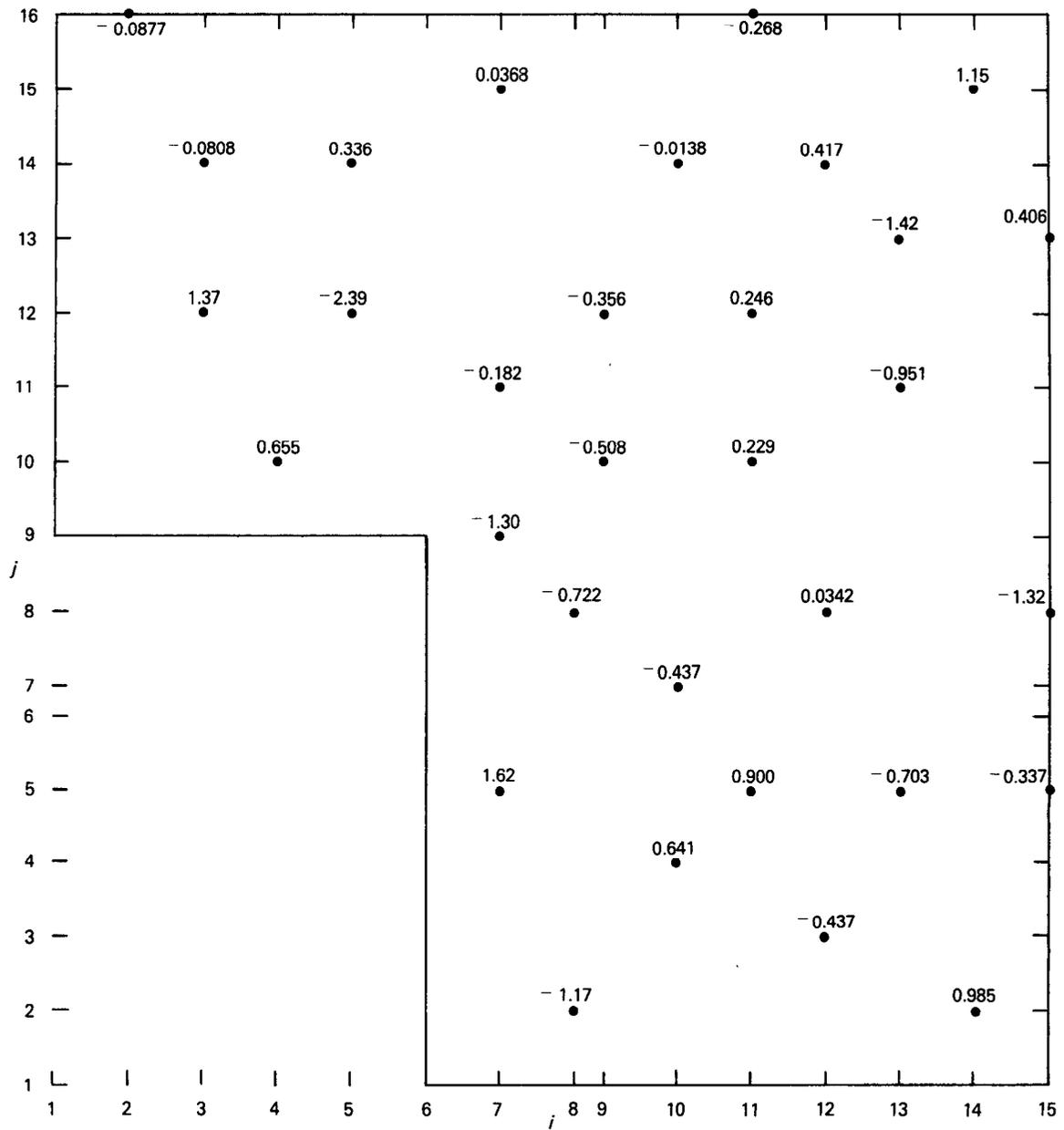
Figure 6

Problem 5.6-1

a. Compute

$$\underline{H}(\underline{X}^T\omega\underline{X})^{-1}\underline{H}^T$$

where $\underline{H}=[0\ 0\ 1]$. Then,

$$\underline{H}(\underline{X}^T\omega\underline{X})^{-1}\underline{H}^T=\widehat{\text{Var}}(\hat{b}_3)/s^2\ .$$

By using equation 5.6-11,

$$W=\frac{(\tilde{\beta}_3-\hat{b}_3)(\text{Var}(\hat{b}_3)/s^2)^{-1}(\tilde{\beta}_3-\hat{b}_3)}{s^2}$$

$$=\frac{(\tilde{\beta}_3-\hat{b}_3)^2}{\widehat{\text{Var}}(\hat{b}_3)}$$

$$H_0:\beta_3=\frac{0.0003}{10}=3\times10^{-5}$$

$$H_1:\beta_3\neq3\times10^{-5}$$

$$w=\frac{(3\times10^{-5}-2.3025\times10^{-5})^2}{2.1307\times10^{-11}}$$

$$=2.273\text{ for data set 1.}$$

$F_{0.05}(1,8)=5.318\ .\ \therefore H_0$ accepted

Values of the ratio of recharge to transmissivity, $\beta_3=W/T$, computed by the two methods are not significantly different.

$$w=\frac{(3\times10^{-5}-2.3430\times10^{-5})^2}{2.9684\times10^{-11}}$$

$$=1.454\text{ for data set 2}$$

$F_{0.05}(1,7)=5.591\ .\ \therefore H_0$ accepted.

The Maxey-Eakin estimate of W/T could be used as prior information in the regression model, but an estimate of $\text{Var}(W/T)$ would also be needed.

b. This test is the same as the one in a except that

$$H_0:\beta_3=0\ \ H_1:\beta_3\neq0$$

$$w=\frac{(0-2.3025\times10^{-5})^2}{2.1307\times10^{-11}}$$

$$=24.88\text{ for data set 1}$$

$\therefore H_0$ rejected.

$$w=\frac{(0-2.3430\times10^{-5})^2}{2.9684\times10^{-11}}$$

$$=18.49\text{ for data set 2}$$

$\therefore H_0$ rejected.

W/T is significantly different from zero. Thus, recharge is a significant variable in the regression equation.

c. $\tilde{\beta}_3=\hat{b}_3\pm\sqrt{qF_\alpha(q,n-p)}s_{b3}$

$$=2.3025\times10^{-5}\pm\sqrt{5.318}\times4.6159\times10^{-6}$$

$$=2.3025\times10^{-5}$$
$$\pm1.0645\times10^{-5}\text{ for data set 1 .}$$

$$\tilde{\beta}_3=2.3430\times10^{-5}\pm\sqrt{5.591}\times5.4483\times10^{-6}$$

$$=2.3430\times10^{-5}$$
$$\pm1.2883\times10^{-5}\text{ for data set 2.}$$

Problem 5.6-2

For T extreme:

$$T=\hat{T}\pm\frac{\sqrt{2F_{0.05}(2,5)}}{s_{bT}}\widehat{\text{Var}}(T)$$

$$=0.11349\pm\frac{\sqrt{2\times5.7861}}{0.0030827}(0.95030\times10^{-5})$$

$$=0.11349\pm0.010487$$

$$S=\hat{S}\pm\frac{\sqrt{2F_{0.05}(2,5)}}{s_{bT}}\widehat{\text{Cov}}(T,S)$$

$$=0.00055221\pm\frac{\sqrt{2\times5.7861}}{0.0030827}(-0.11369\times10^{-6})$$

$$=0.00055221\mp0.00012546.$$

For S extreme:

$$T=\hat{T}\pm\frac{\sqrt{2F_{0.05}(2,5)}}{s_{bS}}\ \widehat{\text{Cov}}(T,S)$$

$$=0.11349\pm\frac{\sqrt{2\times5.7861}}{3.8203\times10^{-5}}(-0.11369\times10^{-6})$$

$$=0.11349\mp0.010124$$

$$S=\hat{S}\pm\frac{\sqrt{2F_{0.05}(2,5)}}{s_{bS}}\ \widehat{\text{Var}}(S)$$

$$=0.00055221\pm\frac{\sqrt{2\times5.7861}}{3.8203\times10^{-5}}(0.14595\times10^{-8})$$

$$=0.00055221\pm0.00012996\ .$$

Problem 5.6-3

$$\tilde{\underline{b}}=\hat{\underline{b}}\pm\frac{\sqrt{qF_\alpha(q,n-p)}}{s_{bi}}\ \underline{V}_{bi}$$

where $\underline{V}_b=(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2$. Let $(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2=\underline{A}$ $=\{A_{ij}\}$. Then the above equation can be written in algebraic form for calculations as

$$\tilde{b}_j=\hat{b}_j\pm\sqrt{qF_\alpha(q,n-p)}\ \frac{A_{ij}}{\sqrt{A_{ii}}}\ .$$

Calculations:

$$\sqrt{qF_\alpha(q,n-p)}=\sqrt{2F_{0.05}(2,27)}=\sqrt{2\times3.359}=2.592$$

For $T_3=\beta_{12}$ extreme:

$\tilde{b}_1=80.978\pm9.2811$
$\tilde{b}_2=935.15\pm208.42$
$\tilde{b}_3=-97000\mp85.314$
$\tilde{b}_4=-50961\mp29.377$
$\tilde{b}_5=10.198\pm0.11520$
$\tilde{b}_6=5.1211\pm1.3758\times10^{-2}$
$\tilde{b}_7=5.4730\pm1.2758\times10^{-2}$
$\tilde{b}_8=65.754\pm7.2341$
$\tilde{b}_9=3.1149\times10^{-4}\pm3.0988\times10^{-5}$

$\tilde{b}_{10}=487.89\pm2.5660$
$\tilde{b}_{11}=-1.3995\times10^{-4}\mp4.9215\times10^{-5}$
$\tilde{b}_{12}=13.288\pm9.3998$
$\tilde{b}_{13}=1.3516\times10^{-4}\pm9.3069\times10^{-5}$
$\tilde{b}_{14}=8.0716\times10^{-2}\pm3.5370\times10^{-5}$

For $q_{B1}=\beta_1$ extreme:

$\tilde{b}_1=80.978\pm43.911$
$\tilde{b}_2=935.15\pm489.14$
$\tilde{b}_3=-97000\mp222.36$
$\tilde{b}_4=-50961\mp73.632$
$\tilde{b}_5=10.198\pm0.27770$
$\tilde{b}_6=5.1211\pm3.2747\times10^{-2}$
$\tilde{b}_7=5.4730\mp3.2200\times10^{-3}$
$\tilde{b}_8=65.754\pm25.550$
$\tilde{b}_9=3.1149\times10^{-4}\pm5.4093\times10^{-5}$
$\tilde{b}_{10}=487.89\pm6.2288$
$\tilde{b}_{11}=-1.3995\times10^{-4}\mp1.1701\times10^{-4}$
$\tilde{b}_{12}=13.288\pm1.9832$
$\tilde{b}_{13}=1.3516\times10^{-4}\pm2.0377\times10^{-5}$
$\tilde{b}_{14}=8.0716\times10^{-2}\pm1.5001\times10^{-4}\ .$

Problem 5.7-1

a. Let $\underline{X}=\{X_{ij}\}$, $(\underline{X}^T\underline{\omega}\underline{X})^{-1}=\{A_{ij}\}$. Then

$$\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}=\{\sum_{k=1}^{p}X_{ik}A_{kl}\}=\{C_{il}\}$$

$$\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T=\{\sum_{l=1}^{p}C_{il}X_{jl}\}=$$

$$\{\sum_{l=1}^{p}\sum_{k=1}^{p}X_{ik}A_{kl}X_{jl}\}=\{\sum_{l=1}^{p}X_{jl}\sum_{k=1}^{p}X_{ik}A_{kl}\}\ .$$

If $i=j$, then the entry is

$$\sum_{l=1}^{p}\sum_{k=1}^{p}X_{ik}A_{kl}X_{il}=\sum_{l=1}^{p}X_{il}\sum_{k=1}^{p}X_{ik}A_{kl}.$$

Compute for $i=j=1$ for data set 1.

$\underline{X}_1=[0.95\ 0.05\ 23{,}750]\ .$

$$(\underline{X}^T\underline{\omega}\underline{X})^{-1}=\begin{bmatrix} 0.9157290438 & 0.2752072275 & -5.983931169\times10^{-6} \\ & 0.7831966299 & -5.318014058\times10^{-6} \\ \text{symmetric} & & 6.878625596\times10^{-11} \end{bmatrix}.$$

$\underline{X}_1(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}_1^T=0.95(0.95\times0.9157290438$
$+0.05\times0.2752072275$
$+23{,}750\times(-5.983931169\times10^{-6}))$
$+0.05(0.95\times0.2752072275$
$+0.05\times0.7831966299$
$+23{,}750\times(-5.318014058\times10^{-6}))$
$+23{,}750\times(0.95\times(-5.983931169\times10^{-6})$
$+0.05\times(-5.318014058\times10^{-6})$
$+23{,}750\times6.878625596\times10^{-11}))$
$=0.6106927103\,.$

$\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T=\underline{X}\underline{C}\underline{C}^{-1}(\underline{X}^T\underline{\omega}\underline{X})^{-1}(\underline{C}\underline{C}^{-1})^T\underline{X}^T$
$\qquad\qquad=\underline{X}\underline{C}(\underline{C}^T\underline{X}^T\underline{\omega}\underline{X}\underline{C})^{-1}\underline{C}^T\underline{X}^T$
$\qquad\qquad=(\underline{X}\underline{C})((\underline{X}\underline{C})^T\underline{\omega}(\underline{X}\underline{C}))^{-1}(\underline{X}\underline{C})^T$
$\qquad\qquad=\underline{S}(\underline{S}^T\underline{\omega}\underline{S})^{-1}\underline{S}^T\,.$

b. $\widehat{\mathrm{Var}}(\hat{f}_j)=\underline{X}_j\widehat{\mathrm{Var}}(\hat{\underline{b}})\underline{X}_j^T$

Let $j=1$, data set 1.

$\widehat{\mathrm{Var}}(\hat{f}_1)=\underline{X}_1(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2\underline{X}_1^T$

$\qquad\qquad=0.6106927103\times0.30975625$

$\qquad\qquad=0.1891658838\,.$

c. $f_{\beta j}=\hat{f}_j\pm\sqrt{pF_\alpha(p,n-p)}s_{yj}$

$F_{0.05}(3,8)=4.066$

$\qquad\hat{f}_1=0.95\hat{b}_1+0.05\hat{b}_2+23{,}750\hat{b}_3$

$\qquad\quad=0.95(50.1204)+0.05(9.48742)$

$\qquad\quad+23{,}750(2.30248\times10^{-5})$

$\qquad\quad=48.636$

$\qquad f_{\beta1}=48.636\pm\sqrt{3\times4.066}\times0.434932$

$\qquad\quad=48.636\pm1.519$

for $j=1$, data set 1.

Problem 6.2–1

Sets of parameters for the modified Beale's measure:
1. (0.12398,0.00042675)
2. (0.10300,0.00067767)
3. (0.10337,0.00068217)
4. (0.12361,0.00042225)

Only 2 and 4 need be used because the other two are nearly the same. By using 2, drawdown, s, is: 1.5947, 2.2223, 2.5541, 2.8218, 3.0676, 3.3503, 3.5979. By using 4, s is: 1.7821, 2.3239, 2.6058, 2.8320, 3.0390, 3.2763, 3.4839.

The resulting value of the modified Beale's measure is

$$\hat{N}_b=0.027702$$

Because $F_{0.05}(2,5)=5.7861$, $0.09/F=0.0156$ and $1/F=0.173$, so that the model is almost roughly linear.

The input data for the modified Beale's measure program are shown in figure 1.

2	2	7	0	2	.0014328	
.11349	.00055221					
1.6715	2.2521	2.5564	2.8012	3.0256	3.2832	3.5086
1	1	1	1	1	1	1
-8.09604	-1362.96					
-12.9342	-1420.13					
-15.5344	-1436.83					
-17.6456	-1446.30					
-19.5912	-1452.76					
-21.8336	-1458.27					
-23.8030	-1461.84					
.10300	.00067767					
1.5947	2.2223	2.5541	2.8218	3.0676	3.3503	3.5979
.12361	.00042225					
1.7821	2.3239	2.6058	2.8320	3.0390	3.2763	3.4839

Figure 1

Problem 6.2-2

From the computer output, the modified Beale's measure is 0.29808. Based on $F_{0.05}(2,27)$ =3.359, the modified Beale's measure indicates that the model is at the point of being highly nonlinear. Hence, linear theory can be applied to use the W statistic based on $q=2$ only as a very rough approximation.

The input data to the regression code, modified to compute the modified Beale's measure, are shown in figure 1.

```
        CLASS PROBLEM
        TWO-DIMENSIONAL FLOW
        SEVERAL ZONES
    15  16    4   32    7   14    0    4    2   10    0    0    0
         2       .08        0         0         1
DX       1    0
 1  14    1    1        1    1
       1000      1000     1000     1000     1000     1000     1000    400
       1000      1000     1000     1000     1000     1000
DY       1    0
 1  15    1    1        1    1
       1000      1000     1000     1000     1000      400     .1000   1000
       1000      1000     1000     1000     1000     1000     1000
CX       2    0
 6  14    1    8        1    0
 1  14    9   15        1    0
CY       2    0
 6  14    1    8        1    0
 1  14    9   15        1    0
VL       2    0
 8  14    6    6        1    0
 8   8    7   15        1    0
HR       2    0
 8  15    6    7      4.5    0
 8   9    7   16      4.5    0
QR       2    0
 6  14    1    8        1    0
 1  14    9   15        1    0
HC       1    0
 1  15    1   16        0    0
IZN      7    0
 6  14    1    8    2    0
 1  14    9   15    2    0
 1   6   12   15    1    0
 6   7    1    3    3    0
 8  14    1    4    3    0
 8  14    6    6    4    0
 8   8    7   15    4    0
 1   8    2     7000       1000      60.70         1
 2  14    2    12400       1000      75.64         1
 3  12    3    10400       2000      60.27         1
 4  10    4     8400       3000      29.67         1
 5   7    5     6000       4000       4.22         1
 6  11    5     9400       4000       4.37         1
 7  13    5    11400       4000       6.07         1
 8  14    5    13400       4000       5.81         1
 9  10    7     8400       5400       4.57         1
10   8    8     7000       6400       5.21         1
11  12    8    10400       6400     -44.89         1
12  14    8    13400       6400       7.01         1
13   7    9     6000       7400       6.95         1
14   4   10     3000       8400      12.21         1
15   9   10     7400       8400       4.04         1
```

Figure 1

```
16  11  10      9400      8400    -89.36        1
17   7  11      6000      9400      6.68        1
18  13  11     11400      9400    -15.32        1
19   3  12      2000     10400     16.88        1
20   5  12      4000     10400     15.87        1
21   9  12      7400     10400      4.48        1
22  11  12      9400     10400    -18.34        1
23  13  13     11400     11400     -2.47        1
24  14  13     13400     11400      8.10        1
25   3  14      2000     12400     54.12        1
26   5  14      4000     12400     38.27        1
27  10  14      8400     12400      .053        1
28  12  14     10400     12400     -2.92        1
29   7  15      6000     13400      8.30        1
30  14  15     12400     13400      4.54        1
31   2  15      1000     14400     85.82        1
32  11  15      9400     14400      2.26        1
 1   8   8    0      9
 2  10  10    0     11
 3  12  12    0     13
 4  10  10   14      0
 8       0
 9   .00012
10      84
11       0
12       0
13  .000051
14     .008
 1  65.754     65.754         0 .00031149
 2 487.89     487.89          0-.00013995
 3  13.288     13.288         0 .00013516
 4 487.89     487.89    .080716         0
 1  16   7   16   1     80.978         0      .01
 7  16  15   16   2    935.15          0   .00056
11  10  11   10   3    -97000       1940        1
12   8  12    8   4    -50961       1020        1
IN   1   0
15  15   5   16  -1     0
 1  10   5    6         1.04         .48
15  16  10.198
15  15
15  14
15  13
15  12
15  11
15  10
15   9
15   8
15   7   5.1211
 2   2   6    7          .48         .54
15   6   5.1211
15   5   5.4730
 4
```

Figure 1—Continued

```
 1     90.259
 2    1143.6
 3     -97085
 4     -50990
 5     10.313
 6      5.1349
 7      5.4858
 8     72.988
 9   3.4248E-4
10     490.46
11  -1.8916E-4
12     22.688
13   2.2823E-4
14   8.0751E-2
 1     71.697
 2     726.73
 3     -96915
 4     -50932
 5     10.083
 6      5.1073
 7      5.4602
 8     58.520
 9   2.8050E-4
10     485.32
11  -9.0735E-5
12      3.8882
13   4.2091E-5
14   8.0681E-2
 1     124.97
 2    1424.3
 3     -97222
 4     -51035
 5     10.476
 6      5.1538
 7      5.4698
 8     91.304
 9   3.6558E-4
10     494.12
11  -2.5696E-4
12     15.271
13   1.5554E-4
14   8.0866E-2
 1     36.987
 2     446.01
 3     -96778
 4     -50887
 5      9.9203
 6      5.0884
 7      5.4762
 8     40.204
 9   2.5740E-4
10     481.66
11  -2.2940E-5
12     11.305
13   1.1478E-4
14   8.0566E-2
 2      .98677
```

Problem 6.3-1

Compute

$$\hat{\gamma} = (\underline{Y}_p - \underline{X}_p \hat{\underline{b}}*)^T [s^2 \underline{X}_p (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U}]^{-1}$$
$$\cdot (\underline{Y}_p - \underline{X}_p \hat{\underline{b}}*).$$

For data set 1

$$\hat{\underline{b}}* = \begin{bmatrix} 50.018 \\ 9.1954 \\ 2.5008 \times 10^{-5} \end{bmatrix}.$$

For data set 2

$$\hat{\underline{b}}* = \begin{bmatrix} 50.055 \\ 9.7914 \\ 2.2766 \times 10^{-5} \end{bmatrix}.$$

For data set 1

$$(\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} s^2 =$$

$$\begin{bmatrix} 2.5550 \times 10^{-1} & 8.9780 \times 10^{-2} & -1.7349 \times 10^{-6} \\ & 2.5550 \times 10^{-1} & -1.7349 \times 10^{-6} \\ \text{symmetric} & & 2.0715 \times 10^{-11} \end{bmatrix}$$

$$s^2 \underline{X}_p (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T = 1 \times 2.5550 \times 10^{-1} \times 1$$
$$= 0.25550$$

$$\underline{U} = (1.1)^2 = 1.21$$

$$s^2 \underline{X}_p (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U} = 0.25550 + 1.21$$
$$= 1.46550$$

$$[s^2 \underline{X}_p (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U}]^{-1}$$
$$= 0.68236$$

$$\underline{Y}_p - \underline{X}_p \hat{\underline{b}}* = 11 - 9.1954 = 1.8046$$

$$\hat{\gamma} = 1.8046 \times 0.68236 \times 1.8046 = 2.222$$

$$\chi_{0.05}^2(1) = 3.841 \therefore H_0 \text{ accepted.}$$

Prior and pure regression estimates of $\beta_2 = h_b$ are in agreement.

For data set 2

$$(\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} s^2 =$$

$$\begin{bmatrix} 5.4121 \times 10^{-1} & 2.6265 \times 10^{-1} & -3.9795 \times 10^{-6} \\ & 5.4121 \times 10^{-1} & -3.9795 \times 10^{-6} \\ \text{symmetric} & & 4.3413 \times 10^{-11} \end{bmatrix}$$

$$s^2 \underline{X}_p (\underline{X}_s \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T = 1 \times 5.4121 \times 10^{-1} \times 1$$
$$= 0.54121$$

$$\underline{U} = (0.95)^2 = 0.9025$$

$$s^2 \underline{X}_p (\underline{X}_s \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U} = 0.54121 + 0.9025$$
$$= 1.44371$$

$$[s^2 \underline{X}_p (\underline{X}_s \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U}]^{-1}$$
$$= 0.69266$$

$$\underline{Y}_p - \underline{X}_p \hat{\underline{b}}* = 9.5 - 9.7914 = -0.2914$$

$$\hat{\gamma} = -0.2914 \times 0.69266 \times (-0.2914)$$
$$= 0.5882$$
$$\therefore H_0 \text{ accepted.}$$