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

Click here to return to USGS Publications

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

$$\underline{f}_{\beta}-\underline{f}_0\cong\underline{\underline{X}}(\underline{\beta}-\underline{b}_0) \qquad (5.1\text{-}1)$$

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

$$\underline{f}_{\beta}=\underline{f}(\underline{\xi},\underline{\beta}) \qquad (5.1\text{-}2)$$

$$\underline{f}_0=(\underline{\xi},\underline{b}_0)\ . \qquad (5.1\text{-}3)$$

Based on equation 5.1-1, the true regression model is

$$\underline{Y}-\underline{f}_0\cong\underline{\underline{X}}(\underline{\beta}-\underline{b}_0)+\underline{\epsilon} \qquad (5.1\text{-}4)$$

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

$$\underline{Y}-\underline{f}_0\cong\underline{\underline{X}}(\underline{b}-\underline{b}_0)+\underline{e} \qquad (5.1\text{-}5)$$

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 $\hat{\underline{b}}$ of $\underline{\beta}$ and $\hat{\underline{e}}=\underline{Y}-\underline{f}(\underline{\xi},\hat{\underline{b}})$ of $\underline{\epsilon}$ are obtained. For a linear model $\hat{\underline{b}}$ 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 $\hat{\underline{b}}$. Thus, by minimizing $S(\underline{b})$ using equation 5.1-5 as the estimated regression model, normal equations

$$\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0)\cong\underline{\underline{X}}^T\underline{\underline{\omega}}(\underline{Y}-\underline{f}_0) \qquad (5.1\text{-}6)$$

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 $\hat{\underline{b}}$ and $\hat{\underline{e}}$ is

$$\underline{Y}-\underline{f}_0\cong\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0)+\hat{\underline{e}}\ . \qquad (5.1\text{-}7)$$

By utilizing the definition of $\hat{\underline{e}}$ ($\hat{\underline{e}}=\underline{Y}-\hat{\underline{f}}$, where $\hat{\underline{f}}=\underline{f}(\underline{\xi},\hat{\underline{b}})$), in equation 5.1-7, a predictive model

$$\hat{\underline{f}}-\underline{f}_0\cong\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0) \qquad (5.1\text{-}8)$$

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 $\hat{\underline{b}}$ in equation 5.1-6, then $\underline{f}_0=\hat{\underline{f}}$, and

$$\underline{\underline{X}}^T\underline{\underline{\omega}}(\underline{Y}-\hat{\underline{f}})=\underline{0}\ . \qquad (5.1\text{-}9)$$

Because any approximation inherent in equation 5.1-6 is removed as $\underline{b}_0\rightarrow\hat{\underline{b}}$, 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:

$$Y=f(\xi_1,\xi_2,\ldots,\xi_k;\underline{\beta})+\epsilon \quad . \tag{5.2-1}$$

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

$$E(\underline{\epsilon})=\underline{0} \tag{5.2-2}$$

$$\mathrm{Var}(\underline{\epsilon})=\underline{V}\sigma^2 \tag{5.2-3}$$

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 $\underline{V}^{-1/2}$ to obtain

$$E(\underline{V}^{-1/2}\underline{\epsilon})=\underline{0} \tag{5.2-4}$$

$$\mathrm{Var}(\underline{V}^{-1/2}\underline{\epsilon})=\underline{I}\sigma^2 \quad . \tag{5.2-5}$$

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 $\underline{V}^{-1/2}\underline{\epsilon}$ 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 $\underline{V}^{-1/2}$ are equivalent; that is,

$$\underline{\omega}=\underline{V}^{-1} \quad . \tag{5.2-6}$$

For a linear model at least, the Gauss-Markov theorem (Beck and Arnold, 1977, p. 232–234) establishes that the variance of $\hat{b}_j$, $\mathrm{Var}(\hat{b}_j)$, is a minimum if equation 5.2-6 is true. Furthermore, to compute $\mathrm{Var}(\underline{\hat{b}})$ 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 $\underline{\hat{b}}-\underline{b}_0$ and take the expected value of it to obtain:.

$$\begin{aligned}E(\underline{\hat{b}}-\underline{b}_0)&=(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}E(\underline{Y}-\underline{f}_0)\\&=(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}\underline{X}(\underline{\beta}-\underline{b}_0)=\underline{\beta}-\underline{b}_0\end{aligned} \tag{5.2-7}$$

where the fact that $E(\underline{\epsilon})=0$ was used. From equation 5.2-7 it is seen that

$$E(\underline{\hat{b}})=\underline{\beta} \quad . \tag{5.2-8}$$

If $\underline{V}$ is diagonal, then

$$\underline{V}=\begin{bmatrix}1/\omega_1 & & & \\ & 1/\omega_2 & & \\ & & \ddots & \\ & & & 1/\omega_n\end{bmatrix} \quad . \tag{5.2-9}$$

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

$$\underline{V}^{-1/2}\underline{\epsilon} \sim N(\underline{0}, \underline{I}\sigma^2) \; . \qquad (5.2\text{-}11)$$

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 $\underline{V}^{-1/2}\underline{\epsilon}$) be normally distributed implies that the elements of $\underline{V}^{-1/2}\underline{\epsilon}$ 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 $\underline{V}^{-1/2}\underline{\epsilon}$) 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 $\underline{\beta}$, 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, $\underline{\omega}^{1/2}\underline{\hat{e}}$, and weighted disturbances, $\underline{\omega}^{1/2}\underline{\epsilon}$. In this case equation 5.3-3 becomes

$$\underline{\omega}^{1/2}\underline{\hat{e}} = (\underline{I} - \underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2})\underline{\omega}^{1/2}\underline{\epsilon} \; . \qquad (5.3\text{-}4)$$

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:

$$[\underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2}][\underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2}]$$

$$= \underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2}\underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2}$$

$$= \underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2} \; . \qquad (5.3\text{-}5)$$

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

$$\underline{R} = \underline{\omega}^{1/2}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}^{1/2} \qquad (5.3\text{-}6)$$

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

$$S(\hat{\underline{b}})=\hat{\underline{e}}^T\underline{\underline{\omega}}\hat{\underline{e}}$$
$$=(\underline{\underline{\omega}}^{1/2}\hat{\underline{e}})^T(\underline{\underline{\omega}}^{1/2}\hat{\underline{e}}) \quad . \qquad (5.4\text{-}1)$$

Using equations 5.3-4, 5.3-7, and the fact that tr(scaler)=scaler, equation 5.4-1 becomes

$$S(\hat{\underline{b}})=(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})^T(\underline{\underline{I}}-\underline{\underline{R}})^T(\underline{\underline{I}}-\underline{\underline{R}})(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})$$
$$=\mathrm{tr}[(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})^T(\underline{\underline{I}}-\underline{\underline{R}})(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})]$$
$$=\mathrm{tr}[(\underline{\underline{I}}-\underline{\underline{R}})(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})^T] \quad . \qquad (5.4\text{-}2)$$

The expected value of equation 5.4-2 is

$$E(S(\hat{\underline{b}}))=\mathrm{tr}\{(\underline{\underline{I}}-\underline{\underline{R}})E[(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})^T]\}$$
$$=\mathrm{tr}[(\underline{\underline{I}}-\underline{\underline{R}})\mathrm{Var}(\underline{\underline{\omega}}^{1/2}\underline{\epsilon})]$$
$$=\mathrm{tr}[(\underline{\underline{I}}-\underline{\underline{R}})\sigma^2]$$
$$=(n-p)\sigma^2 \qquad (5.4\text{-}3)$$

from which

$$\sigma^2=\frac{E(S(\hat{\underline{b}}))}{n-p} \quad . \qquad (5.4\text{-}4)$$

The fact that $\mathrm{tr}(\underline{\underline{R}})=p$ can be demonstrated by rearranging the matrices within $\underline{\underline{R}}$. The estimate, s^2, of σ^2 is

$$s^2=\frac{S(\hat{\underline{b}})}{n-p}=\frac{(\underline{Y}-\hat{\underline{f}})^T\underline{\underline{\omega}}(\underline{Y}-\hat{\underline{f}})}{n-p} \quad . \qquad (5.4\text{-}5)$$

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

$$s^2=\frac{(\underline{Y}-\underline{f}_0-\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0))^T\underline{\underline{\omega}}(\underline{Y}-\underline{f}_0-\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0))}{n-p}$$
$$=\frac{(\underline{Y}-\underline{f}_0)^T\underline{\underline{\omega}}(\underline{Y}-\underline{f}_0)-2(\hat{\underline{b}}-\underline{b}_0)^T\underline{\underline{X}}^T\underline{\underline{\omega}}(\underline{Y}-\underline{f}_0)+(\hat{\underline{b}}-\underline{b}_0)^T\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0)}{n-p}$$
$$=\frac{(\underline{Y}-\underline{f}_0)^T\underline{\underline{\omega}}(\underline{Y}-\underline{f}_0)-(\hat{\underline{b}}-\underline{b}_0)^T\underline{\underline{X}}^T\underline{\underline{\omega}}(\underline{Y}-\underline{f}_0)}{n-p} \qquad (5.4\text{-}6)$$

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 $\underline{\underline{\omega}}^{1/2}\underline{Y}$ and $\underline{\underline{\omega}}^{1/2}\hat{\underline{f}}$

This measure is defined as

$$R_y=\frac{\underline{d}_y^T\underline{d}_{\hat{y}}}{[(\underline{d}_y^T\underline{d}_y)(\underline{d}_{\hat{f}}^T\underline{d}_{\hat{f}})]^{1/2}} \qquad (5.4\text{-}7)$$

where

$$\underline{d}_y=\underline{\underline{\omega}}^{1/2}\underline{Y}-m_y\underline{l} \qquad (5.4\text{-}8)$$

$$\underline{d}_{\hat{f}}=\underline{\underline{\omega}}^{1/2}\hat{\underline{f}}-m_{\hat{f}}\underline{l} \qquad (5.4\text{-}9)$$

$$m_y=\sum_{i=1}^{n}\underline{\omega}_i^{1/2}\underline{Y}/n \qquad (5.4\text{-}10)$$

$$m_{\hat{f}}=\sum_{i=1}^{n}\underline{\omega}_i^{1/2}\hat{\underline{f}}/n \qquad (5.4\text{-}11)$$

$\underline{l}$=vector of ones,

and $\underline{\omega}_i^{1/2}$ is row i of $\underline{\underline{\omega}}^{1/2}$. 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

$$\text{Var}(\hat{\underline{b}}) = \text{Var}[(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}(\underline{Y}-\underline{f}_0)+\underline{b}_0]$$

$$=(\underline{X}^T\underline{\omega}\underline{X})^{-1}\underline{X}^T\underline{\omega}\,\text{Var}(\underline{Y})\underline{\omega}\underline{X}(\underline{X}^T\underline{\omega}\underline{X})^{-1}$$

$$=(\underline{X}^T\underline{\omega}\underline{X})^{-1}\sigma^2 \qquad (5.4\text{-}12)$$

where assumptions from equations 5.2-3 and 5.2-6 and the fact that $\text{Var}(\underline{Y})=\text{Var}(\underline{Y}-\underline{f}_\beta)=\underline{\omega}^{-1}\sigma^2$ were employed. The estimate of $\text{Var}(\hat{\underline{b}})$ is

$$\text{Var}(\hat{\underline{b}})=(\underline{X}^T\underline{\omega}\underline{X})^{-1}s^2 \; . \qquad (5.4\text{-}13)$$

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

$$r_{ij}=\frac{\text{Cov}(\hat{b}_i,\hat{b}_j)}{[\text{Var}(\hat{b}_i)\text{Var}(\hat{b}_j)]^{1/2}} \qquad (5.4\text{-}14)$$

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 $\text{Var}(\hat{\underline{b}})$. 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

$$\underline{\omega}^{1/2}\underline{\hat{e}}=(\underline{I}-\underline{R})\underline{\omega}^{1/2}\underline{\epsilon} \qquad (5.5\text{-}1)$$

from which

$$E(\underline{\omega}^{1/2}\underline{\hat{e}})=(\underline{I}-\underline{R})\underline{\omega}^{1/2}E(\underline{\epsilon})$$

$$=\underline{0} \qquad (5.5\text{-}2)$$

and

$$\mathrm{Var}(\underline{\omega}^{1/2}\underline{\hat{e}})=\mathrm{Var}[(\underline{I}-\underline{R})\underline{\omega}^{1/2}\underline{\epsilon}]$$

$$=(\underline{I}-\underline{R})\mathrm{Var}(\underline{\omega}^{1/2}\underline{\epsilon})(\underline{I}-\underline{R})^T$$

$$=(\underline{I}-\underline{R})(\underline{I}-\underline{R})^T\sigma^2$$

$$=(\underline{I}-\underline{R})\sigma^2 \ . \qquad (5.5\text{-}3)$$

Hence, if assumptions given by equations 5.2-4 through 5.2-6, and 5.2-11 hold,

$$\underline{\omega}^{1/2}\underline{\hat{e}}=\underline{\hat{u}} \sim N(\underline{0},(\underline{I}-\underline{R})\sigma^2) \qquad (5.5\text{-}4)$$

where for convenience, the definition is made that $\underline{\omega}^{1/2}\underline{\hat{e}}=\underline{\hat{u}}$.

It can be shown that $\underline{I}-\underline{R}\rightarrow\underline{I}$ as $n-p\rightarrow\infty$. Whenever $n-p$ 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 $\underline{\hat{u}}$ to help decide whether the distribution of $\underline{\hat{u}}$ 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

$$\mathrm{Var}(\underline{g})=\underline{\Omega}\mathrm{Var}(\underline{d})\underline{\Omega}^T$$

$$=\underline{\Omega}^2s^2 \qquad (5.5\text{-}7)$$

where the definition of $\mathrm{Var}(\underline{d})$ and the symmetry of $\underline{\Omega}$ were used. Hence, $\underline{\Omega}$ must be defined so that

$$\underline{\Omega}^2=\underline{I}-\underline{R} \ . \qquad (5.5\text{-}8)$$

However, because $\underline{I}-\underline{R}$ is idempotent, $(\underline{I}-\underline{R})=(\underline{I}-\underline{R})^2$, and equation 5.5-8 may be simplified to become

$$\underline{\Omega}=\underline{I}-\underline{R} \qquad (5.5\text{-}9)$$

so that equation 5.5-5 assumes as its final form

$$\underline{g}=(\underline{I}-\underline{R})\underline{d} \ . \qquad (5.5\text{-}10)$$

To generate a set of simulated residuals, $\underline{g}$, 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

$$F_i = m_i/(n+1), i = 1,2, \ldots, n \qquad (5.5\text{–}11)$$

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

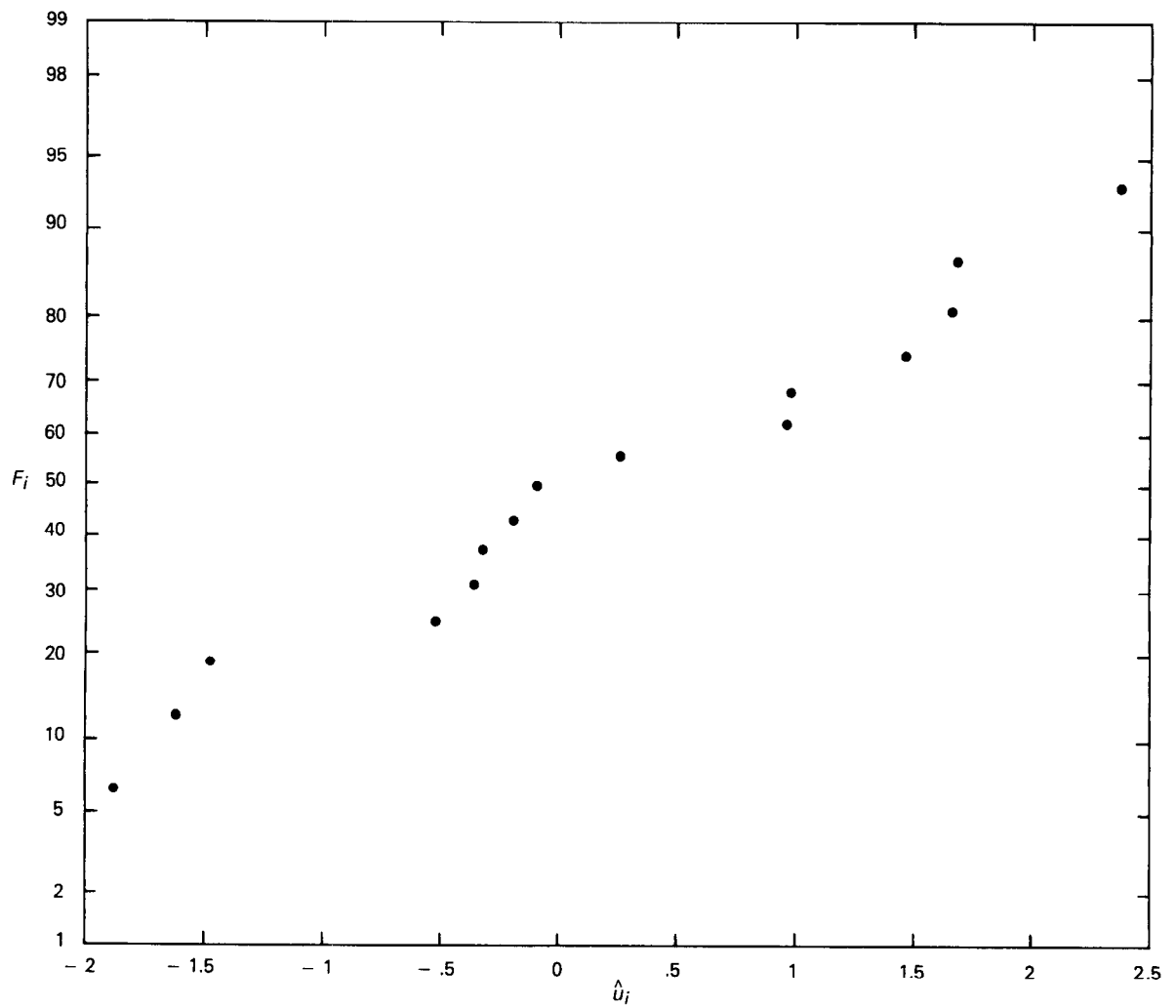

Figure 5.5–1

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, p. 157-162) could be used to test for randomness of signs of $\hat{u}_j$ along the $\hat{f}_j$ axis.

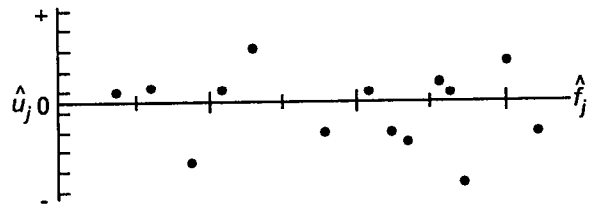

Figure 5.5-2

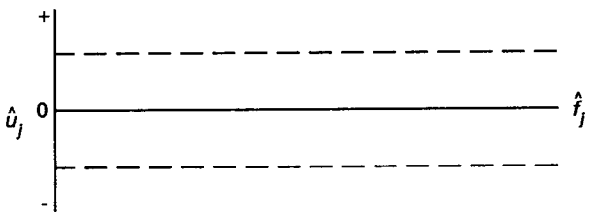

Figure 5.5-3

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,

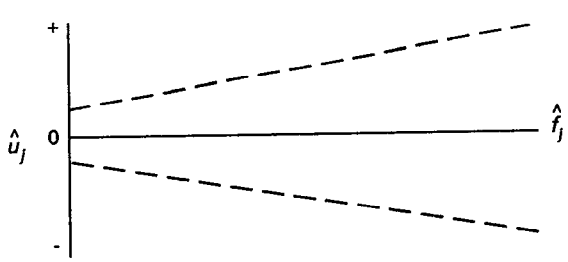

Figure 5.5-4

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 $\mathrm{Cov}(\hat{\underline{u}},\hat{\underline{f}})=\underline{0}$ 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.

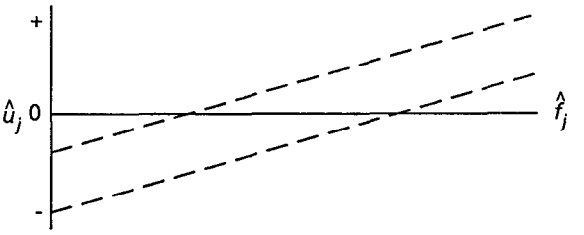

Figure 5.5-5

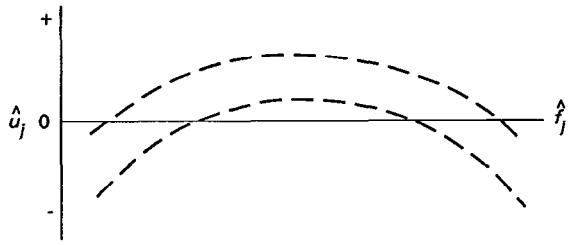

Figure 5.5-6

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 $\widehat{\mathrm{Var}}(\hat{\underline{b}})$ using equation 5.4-13. You will have to unscale the entries of $(\underline{Z}^T\underline{Z})^{-1}$ to obtain $\widehat{\mathrm{Var}}(\hat{\underline{b}})$. Examination of the calculation 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 $\hat{\underline{e}}$ 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 $\hat{\underline{e}}$ differ significantly from the plots of $\underline{g}$? What can you conclude about the distribution of $\hat{\underline{e}}$? 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 $\underline{d}$ and simulated residuals $\underline{g}$, computed by using the residuals analysis program (appendix 5.8.1). Plot $\underline{d}$ and $\underline{g}$ on normal probability paper using equation 5.5-11. Are correlation effects evident? Plot $\hat{\underline{u}}$ (why $\hat{\underline{u}}$ instead of $\hat{\underline{e}}$?) on normal probability paper. Does the plot differ significantly from the plots of $\underline{g}$? 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 $\underline{\beta}_2$ of parameter set $\underline{\beta}$ cannot be distinguished from some corresponding given subset $\tilde{\underline{\beta}}_2$. That is, test

$$H_0: \underline{\beta}_2 = \tilde{\underline{\beta}}_2 \text{ versus } H_1: \underline{\beta}_2 \neq \tilde{\underline{\beta}}_2.$$

The linearized model assumed is

$$\underline{Y} = \underline{f}_0 + \underline{\underline{X}}(\underline{\beta} - \underline{b}_0) + \underline{\epsilon}$$

$$= \underline{f}_0 + \underline{\underline{X}}_1(\underline{\beta}_1 - \underline{b}_{01}) + \underline{\underline{X}}_2(\underline{\beta}_2 - \underline{b}_{02}) + \underline{\epsilon} \quad (5.6\text{-}1)$$

where $\underline{\underline{X}}$, $\underline{\beta}$, and $\underline{b}_0$ are conformably partitioned as follows:

$$\underline{\underline{X}} = [\underline{\underline{X}}_1, \underline{\underline{X}}_2] \quad (5.6\text{-}2)$$

$$\left.\begin{aligned} \underline{\beta} &= \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{bmatrix} \\ \underline{b}_0 &= \begin{bmatrix} \underline{b}_{01} \\ \underline{b}_{02} \end{bmatrix} \end{aligned}\right\} \quad (5.6\text{-}3)$$

Based on equation 5.6-1, we may state a predictive model of the form

$$\tilde{\underline{f}} = \underline{f}_0 + \underline{\underline{X}}_1(\tilde{\underline{b}}_1 - \underline{b}_{01}) + \underline{\underline{X}}_2(\tilde{\underline{\beta}}_2 - \underline{b}_{02}) \quad (5.6\text{-}4)$$

where $\tilde{\underline{b}}_1$ is an estimate of $\underline{\beta}_1$ and, under the null hypothesis, $\underline{\beta}_2$ is assumed to be given by $\tilde{\underline{\beta}}_2$.

The W statistic, which is stated explicitly later on, is based on a comparison of the restricted sum of squares $(\underline{Y} - \tilde{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \tilde{\underline{f}})$ and the unrestricted sum of squares $(\underline{Y} - \hat{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \hat{\underline{f}})$. The unrestricted sum of squares is obtained from the standard least squares analysis. The restricted sum of squares is obtained by minimizing

$$S_0(\tilde{\underline{b}}_1) = (\underline{Y} - \tilde{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \tilde{\underline{f}}) \quad (5.6\text{-}5)$$

with respect to $\tilde{\underline{b}}_1$ while holding $\tilde{\underline{\beta}}_2$ constant. For a linear model this results in the normal equations:

$$\tilde{\underline{b}}_1 - \underline{b}_{01} = (\underline{\underline{X}}_1^T \underline{\underline{\omega}} \underline{\underline{X}}_1)^{-1} \underline{\underline{X}}_1^T \underline{\underline{\omega}} (\underline{Y} - \underline{f}_0 - \underline{\underline{X}}_2(\tilde{\underline{\beta}}_2 - \underline{b}_{02})) \quad (5.6\text{-}6)$$

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 $\tilde{\underline{b}}_1$ by using the standard procedures, except that $\tilde{\underline{\beta}}_2$ 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

$$W = \frac{[(\underline{Y} - \tilde{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \tilde{\underline{f}}) - (\underline{Y} - \hat{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \hat{\underline{f}})]/q}{(\underline{Y} - \hat{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \hat{\underline{f}})/(n-p)}$$

$$= \frac{[(\underline{Y} - \tilde{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \tilde{\underline{f}}) - (\underline{Y} - \hat{\underline{f}})^T \underline{\underline{\omega}} (\underline{Y} - \hat{\underline{f}})]/q}{s^2}$$

$$\sim F(q, n-p) \quad (5.6\text{-}7)$$

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

$$(\underline{Y}-\tilde{\underline{f}})^T\underline{\underline{\omega}}(\underline{Y}-\tilde{\underline{f}})-(\underline{Y}-\hat{\underline{f}})^T\underline{\underline{\omega}}(\underline{Y}-\hat{\underline{f}})$$

$$=(\tilde{\underline{\beta}}_2-\hat{\underline{b}}_2)^T[\underline{\underline{H}}(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{\underline{H}}^T]^{-1}(\tilde{\underline{\beta}}_2-\hat{\underline{b}}_2) \quad (5.6\text{-}8)$$

where

$$\underline{\underline{H}}=[\underline{\underline{0}},\ \underline{\underline{I}}_q]_{(q\times p)}, \quad (5.6\text{-}9)$$

$\underline{\underline{I}}_q$=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. (5.6-10)

Thus, the alternative form for equation 5.6-7 is

$$W=\frac{(\tilde{\underline{\beta}}_2-\hat{\underline{b}}_2)^T[\underline{\underline{H}}(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{\underline{H}}^T]^{-1}(\tilde{\underline{\beta}}_2-\hat{\underline{b}}_2)/q}{s^2}. \quad (5.6\text{-}11)$$

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 $\hat{\underline{f}}$, but equation 5.6-7 requires, in addition, a restricted regression using equation 5.6-6 to obtain $\tilde{\underline{b}}_1$ and $\tilde{\underline{f}}$. 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 $\underline{Y}=\underline{f}_0+\underline{\underline{X}}(\underline{\beta}-\underline{b}_0)+\underline{\epsilon}$ 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 $\tilde{\underline{b}}$, 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 joint confidence region 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

$$(\underline{Y}-\tilde{\underline{f}})^T\underline{\underline{\omega}}(\underline{Y}-\tilde{\underline{f}})-(\underline{Y}-\hat{\underline{f}})^T\underline{\underline{\omega}}(\underline{Y}-\hat{\underline{f}})\leq s^2qF_\alpha(q,n-p) \quad (5.6\text{-}12)$$

or

$$(\underline{\beta}_2-\underline{\hat{b}}_2)^T[\underline{\underline{H}}(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{\underline{H}}^T]^{-1}(\underline{\beta}_2-\underline{\hat{b}}_2) \leq s^2qF_\alpha(q,n-p)\,. \qquad (5.6\text{-}13)$$

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 $\underline{\hat{b}}$. 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

$$\underline{\tilde{b}}=\underline{\hat{b}}\pm\frac{\sqrt{qF_\alpha(q,n-p)}}{s_{bi}}\underline{V}_{bi} \qquad (5.6\text{-}14)$$

where $\underline{\tilde{b}}^T=[\underline{\tilde{b}}_1^T,\underline{\tilde{\beta}}_2^T]$ and $\underline{V}_{bi}$ is the ith column of $\underline{\underline{V}}_b=(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}s^2$. 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 $\underline{\tilde{b}}$ 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 $\underline{\tilde{\beta}}_2$ 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 $\underline{\tilde{b}}_1$ were computed using equation 5.6-6.

Parameter sets computed using equation 5.6-14 will exactly satisfy the relationship $s^2qF_\alpha(q,n-p)=(\underline{Y}-\underline{\tilde{f}})^T\underline{\underline{\omega}}(\underline{Y}-\underline{\tilde{f}})-(\underline{Y}-\underline{\hat{f}})^T\underline{\underline{\omega}}(\underline{Y}-\underline{\hat{f}})=(\underline{\tilde{\beta}}_2-\underline{\hat{b}}_2)^T[\underline{\underline{H}}(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{\underline{H}}^T]^{-1}(\underline{\tilde{\beta}}_2-\underline{\hat{b}}_2)$ 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 $(\underline{Y}-\underline{\tilde{f}})^T\underline{\underline{\omega}}(\underline{Y}-\underline{\tilde{f}})-(\underline{Y}-\underline{\hat{f}})^T\underline{\underline{\omega}}(\underline{Y}-\underline{\hat{f}})$ 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 ft^2/day. Using these estimates, test the null hypothesis that there is no significant difference at $\alpha=0.05$ between W/T as estimated above

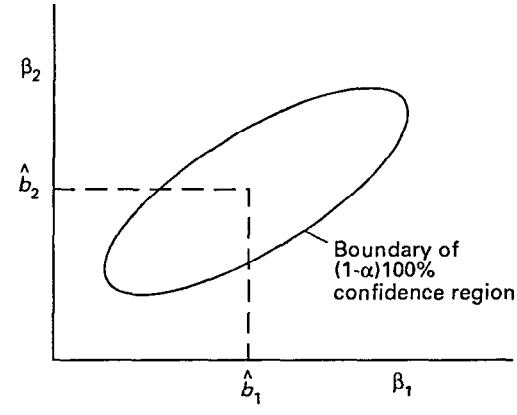

Figure 5.6-1

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

$$\mathrm{Var}(\underline{\underline{\omega}}^{1/2}\hat{\underline{f}})=\mathrm{Var}[\underline{\underline{\omega}}^{1/2}\underline{\underline{X}}(\hat{\underline{b}}-\underline{b}_0)+\underline{\underline{\omega}}^{1/2}\underline{f}_0]$$
$$=\underline{\underline{\omega}}^{1/2}\underline{\underline{X}}\,\mathrm{Var}(\hat{\underline{b}})\underline{\underline{X}}^T\underline{\underline{\omega}}^{1/2} \ . \qquad (5.7\text{-}2)$$

By using equations 5.3-6 and 5.4-12, equation 5.7-2 can be written

$$\mathrm{Var}(\underline{\underline{\omega}}^{1/2}\hat{\underline{f}})=\underline{\underline{R}}\sigma^2 \ . \qquad (5.7\text{-}3)$$

Estimates corresponding to equations 5.7-1, 5.7-2, and 5.7-3 are

$$\widehat{\mathrm{Var}}(\hat{\underline{f}})=\underline{\underline{X}}\,\widehat{\mathrm{Var}}(\hat{\underline{b}})\underline{\underline{X}}^T \qquad (5.7\text{-}4)$$

$$\widehat{\mathrm{Var}}(\underline{\underline{\omega}}^{1/2}\hat{\underline{f}})=\underline{\underline{\omega}}^{1/2}\underline{\underline{X}}\,\widehat{\mathrm{Var}}(\hat{\underline{b}})\underline{\underline{X}}^T\underline{\underline{\omega}}^{1/2} \qquad (5.7\text{-}5)$$

$$=\underline{\underline{R}}s^2 \ . \qquad (5.7\text{-}6)$$

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 $\mathrm{Var}(\underline{\underline{\omega}}^{1/2}\hat{\underline{f}})$ (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, $\underline{\underline{\omega}}$ must be available for all points in $\hat{\underline{f}}$. Matrix $(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}s^2$ 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

$$s_{yj}=s\sqrt{\underline{X}_j(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{X}_j^T} \qquad (5.7\text{-}8)$$

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 confidence band.

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 prediction interval. 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):

$$Y_j^{pred}=\hat{f}_j\pm\sqrt{kF_\alpha(k,n-p)}\,\sqrt{s^2/\omega_j+s_{yj}^2}\,, \qquad j=1, 2, ...,k \qquad (5.7\text{-}9)$$

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, $\mathrm{Var}(Y_j^{pred})=\mathrm{Var}(Y_j^{pred}-\hat{f}_j)+\mathrm{Var}(\hat{f}_j)$ or $\mathrm{Var}(Y_j^{pred})=s^2/\omega_j+s_{yj}^2$.

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 $\underline{X}(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{X}^T$, which is used in equation 5.7-1. For this exercise let $(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}}^{-1}=\underline{\underline{A}}$ and write the result in terms of $\underline{\underline{A}}$ 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 $\underline{X}(\underline{\underline{X}}^T\underline{\underline{\omega}}\underline{\underline{X}})^{-1}\underline{X}^T$, 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 $\mathrm{Var}(\hat{f}_j)$, 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 $\underline{\underline{X}}_p=[\underline{\underline{I}}_p, \underline{0}]$. However, matrix $[\underline{\underline{I}}_p, \underline{0}]$ 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 $\underline{g}$ 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) ⋮ COV(NVAR,1) COV(2,2) ⋮ COV(NVAR,2) ⋮ 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) ⋮ 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) ⋮ X(NVAR,1) X(1,2) ⋮ X(NVAR,2) ⋮ 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=O.

Data Set F.
Parameter numbers having prior information (format 16I5).

Line columns	Variable	Definition
1-5	IPR(1)	Array subscript numbers for regression parameters in $\underline{\hat{b}}$ having prior information. For use with the regression groundwater program, the array subscript numbers must be the subscript numbers in the parameter vector computed by that program.
6-10	IPR(2)	
⋮	⋮	
	IPR(NPRIR)	

Omit data set if NPRIR=0.

Data Set G.
Standard deviation matrix for prior information $\underline{U}^{1/2}$ (format 8F10.0).

Line columns	Variable	Definition
1-10	F(1)	Diagonal standard deviation matrix for prior information, entered in the same order as IPR(I) from 1 through NPRIR.
11-20	F(2)	
⋮	⋮	
	F(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 $\underline{g}$. 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 entry is paired with its theoretical frequency as computed by using equation 5.5-11.
4. Vector $\underline{g}$. Each entry is printed in its natural position corresponding to its position in a row or column of $\underline{\underline{R}}$. Rows and columns of $\underline{\underline{R}}$ are ordered by first sample observation numbers followed by prior information numbers, which are the subscripts I of IPR(I).
5. Vector $\underline{g}$. This vector is ordered and paired with its theoretical frequency in the same way as $\underline{d}$ is.
6. Covariance matrix $(\underline{\underline{I}}-\underline{\underline{R}})s^2$. This matrix, which is an estimate of the one defined by equation 5.5-3, is for weighted residuals $\underline{\hat{u}}$ composed of both sample and prior information, with the prior information occupying the last n_p rows and columns.
7. Correlation matrix for $\underline{\hat{u}}$. This matrix is derived from $(\underline{\underline{I}}-\underline{\underline{R}})s^2$.

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.