AN IMPROVED MULTIPLE LINEAR REGRESSION AND DATA ANALYSIS COMPUTER PROGRAM PACKAGE

by Steven M. Sidik

Lewis Research Center
Cleveland, Ohio 44135

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION - WASHINGTON, D. C. - APRIL 1972
NEWRAP, an improved version of a previous multiple linear regression program called RAPIER, CREDUC, and CRSPLT, allows for a complete regression analysis including cross plots of the independent and dependent variables, correlation coefficients, regression coefficients, analysis of variance tables, t-statistics and their probability levels, rejection of independent variables, plots of residuals against the independent and dependent variables, and a canonical reduction of quadratic response functions useful in optimum seeking experimentation. A major improvement over RAPIER is that all regression calculations are done in double precision arithmetic.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMMARY</td>
<td>1</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>2</td>
</tr>
<tr>
<td>SYMBOLS</td>
<td>3</td>
</tr>
<tr>
<td>ESTIMATION OF BASIC LINEAR MODEL</td>
<td>5</td>
</tr>
<tr>
<td>BASIC LINEAR MODEL</td>
<td>5</td>
</tr>
<tr>
<td>ESTIMATING $b$</td>
<td>8</td>
</tr>
<tr>
<td>CORRELATION MATRIX</td>
<td>11</td>
</tr>
<tr>
<td>ESTIMATING $\sigma^2$</td>
<td>11</td>
</tr>
<tr>
<td>HYPOTHESIS TESTING</td>
<td>13</td>
</tr>
<tr>
<td>NORMALITY OF $\epsilon$</td>
<td>13</td>
</tr>
<tr>
<td>ANALYSIS OF VARIANCE TABLE</td>
<td>14</td>
</tr>
<tr>
<td>CHOICE OF ESTIMATOR FOR $\sigma^2$</td>
<td>15</td>
</tr>
<tr>
<td>TEST OF OVERALL REGRESSION</td>
<td>16</td>
</tr>
<tr>
<td>t-TESTS</td>
<td>17</td>
</tr>
<tr>
<td>PREDICTING VALUES FROM ESTIMATED REGRESSION EQUATION</td>
<td>18</td>
</tr>
<tr>
<td>NEWRAP PROGRAM</td>
<td>19</td>
</tr>
<tr>
<td>USERS GUIDE TO NEWRAP INPUT</td>
<td>19</td>
</tr>
<tr>
<td>Illustrative Regression Problem Requiring No Transformations</td>
<td>19</td>
</tr>
<tr>
<td>Detailed Description of Input Cards</td>
<td>21</td>
</tr>
<tr>
<td>Illustrative Problem Requiring Transformations</td>
<td>31</td>
</tr>
<tr>
<td>NEWRAP DOCUMENTATION AND LISTINGS</td>
<td>40</td>
</tr>
<tr>
<td>General Mathematical and Logical Flow of Program</td>
<td>40</td>
</tr>
<tr>
<td>Routines and Their Major Functions</td>
<td>42</td>
</tr>
<tr>
<td>Call Structure of Program</td>
<td>42</td>
</tr>
<tr>
<td>Dictionary of Program</td>
<td>43</td>
</tr>
<tr>
<td>Program Listing</td>
<td>46</td>
</tr>
</tbody>
</table>
| CRSPLT PROGRAM | 76 
| CREDUC PROGRAM | 85 |
| APPENDIX - BORROWED ROUTINES | 91 |
| REFERENCES | 94 |
NEWRAP is a digital computer program which can be used with ease to perform extensive regression analyses or a simple least-squares curve fit. The program is written in FORTRAN IV, version 13, for the IBM 7094/7044 DCS. The major value of the program is the comprehensiveness of its calculations and options.

NEWRAP computes the variance-covariance matrix of the independent variables, regression coefficients, t-statistics for individual tests, and analysis of variance tables for overall testing of regression. There is a provision for a choice of three strategies for the variance estimate to be used in computing t-statistics.

Also, more than one set of responses of dependent variables can be analyzed for the same set of independent variables.

A backward rejection option method based on the first dependent variable may be used to delete nonsignificant terms from the model. In this case, a critical significance level is supplied as input. The least significant independent variable is deleted and the regression recomputed. This process is repeated until all remaining variables have significantly nonzero coefficients.

The NEWRAP program uses the triangular form of symmetric matrices throughout. It also allows for the use of weighted regression, computation of predicted values at any combination of independent variables, a table of residuals, and plots of residuals.

By use of CRSPLT, a preregression analysis may be performed which may aid in the choice of model to use in NEWRAP. This program accepts the same raw data in the same format and computes the variance-covariance matrix and correlation matrix of all the variables and an eigenvector decomposition of the variance-covariance matrix corresponding to the independent variables. Microfilm plots are then printed of specified pairs of variables. Punched output of residuals and predicted values from NEWRAP can also be used for more complicated residual plots than the direct use of the plotting option NEWRAP permits.

When a quadratic response function has been estimated (as for example in optimum-seeking experimentation) CREDUC may be used to obtain all information necessary for a canonical analysis of the function.

The three programs together provide a useful data analysis package that can be applied to a large variety of common research and development situations.
INTRODUCTION

RAPIER (ref. 1) is a very flexible multiple linear regression analysis computer program which has been in frequent use at the NASA Lewis Research Center. It was tested with the data presented in Wampler (ref. 2) and performed quite poorly. This alone was not very disturbing since real data are seldom even nearly as ill-conditioned as that set of data. A second factor, however, is that Wampler's data leads to a 5 by 5 matrix to be inverted whereas RAPIER is designed to handle matrices of up to 60 by 60. With real data it is not uncommon for the matrix to become more ill-conditioned as the dimension increases. Often the user increases the size of the model by adding terms which are functions of the original independent variables (as for example in polynomial models) and this often leads to increased correlations and ill-conditioning. For this reason, RAPIER was modified primarily by rearranging the storage of variables in COMMON blocks and performing all the regression calculations in double precision. This was done without losing any of the capabilities of the original program (in fact adding new options). The resulting version is called NEWRAP.

It may be of interest to some RAPIER users that in a number of sample calculations the major numerical inaccuracies arising in the regression calculations were not involved in the actual inversion of the $X'X$ matrix but in the calculation of the inner products which give

$$\hat{b} = (X'X)^{-1}(X'y)$$

Thus a major improvement might be made by computing inner products in double precision arithmetic and truncating to single precision answers without going to complete double precision arithmetic although the latter alternative would further increase the accuracy. As a matter of fact, the double precision inner product calculation is used in a different least-squares method proposed by Golub (ref. 3) which is reference 19 of Wampler's paper.

It should be pointed out that in estimation problems an alternative to the obvious step of more accurate routines is provided by Hoerl and Kennard (ref. 4). They present a technique called "ridge regression" which uses the method of minimum mean squared error estimation in place of minimum variance unbiased estimation. The ridge regression technique should have some definite appeal to statisticians, because it recognizes the fact that existence of ill-conditioned data indicates a problem which should be accounted for statistically as well as computationally. They do consider the problem of rejecting terms but their methods are not amenable to incorporation in NEWRAP in its present form.

A second reason for modifying the program was the desire to provide plots of the residuals as was strongly recommended in chapter 3 of Draper and Smith (ref. 5). With
the microfilm plotting capabilities provided by CINEMATIC (ref. 6) available at the
Lewis Research Center computer facility, this feature was also added to NEWRAP without significantly increasing printed output. CINEMATIC is a very specialized set of routines for the 7094/7044 DCS and 360/67 systems. If microfilm plotting is not available at other computer installations, the subroutines used in plotting may readily be changed to routines which produce line printer plots or CALCOMP plots however.

The RAPIER program used an algorithm for the coefficient calculations that inverted the correlation matrix and then converted this to the \((X'X)^{-1}\) matrix to calculate \(\hat{b}\). After inspection of several test cases, it seemed that this method did not improve the accuracy of the calculation of \((X'X)^{-1}\). Thus it was dropped and NEWRAP inverts \(X'X\) directly.

As with the RAPIER report, only the statistics and mathematics necessary to explain the program capabilities will be presented along with illustrative input and output listings and listings of the programs.

### SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>matrix</td>
</tr>
<tr>
<td>b</td>
<td>vector (column)</td>
</tr>
<tr>
<td>(b_i)</td>
<td>true regression coefficient</td>
</tr>
<tr>
<td>(\hat{b}_i)</td>
<td>estimated regression coefficient</td>
</tr>
<tr>
<td>(b_0)</td>
<td>constant term</td>
</tr>
<tr>
<td>(b_1, \ldots, b_J)</td>
<td>unknown parameters</td>
</tr>
<tr>
<td>C</td>
<td>correlation matrix</td>
</tr>
<tr>
<td>(C_{ij})</td>
<td>elements of (C)</td>
</tr>
<tr>
<td>D</td>
<td>indicator variable, equal to 0 if no (b_0) coefficient is estimated and equal to 1 if (b_0) is estimated</td>
</tr>
<tr>
<td>(E(x))</td>
<td>expected value of (x) (i.e., mean of (x) over all possible values of (x))</td>
</tr>
<tr>
<td>e</td>
<td>vector of observed values minus predicted values</td>
</tr>
<tr>
<td>(F_{a,d})</td>
<td>statistic distributed as variance ratio with (a) and (d) degrees of freedom</td>
</tr>
<tr>
<td>(f_j(z_1, \ldots, z_K))</td>
<td>term of regression equation</td>
</tr>
<tr>
<td>(H_0)</td>
<td>statistical hypothesis to be tested</td>
</tr>
</tbody>
</table>
$H_1$ alternate hypothesis to be accepted if $H_0$ is judged to be false

$J$ number of coefficients estimated, excluding $b_0$

$K$ number of independent variables observed

$k$ number of segments or cells in range of possible studentized residuals

LOF lack of fit

$M$ total number of independent and dependent variables

$M S(source)$ mean square due to source, where source is REG, RES, etc.

$N$ number of observations

NPDEG pooled degrees of freedom for replication error

$N(μ, σ^2)$ normal distribution with mean $μ$ and variance $σ^2$

$R$ number of sets of replicates

REG regression

REP replication

RES residual

$r_i$ number of replicates in set $i$

$S$ diagonal matrix

$S_c$ sum of squares correction if $D = 1$, and 0 if $D = 0$

$SSQ(source)$ sum of squares due to source, where source is REG, RES, etc.

$s_j$ elements of diagonal matrix

TOT total

$t_n$ statistic distributed as Student's $t$ with $n$ degrees of freedom

$V(x)$ variance of $x$, expected value of $(x - E(x))^2$

$W, X$ matrices

$w, x$ vectors (column)

$X_s$ stationary point of estimated quadratic surface

$x(J)$ $x_j$

$\bar{x}_{.j} = \frac{1}{N} \sum_{i=1}^{N} x_{ij}$

$y$ vector (column)

$Z_i$ studentized residual
**ESTIMATION OF BASIC LINEAR MODEL**

**BASIC LINEAR MODEL**

In multiple linear regression, a dependent or response variable $Y$ (such as temperature or pressure) measured on an object or experiment is assumed to be correlated with a function of one or more other variables $(z_1, \ldots, z_K)$ measured on the same object or experiment. This function includes a number of unknown parameters $(b_1, \ldots, b_J)$ and can be represented as

$$y = h(b_1, \ldots, b_J, z_1, \ldots, z_K) + \epsilon$$  \hspace{1cm} (1)

The only restriction imposed on this function is that it be linear in the parameters; that is, the function is of the form

$$y = \sum_{j=1}^{J} b_j f_j(z_1, \ldots, z_K) + \epsilon$$  \hspace{1cm} (2)

where $f_j(z_1, \ldots, z_K)$ is a TERM of the regression equation. (A TERM is a quantity which may be a variable or a function of a variable, e.g., $T$ is a TERM and $Z$, after it is defined as $Z = \log T$, is also a TERM.)

Suppose that there are $N$ observations of the dependent variable. Let the subscript $i$ indicate that the values are associated with the $i$\textsuperscript{th} observation; in particular, the value of the response variable $y_i$ would depend on the observed values of the variables $(z_{i1}, \ldots, z_{iK})$. Also, let the subscript $j$ denote the $j$\textsuperscript{th} term in the regression
model so that \( x_{ij} = f_j(z_{i1}, \ldots, z_{iK}) \) describes the transformations of the \( z_{i1}, \ldots, z_{iK} \) to produce the value of \( x_{ij} \) for the \( j \)th term at the \( i \)th observation.

The regression model can now be rewritten as

\[
y_i = b_1 x_{i1} + b_2 x_{i2} + \ldots + b_J x_{iJ} + \epsilon_i \quad i = 1, \ldots, N
\]  

where \( \epsilon_i \) denotes the difference between the observed value and the expected value of \( y_i \). For the \( N \) observations, it is convenient to write this regression model in matrix notation as \( y = Xb + \epsilon \) where

\[
y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}, \quad X = \begin{pmatrix} x_{11} & \cdots & x_{1J} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NJ} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_J \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_N \end{pmatrix}
\]
More often than not, the analyst feels the following model is more appropriate:

\[ y_i = b_0 + b_1x_{i1} + \ldots + b_Jx_{iJ} + \epsilon_i \quad i = 1, \ldots, N \]  

(5)

Let \( a_0 = b_0 + b_1\bar{x}_1 + \ldots + b_J\bar{x}_J \). Then, the result of adding this equation to and subtracting it from equation (5) and rearranging the terms is

\[ y_i = (b_0 + b_1\bar{x}_1 + \ldots + b_J\bar{x}_J) \]

\[ + b_1(x_{i1} - \bar{x}_1) + \ldots + b_J(x_{iJ} - \bar{x}_J) + \epsilon_i \quad i = 1, \ldots, N \]  

(6)

If, then, a dummy variable \( x_0 \) is introduced such that, for all values of \( i \), \( x_{i0} = 1.0 \), equation (6) may be written as

\[ y_i = a_0 x_{i0} + b_1(x_{i1} - \bar{x}_1) + \ldots + b_J(x_{iJ} - \bar{x}_J) + \epsilon_i \quad i = 1, \ldots, N \]  

(6a)

Equation (6a) now resembles equation (3) and may be written in matrix notation, similar to equation (4), as \( y = Xb + \epsilon \) where now

\[
\begin{bmatrix}
  y_1 \\
  \vdots \\
  y_N
\end{bmatrix}
= \begin{bmatrix}
  1.0 & x_{11} - \bar{x}_1 & \ldots & x_{1J} - \bar{x}_J \\
  1.0 & x_{21} - \bar{x}_1 & \ldots & x_{2J} - \bar{x}_J \\
  \vdots & \vdots & \ddots & \vdots \\
  1.0 & x_{N1} - \bar{x}_1 & \ldots & x_{NJ} - \bar{x}_J
\end{bmatrix}
\begin{bmatrix}
  a_0 \\
  b_1 \\
  \vdots \\
  b_J
\end{bmatrix}
+ \begin{bmatrix}
  \epsilon_1 \\
  \vdots \\
  \epsilon_N
\end{bmatrix}
\]

(7)
ESTIMATING \( \hat{b} \)

Equations (4) and (7) are similar in form and for \( N > J \) are an overdetermined set of linear equations. There will be some vector \( \hat{b} \) which is a "best" vector to use. If the vector \( \epsilon \) is composed of random variables \( \epsilon_i \) such that \( \mathbb{E}(\epsilon_i) = 0, \mathbb{V}(\epsilon_i) = \sigma^2 < +\infty \), and the \( \epsilon_i \) are uncorrelated, then as is well known, the method of least squares gives the linear minimum variance unbiased estimators \( \hat{b} \) for \( b \). And \( \hat{b} \) is given by

\[
\hat{b} = (X'X)^{-1}X'y
\]  

(8)

The matrix \( X'X \) divided by \( N - 1 \) is called the moment matrix of the experiment. The variance-covariance matrix of \( \hat{b} \) is given by

\[
\mathbb{V}(\hat{b}) = \sigma^2(X'X)^{-1}
\]  

(9)

It is important to note that when the form of equation (7) is used, \( X'X \) is given by

\[
X'X = 
\begin{pmatrix}
N & 0 & \ldots & 0 \\
0 & \sum_{i=1}^{N} (x_{i1} - \bar{x}_1)^2 & \ldots & \sum_{i=1}^{N} (x_{i1} - \bar{x}_1)(x_{iJ} - \bar{x}_J) \\
\vdots & \vdots & \ddots & \vdots \\
0 & \sum_{i=1}^{N} (x_{i1} - \bar{x}_1)(x_{iJ} - \bar{x}_J) & \ldots & \sum_{i=1}^{N} (x_{iJ} - \bar{x}_J)^2
\end{pmatrix}
\]  

(10)

This is seen to be symmetric and of the form

\[
X'X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}
\]

Hence,

\[
(X'X)^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}
\]
NEWRAP uses this relation to advantage by storing only the lower triangular part of $B$ and computing only the coefficients $b_1, \ldots, b_J$ by matrix manipulations. Then $b_0$ is given by the simple equation

$$b_0 = \bar{y} - \hat{b}_1 \bar{x}_1 - \hat{b}_2 \bar{x}_2 - \cdots - \hat{b}_J \bar{x}_J$$

where $\bar{y} = \sum y_i / N = \bar{a}_0$. It can also be shown that

$$V(\hat{b}_0) = V(\bar{y}) + V(\hat{b}'\bar{x}) = \left[\frac{1}{N} + \bar{x}'(X'X)^{-1} \bar{x}\right] \sigma^2$$

$$\text{COV}(\hat{b}_0, \hat{b}) = -(X'X)^{-1} \bar{x} \sigma^2$$

When there is no $b_0$ term in the regression model,

$$X'X = \begin{bmatrix}
\sum_{i=1}^{N} x_{i1}^2 & \sum_{i=1}^{N} x_{i1}x_{i2} & \cdots & \sum_{i=1}^{N} x_{i1}x_{iJ} \\
\sum_{i=1}^{N} x_{i2}x_{i1} & \sum_{i=1}^{N} x_{i2}x_{i2} & \cdots & \sum_{i=1}^{N} x_{i2}x_{iJ} \\
\cdots & \cdots & \cdots & \cdots \\
\sum_{i=1}^{N} x_{iJ}x_{i1} & \sum_{i=1}^{N} x_{iJ}x_{i2} & \cdots & \sum_{i=1}^{N} x_{iJ}x_{iJ}
\end{bmatrix}$$

(12)

Comparing this to equation (10) shows this form of $X'X$ to be similar to the lower right submatrix in equation (10). This similarity is used to simplify notation by assuming that $X'X$ represents either the form of equation (12) or the lower right portion of equation (10) and considering the calculation of $b_0$ as a special case. Thus, further reference to $b$ implies

$$b = \begin{pmatrix}
b_1 \\
\cdot \\
\cdot \\
b_J
\end{pmatrix}$$

There are two different methods of computing the regression coefficients which may
be used in NEWRAP. The first method uses bordering (ref. 7) on the full $X'X$ matrix. If $X'X$ is a nearly singular matrix, there may be problems with accuracy resulting in overflows or underflows causing execution to terminate without any results being printed. The second method uses a method of bordering which enters one term at a time into the model equation. After each term is entered, a full regression analysis is printed. Typically, if $X'X$ is nearly singular, a number of terms will have been added to the model before the results become unreliable or cause execution to be terminated. Thus, at least a partial analysis of the full model is available to aid in selection of further models to submit. After all the terms have been entered, the program then switches to the procedure which inverts the appropriate full $X'X$ matrix at each stage for further analyses.

The use of the bordering method leads to a large volume of printed output and is not recommended as a standard procedure. Through use of CRSPLT as a preregression analysis program it may be easier to determine if bordering should be used. CRSPLT can also help indicate the order of arrangement of the terms of the model so that those thought to be most important can be entered into the model first.

Also note that the individual observations may be weighted to perform a weighted regression analysis. NEWRAP permits the use of weights (ref. 5). In this case, the $X'X$ and $X'y$ matrices take the following form:

$$X'X = \begin{pmatrix}
\sum_{i=1}^{N} [(x_{i1} - \bar{x}_1)^2 w_i] & \cdots & \sum_{i=1}^{N} [w_i(x_{i1} - \bar{x}_1)(x_{iJ} - \bar{x}_J)] \\
\vdots & \ddots & \vdots \\
\sum_{i=1}^{N} [(x_{iJ} - \bar{x}_J)(x_{i1} - \bar{x}_1)w_i] & \cdots & \sum_{i=1}^{N} [(x_{iJ} - \bar{x}_J)^2 w_i]
\end{pmatrix}$$

(12a)

$$X'y = \begin{pmatrix}
\sum_{i=1}^{N} x_{i1}y_i w_i \\
\vdots \\
\sum_{i=1}^{N} x_{iJ}y_i w_i
\end{pmatrix}$$
CORRELATION MATRIX

Another matrix of interest both computationally and statistically in the correlation matrix $C$. The elements of $C$, which are denoted $C_{ij}$, are the sample correlation coefficients between the terms $X_i$ and $X_j$. These are

$$C_{ij} = \sum_{l=1}^{N} \frac{(x_{l1} - \bar{x}_1)(x_{lj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^{N} (x_{ki} - \bar{x}_1)^2} \sqrt{\sum_{k=1}^{N} (x_{kj} - \bar{x}_j)^2}}$$

(13)

and all these numbers are between 1.0 and -1.0.

The calculation of $C$ can be expressed in matrix notation conveniently by defining a diagonal matrix $S = \text{diag}(s_1, s_2, \ldots, s_J)$ with elements

$$s_j = \frac{1.0}{\sqrt{(X'X)_{jj}}} \quad j = 1, \ldots, J$$

(14)

Then

$$C = S(X'X)S$$

(15)

It may also be that the independent variables are random variables. Then $X'X$ divided by $N - 1$ represents the sample variance-covariance matrix and $C$ the sample correlation matrix. If the independent variables are considered to be from a multivariate distribution, it is useful in some cases to consider the eigenvalues and eigenvectors of $X'X$. For these reasons, NEWRAP includes options to compute and print these quantities. These may also be computed and printed through use of the CRSPLT program.

ESTIMATING $\sigma^2$

For any regression model $y = Xb + \epsilon$, there are possibly two methods of estimating $\sigma^2$. First, if the assumed regression model is, in reality, the true model, it is well known that an unbiased estimator is given by
Second, where there are replicated data points, another estimator of \( \sigma^2 \), depending only on \( V(\epsilon_i) = \sigma^2 \) for all \( i \) and not on the validity of the assumed model, is the pooled mean squares computed from the replicated data points.

Assume the observations are grouped into replicate sets in sequence. Let \( R \) be the number of sets of replicates and \( r_i \) be the number of replicates in the \( i^{th} \) replicate set. Let

\[
SSQ(i) = \sum_{k=r^*+1}^{r^*+r_i} (y_k - \bar{y}_i)^2
\]  

(17)

where

\[
r^* = \sum_{j=1}^{i-1} r_j
\]

It is assumed \( y_n \) is from the \( i^{th} \) replicate set and \( \bar{y}_i \) is calculated only from those \( y_n \) in the \( i^{th} \) replicate set. Then define the pooled sum of squares due to replication as

\[
SSQ(REP) = \sum_{i=1}^{R} SSQ(i)
\]

and the pooled degrees of freedom as \( NPDEG = \sum_{i=1}^{R} (r_i - 1) \). The second estimator of \( \sigma^2 \) becomes

\[
\sigma^2_{REP} = \frac{SSQ(REP)}{NPDEG}
\]

(18)

It can be shown (ref. 5, p. 26) that the sums of squares due to residuals can be partitioned into a component due to replication and a component due to lack of fit; that is,

\[
SSQ(RES) = SSQ(LOF) + SSQ(REP)
\]  

(19)
This partitioning is used later to determine the estimate of $\sigma^2$ to use in tests of hypotheses.

**HYPOTHESIS TESTING**

**NORMALITY OF $\varepsilon$**

As stated before, the only assumption necessary for $\hat{\beta}$ to be a linear minimum variance unbiased estimator is that $E(\varepsilon_i) = 0.0$, $V(\varepsilon_i) = \sigma^2 < +\infty$, and $\varepsilon_i$ be uncorrelated. If it can further be assumed that $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, a number of standard tests become available. NEWRAP computes a chi-squared statistic which can be used as an approximate test.

Under the hypothesis $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, the studentized residuals defined by

$$Z_i = \frac{y_i - \hat{y}_i}{\hat{\sigma}} = \frac{e_i}{\hat{\sigma}}$$

will be distributed as Student's $t$ with the degrees of freedom associated with the estimate $\sigma$. If the degree of freedom is 30 or more, the $t$ distribution is very close to the normal.

The range of possible studentized residuals is $(-\infty, +\infty)$ and may be divided into $k$ segments or cells each with probability $p_i$, so that each segment will have $Np_i$ as the expected number of observations falling into it. Let $n_i$ denote the number of studentized residuals in the $i^{th}$ cell. Then a chi-squared goodness-of-fit statistic may be calculated as

$$\chi^2_{k-1} = \sum_{i=1}^{k} \frac{(n_i - Np_i)^2}{Np_i}$$

NEWRAP computes this statistic by using an even number of cells greater than or equal to four and less than or equal to 20, such that the expected numbers of observations per cell is five or more. This statistic is not computed when there are less than 20 observations. The bounding values for the $i^{th}$ cell are $Z_{i-1}, Z_i$ where $F(Z_i) = (i \cdot k)/N$ and $F(Z)$ is the cumulative normal distribution function. Then each cell has the same expected number of observations, say $f = N/k$. Then
There is a point to be made concerning the chi-squared calculations. The validity of the use of the chi-squared statistic in a test depends upon the residuals forming a sample of independent and identically distributed random variables. This is not usually the case for regression residuals. Although the tail probabilities of the chi-squared tests might be in error, they should still be able to tell the statistician whether one intended normalizing transformation was more successful than another.

**ANALYSIS OF VARIANCE TABLE**

For most hypothesis testing of the regression model, it is convenient to summarize the available information in an Analysis of Variance (ANOVA) table, as follows:

<table>
<thead>
<tr>
<th>Source</th>
<th>Sums of squares</th>
<th>Degrees of freedom</th>
<th>Mean squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>S_SQ(REG) = ( \hat{b}'X'y - S_c^a )</td>
<td>( J )</td>
<td>MS(REG) = S_SQ(REG) ( J )</td>
</tr>
<tr>
<td>Residual</td>
<td>S_SQ(RES) = y'y - ( \hat{b}'X'y )</td>
<td>( N - J - D^b )</td>
<td>MS(RES) = S_SQ(RES) ( N - J - D )</td>
</tr>
<tr>
<td>Total</td>
<td>S_SQ(TOT) = y'y - S_c</td>
<td>( N - D )</td>
<td></td>
</tr>
</tbody>
</table>

\[ a_{S_c} = \begin{cases} 0 & \text{if no } b_0 \text{ coefficient is estimated.} \\ N_y^2 & \text{if a } b_0 \text{ coefficient is estimated.} \end{cases} \]

\[ b_D = \begin{cases} 0 & \text{if no } b_0 \text{ coefficient is estimated.} \\ 1 & \text{if } b_0 \text{ is estimated.} \end{cases} \]

If there are replicated data points, another ANOVA table can be constructed to show the separation of the residual sums of squares into components from lack of fit and replication, as in the following table:
Sums of squares

<table>
<thead>
<tr>
<th>Source</th>
<th>Sums of squares</th>
<th>Degrees of freedom</th>
<th>Mean squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of fit</td>
<td>SSQ(LOF) = SSQ(RES) - SSQ(REP)</td>
<td>N - J - D - NPDEG</td>
<td>MS(LOF) = SSQ(LOF) \cdot (N - J - D - NPDEG)</td>
</tr>
<tr>
<td>Replication</td>
<td>SSQ(REP)</td>
<td>NPDEG</td>
<td>MS(REP) = SSQ(REP) \cdot NPDEG</td>
</tr>
<tr>
<td>Residual</td>
<td>y'y - b' X y</td>
<td>N - J - D</td>
<td></td>
</tr>
</tbody>
</table>

**CHOICE OF ESTIMATOR FOR $\sigma^2$**

As mentioned previously, there are two possible methods of estimating $\sigma^2$ depending on whether there are replicated data points. This is true for any given model equation. When the backward rejection option of NEWRAP is used, there is no longer one hypothetical model but a series of models. Thus, there is the choice of estimator for $\sigma^2$ to be made after each rejection of a term in the previous model.

As an example, consider the model

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \epsilon$$  \hspace{1cm} (20)

with replicated data points. The first step is to estimate $b_0$, $b_1$, $b_2$, and $b_3$. There will then be the estimators $\hat{\sigma}^2_{RES(J)}$ and $\hat{\sigma}^2_{REP}$. If the model in equation (20) has not left out any important terms, $\hat{\sigma}^2_{RES(J)}$ as well as $\hat{\sigma}^2_{REP}$ is a valid estimator.

The ratio $F = MS(LOF)/MS(REP)$ can be used to test the hypothesis that there is no lack of fit, where $F \sim F_{a,d}$ with $a = N - J - D - NPDEG$ and $d = NPDEG$ degrees of freedom. If the test accepts the hypothesis of no lack of fit, MS(RES) is a pooled estimate of $\sigma^2$ with more degrees of freedom. But there is the possibility that the hypothesis was accepted as a result of random fluctuation when there really is some lack of fit; that is, there is the possibility that $\hat{\sigma}^2_{RES(J)}$ is a biased estimator. If lack of fit is not concluded to be significant, the decision to pool or not is usually made on the basis of the number of degrees of freedom for replication. If this is "large" (no definition of large is given herein), $\hat{\sigma}^2_{REP}$ is used. If "small," the pooled estimate $\hat{\sigma}^2_{RES(J)}$ is used.

In testing equation (20), should it be decided that $b_3$ is not significantly different from zero (see section t-TESTS), the coefficients of the following model would be estimated:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \epsilon$$
From this model there is an estimate $\hat{\sigma}^2_{\text{RES}(J-1)}$. This estimate could also be biased since $b_3$ may be small but nonzero and the decision of $b_3 = 0$ may have been due to the low power of the test.

At the first step, the lack of fit can be considered a random sample of an infinite possibility of biases. But the biases due to pooling mean squares after rejecting terms can be considered to be systematic biasing. In such a case the use of Cochran's test for "the largest of a set of estimated variances as a fraction of their total" might be appropriate.

NEWRAP provides three strategies of pooling estimates for use in the decision procedure:

1. Never pool. This is usable only when there are replicated data points. The estimator used in all t-tests is $\hat{\sigma}^2_{\text{REP}}$.

2. Pool initial residual. This will pool the lack of fit and replication (if any) from the first model only. Additional mean squares due to rejected terms will be ignored.

3. Always pool. This strategy will always use $\hat{\sigma}^2_{\text{RES}(J-i)}$ for the model with $i$ rejected terms.

Wherever a $\hat{\sigma}$ or $\hat{\sigma}^2$ is indicated, NEWRAP always uses the value calculated according to the strategy chosen by the user.

**TEST OF OVERALL REGRESSION**

One of the first tests usually applied to a regression model is the test of the overall significance of the model. In the notation of hypothesis testing this is stated $H_0$: $b = 0$; $H_1$: $b \neq 0$ where

$$
\begin{pmatrix}
 b_1 \\
 . \\
 . \\
 . \\
 b_J
\end{pmatrix}
$$

The statistic for this test is $F = \frac{\text{MS(REG)}}{\hat{\sigma}^2}$. The $F \sim F_{a, d}$ with $a = J - D$, and $d$ equals the degrees of freedom associated with $\hat{\sigma}^2$.

Another useful statistic for judging the significance of overall regression is $R^2 = \frac{\text{SSQ(REG)}}{\text{SSQ(TOT)}}$. The sampling distribution of $R$ does not lend itself to very simple tests except in the case of $H_0$: $b = 0$. The main value of $R^2$ is that it is a
number in the range 0 to 1 and $100 \, R^2$ is a measure of the percentage of variation in the $y$ values that is accounted for by the regression model.

**t-TESTS**

In many cases, the regression model contains terms whose estimated coefficients are "small." This may be an indication that the term does not have a real effect on the dependent variable and that the estimate is nonzero due to random sampling variation. If this is true, it is desirable to delete the term from the regression model. A test statistic for deciding this is

$$
t = \frac{\hat{b}_i}{\sqrt{\hat{\sigma}^2 (X'X)^{-1}_{ii}}} \tag{21}
$$

where $(X'X)^{-1}_{ii}$ denotes the $i^{th}$ diagonal element of the $(X'X)^{-1}$ matrix. The statistic $t \sim t_{N-J-D}$. An equivalent test statistic is

$$
F = t^2 = \frac{\hat{b}_i^2}{\hat{\sigma}^2 (X'X)^{-1}_{ii}} \tag{22}
$$

where $F \sim F_{1, N-J-D}$. This is often referred to (ref. 5) as the partial F-test. The quantity $\hat{b}_i^2 / [(X'X)^{-1}]_{ii}$ is called the additional sum of squares due to $b_i$, if $x_i$ were last to enter the equation. NEWRAP computes and prints the $t$-statistics, the probability associated with the interval $(-t, t)$, and the additional sums of squares for each term.

This particular test is the basis for the rejection option of NEWRAP. The analyst initially chooses which $\hat{\sigma}^2$ estimator to use by the choice of strategy. Then the analyst may choose a confidence level which all coefficients must meet. For example, suppose a confidence level of 0.900 is chosen. The $t$-statistic is then computed for each coefficient, and the coefficient with minimum $|t|$ is identified. If $\min |t| > t_{N-J-D, 0.950}$, all terms are concluded to be significant at the 0.1 level (or 90.0 percent level of confidence). If $\min |t| < t_{N-J-D, 0.950}$, the term corresponding to the minimum $|t|$ is dropped from the hypothetical model, and the regression is recomputed. This process is repeated until all remaining coefficients are significant at the specified level of probability. This procedure can be overridden by an option which allows certain specified terms of the model to be retained regardless of the significance of the coefficient. Ken-
nedy and Bancroft (ref. 8) present a study indicating the backward deletion method is slightly more efficient than forward selection in special situations.

**PREDICTING VALUES FROM ESTIMATED REGRESSION EQUATION**

Regression equations are often used to predict an estimated response at some condition of the independent variables. Useful estimates of parameters to know are the variance of the regression equation and the variance of a single further observation at the desired combination of the independent variables.

Let $x' = (x_1, \ldots, x_J)$ denote the vector of independent variables at which a prediction is desired. Let $x^* = x - \bar{x}$. Let $\hat{\sigma}^2_{\mu \cdot x}$ denote the estimated variance of the regression equation at $x$. Let $\hat{\sigma}^2_{y \cdot x}$ denote the estimated variance of a single further observation at $x$. Then,

$$\hat{\sigma}^2_{\mu \cdot x} = \hat{\sigma}^2 \left[ \frac{D}{N} + x^*(X'X)^{-1} x^* \right] \quad (23)$$

$$\hat{\sigma}^2_{y \cdot x} = \hat{\sigma}^2 \left[ 1.0 + \frac{D}{N} + x^*(X'X)^{-1} x^* \right] \quad (24)$$

where, as before, $D = 1$ if a $b_0$ coefficient is estimated and $D = 0$ if a $b_0$ coefficient is not estimated. The quantity $s = \hat{\sigma}_{\text{RES}(J)}$ is called the standard error of estimate and often is used as a simple approximation to $\hat{\sigma}_{y \cdot x}$. This approximation is close if $N$ is very large and $x = \bar{x}$, in which case,

$$\hat{\sigma}^2_{y \cdot \bar{x}} = s^2 \left( 1.0 + \frac{D}{N} \right) \approx s^2$$

When $x \neq \bar{x}$, this may be a poor approximation. NEWRAP accepts input vectors $x$ and computes $\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \ldots + \hat{b}_J x_J$, as well as $\hat{\sigma}_{\mu \cdot x}^2, \hat{\sigma}_{\mu \cdot x'}^2, \hat{\sigma}_{y \cdot x}^2, \hat{\sigma}_{y \cdot x'}^2$ and the standard error of estimate.
NEWRAP PROGRAM

USERS GUIDE TO NEWRAP INPUT

Illustrative Regression Problem Requiring No Transformations

The illustrative example is described in chapter 7 of reference 5. The data is reproduced in table I. Figure 1 presents this data in a sample input form. The basic model to be fitted is

\[ y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 \]  

(25)

<table>
<thead>
<tr>
<th>Unit number</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( x_3 )</th>
<th>( x_4 )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-75</td>
<td>0</td>
<td>0</td>
<td>-65</td>
<td>1.4</td>
</tr>
<tr>
<td>2</td>
<td>175</td>
<td>0</td>
<td>0</td>
<td>150</td>
<td>26.3</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>-65</td>
<td>...</td>
<td>29.4</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>165</td>
<td>-6</td>
<td>9.7</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>150</td>
<td>32.9</td>
</tr>
<tr>
<td>10</td>
<td>-75</td>
<td>-75</td>
<td>0</td>
<td>150</td>
<td>26.4</td>
</tr>
<tr>
<td>11</td>
<td>175</td>
<td>175</td>
<td>0</td>
<td>-65</td>
<td>8.4</td>
</tr>
<tr>
<td>14</td>
<td>-75</td>
<td>-75</td>
<td>-65</td>
<td>150</td>
<td>28.4</td>
</tr>
<tr>
<td>15</td>
<td>175</td>
<td>175</td>
<td>165</td>
<td>-65</td>
<td>11.5</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0</td>
<td>-65</td>
<td>-65</td>
<td>1.3</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0</td>
<td>165</td>
<td>150</td>
<td>21.4</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>-75</td>
<td>-65</td>
<td>-65</td>
<td>26.4</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>175</td>
<td>165</td>
<td>150</td>
<td>22.9</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-65</td>
<td>3.7</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>-75</td>
<td>0</td>
<td>150</td>
<td>26.5</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>-75</td>
<td>0</td>
<td>150</td>
<td>23.4</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>-75</td>
<td>0</td>
<td>150</td>
<td>26.5</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>175</td>
<td>0</td>
<td>-65</td>
<td>5.8</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>175</td>
<td>0</td>
<td>-65</td>
<td>5.8</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>175</td>
<td>0</td>
<td>-65</td>
<td>5.8</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>-75</td>
<td>-65</td>
<td>150</td>
<td>28.8</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>-75</td>
<td>-65</td>
<td>150</td>
<td>26.4</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>175</td>
<td>165</td>
<td>-65</td>
<td>11.8</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>175</td>
<td>165</td>
<td>-65</td>
<td>11.4</td>
</tr>
</tbody>
</table>
SAMPLE INPUT

FORTRAN STATEMENT

15 SAMPLE NEWRAP PROBLEM 2
DATA IS FROM DRAPER AND SMITH, APPLIED REGRESSION ANALYSIS (REFERENCE 4 OF NEWRAP REPORT)

CHAPTER 7
INITIAL MODEL EQUATION IS

\[ Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + \text{ERR} \]

\[ Y = \text{CHAMBER PRESSURE} \]
\[ x_1 = \text{THERMAL LEVEL} \]
\[ x_2 = \text{VIBRATION LEVEL} \]
\[ x_3 = \text{DROP (SHOCK)} \]
\[ x_4 = \text{STATIC FIRE} \]

RESIDUALS ARE BEEN REQUESTED TO BE PUNCHED (FOR CERFILP PROGRAM)
FOR RESIDUAL PLOTTING ANALYSES

1 4

TA IS FROM DRAPER AND SMITH, APPLIED REGRESSION ANALYSIS.

CHAPTER 7

SAMPLE INPUT (Concluded)

FORTRAN STATEMENT

UNIT NO. 1 75 0 66 3
2 75 0 150 28.8
3 75 0 150 28.9
8 0 0 150 10.5

Figure 1 - Sample input form.
The preceding model requires no transformations of the tabulated data for the dependent or independent variables. Suitable input statements are also given in figure 1.

A subsequent example will illustrate the requirements on the input cards when transformations are involved.

**Detailed Description of Input Cards**

This section of the report describes the input cards as classified into nine sets according to table II.

<table>
<thead>
<tr>
<th>Set number</th>
<th>Name of set</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IDENTIFICATION</td>
<td>Identify and describe problem</td>
</tr>
<tr>
<td>2</td>
<td>PROBLEM SIZE</td>
<td>Define problem size</td>
</tr>
<tr>
<td>3</td>
<td>LOGIC</td>
<td>Specify general logical controls</td>
</tr>
<tr>
<td>4</td>
<td>MODEL</td>
<td>Define model equation</td>
</tr>
<tr>
<td></td>
<td>(a) MODEL SIZE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) TERMS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(c) TRANSFORMATIONS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(d) CONSTANTS</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>REJECTION</td>
<td>Backward rejection controls</td>
</tr>
<tr>
<td>6</td>
<td>REPLICALES</td>
<td>Identify replicated data</td>
</tr>
<tr>
<td>7</td>
<td>FORMAT</td>
<td>Give data format</td>
</tr>
<tr>
<td>8</td>
<td>DATA</td>
<td>Input observed data</td>
</tr>
<tr>
<td>9</td>
<td>PREDICTIONS</td>
<td>Predicted values data</td>
</tr>
</tbody>
</table>

The model equation is defined by set 4 of table II. An example illustrating the use of one blank card for input set 4 which can be used for simple linear regressions is presented by figure 1 and table I. A second example illustrating the use of the set of MODEL cards in the presence of prior constants and transformations will be given at the end of this section of the report. A pictorial representation of an input deck is given by figure 2.
Figure 2. - Sample input deck. (Asterisk denotes the card is optional and its use depends upon data input on previous cards.)
Figure 2. - Concluded.
Details of the input cards are as follows:

1. **IDENTIFICATION (I, IDENT)(I2, 13A6):** IDENT is Hollerith data used to identify the problem. I indicates the number of additional cards to read for further identification or description (columns 1 to 78).

2. **PROBLEM SIZE (NOVAR, NODEP, NOTERM, NOOB, NTKEEP)(3I4, I5, I4)**
   - **NOVAR** Number of input independent variables (number of \( z \)'s in eq. (2))
   - **NODEP** Number of input dependent variables
   - **NOTERM** Number of terms in model equation (number of \( x \)'s in eq. (3)). Note that \( b_0 \) is not counted as a term.
   - **NOOB** Number of observations
   - **NTKEEP** First NTKEEP independent terms of model equation will be retained in model regardless of significance level

3. **LOGIC:** One card with nine one-column fields
   - **BZERO** \( b_0 \) term appears in model equation (T or F)
   - **IFTT** t-statistics and their descriptive confidence levels are to be computed (T or F)
   - **IFWT** Weight of 1.0 is applied to all observations (T or F). If this is F, see input sets 7 and 8 for further information.
   - **IFCHI** Compute and plot residuals (T or F)
   - **STORYX** Calculate eigenvalues and eigenvectors of \( X'X \) (T or F)
   - **IFSSR** Model shall be increased by one term at a time using bordering method for matrix inversion (T or F)
   - **ECONMY** Use economy version of output (T or F). NEWRAP does not print \( X'y \), \( (X'X)^{-1} \), or C when set to F.
   - **ISTRAT** Pooling strategy is 1, 2, or 3:
     1. Never pool. Use replication error as estimate of error. If 1 is selected and no replication is found, strategy 3 is used.
     2. Pool initial residual only.
     3. Pool all residuals.
PUNCH

Punch residuals and predicted values (T or F). If T, observation number is punched and then residuals and predicted values are punched in (I6, 4E16.8/(6X, 4E16.8)) format in pairs (observation number, e₁, ŷ₁, e₂, ŷ₂, etc.).

(4) MODEL: The MODEL cards are used to manipulate the observed input data, supplied by input set 8, into the form of the desired model equation. There are four subsets of this input set 4, namely, MODEL SIZE, TERMS, TRANSFORMATIONS, and CONSTANTS, of which the latter three are used only in the development of complex models.

If a simple linear model is being analyzed, the MODEL SIZE card is left blank, indicating that the number of transformations is zero and the number of constants to be read in is zero. In this case, the TERMS, TRANSFORMATIONS, and CONSTANTS cards of this input set are not expected by the program, and the program assumes the independent and dependent variables are arranged on the input cards of input set 8 as

\[ x_1, x_2, \ldots, x_J, y_1, \ldots, y_{\text{NODEP}} \]

where NODEP is the number of dependent variables.

If a weighting factor other than 1.0 is to be used (i.e., if item 3 of the LOGIC card contains an F), the value of the weighting factor for each observation must appear as the last item in the list, so that in this case the data for each observation is entered on the cards as

\[ x_1, x_2, \ldots, x_J, y_1, \ldots, y_{\text{NODEP}}, WT \]

If the weighting factor is identically 1.0, NEWRAP reads a total of M numerical values for each observation, where M is the sum of the number of independent and dependent variables. The variables are stored consecutively in an array called X, beginning with location 01 and ending with location M. If the weighting factor is not identically 1.0, then M + 1 numerical values are read for each observation, but the last value, being the weighting factor, is treated and stored separately. The data in X are used with their appropriate weighting factors to cumulatively create \( X'X \) and \( X'y \) as shown in equations (12a).

The remaining discussion of this set explains the use of transformations and/or constants to build more complex models. Therefore, the reader who does not immediately need a complex model may skip this material and proceed directly to the description of input set 5.

As mentioned previously, there are up to four subsets of the MODEL cards. Their purpose is to give the structure of the model equation and thereby specify the initial
operations to be performed on the input data. As used here, CONSTANTS means any numerical value specified to be in the model equation in advance of parameter estimation. These numerical values are read from the CONSTANTS cards.

Also, the word TRANSFORMATIONS is to be interpreted as the operations performed on the input data (read from data cards) to compute the $f_j$ values (eq. (2)) of the model equation. The structure of these functions (and of any transformations of the dependent variables) is read from the TRANSFORMATION cards. Finally, the word TERMS is to be interpreted as the computed results of the operations specified by the transformation (including any operations that leave the input data unchanged). The results of the TRANSFORMATIONS are stored in an array CON, and the TERMS cards designate the order of the relative locations in CON where the final values for the terms of the model equations are to be found.

The four subsets, MODEL SIZE, TERMS, TRANSFORMATIONS, and CONSTANTS, will be described in detail now. Also, at the end of the description of this input set, a summary of these cards, with the formats used, is given for convenience.

The MODEL SIZE card specifies NTRANS and KONNO(214) where

\begin{align*}
NTRANS & \quad \text{Number of transformations that will be performed} \\
KONNO & \quad \text{Number of constants that will be read in which are required to specify model equation}
\end{align*}

If the number of transformations is zero, and therefore, the number of constants is zero, the TERMS, TRANSFORMATIONS, and CONSTANTS cards are not expected by the program. This being a simple linear model case, only the MODEL SIZE card, which can be blank, is necessary in this subset, but the values for the observations which are provided in input set 8 must conform to the order as specified in the first three paragraphs describing this input set.

When, however, a more complex model is desired, information must be supplied instructing the program as to (1) where to find the values for the TERMS of the equation, (2) how to create the terms from the variables and the constants, and (3) what the values of the constants are. This information is supplied on the TERMS, TRANSFORMATIONS, and CONSTANTS cards.

The numerical values to be used in the transformations are stored in two arrays called X and CON. The transformations always require that an operator (some value from CON) performs an operation (see table III) on an operand (some specified value from X) to produce a result which will be stored in CON. Thus CON serves two purposes. First, if the number of constants (KONNO) specified on the MODEL SIZE card is nonzero, that many constants will be read from the CONSTANTS cards and stored in CON beginning with location 01. If the number of constants is zero, a CONSTANTS card is not expected by the program. Second, all intermediate and final results of
TABLE III. - OPERATIONS\(^a\) AND CODE NUMBERS

[X indicates a value from X and C a value from CON.]

<table>
<thead>
<tr>
<th>Operation code (OP)</th>
<th>Resulting operation</th>
<th>Operation code (OP)</th>
<th>Resulting operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>No operation</td>
<td>16</td>
<td>1.0 SQRT(X)</td>
</tr>
<tr>
<td>01</td>
<td>N + C</td>
<td>17</td>
<td>C**X</td>
</tr>
<tr>
<td>02</td>
<td>X*C</td>
<td>18</td>
<td>10.0**X</td>
</tr>
<tr>
<td>03</td>
<td>C/X</td>
<td>19</td>
<td>SINH(X)</td>
</tr>
<tr>
<td>04</td>
<td>EXP(X)</td>
<td>20</td>
<td>COSH(X)</td>
</tr>
<tr>
<td>05</td>
<td>X**C</td>
<td>21</td>
<td>(1.0-COS(X)).2.0</td>
</tr>
<tr>
<td>06</td>
<td>ALOG(X)</td>
<td>22</td>
<td>ATAN(X)</td>
</tr>
<tr>
<td>07</td>
<td>ALOG10(X)</td>
<td>23</td>
<td>ATAN2(X,C)</td>
</tr>
<tr>
<td>08</td>
<td>SIN(X)</td>
<td>24</td>
<td>X**2</td>
</tr>
<tr>
<td>09</td>
<td>COS(X)</td>
<td>25</td>
<td>X**3</td>
</tr>
<tr>
<td>10</td>
<td>SIN(π<em>C</em>X)</td>
<td>26</td>
<td>ARCSIN(SQRT(X))</td>
</tr>
<tr>
<td>11</td>
<td>COS(π<em>C</em>X)</td>
<td>27</td>
<td>2.0<em>π</em>X</td>
</tr>
<tr>
<td>12</td>
<td>1.0/X</td>
<td>28</td>
<td>1.0.(2.0<em>π</em>X)</td>
</tr>
<tr>
<td>13</td>
<td>EXP(C/X)</td>
<td>29</td>
<td>ERF(X)</td>
</tr>
<tr>
<td>14</td>
<td>EXP(C;X**2)</td>
<td>30</td>
<td>GAMMA(X)</td>
</tr>
<tr>
<td>15</td>
<td>SQRT(X)</td>
<td>31</td>
<td>X C</td>
</tr>
</tbody>
</table>

\(^a\)All function names and operations are consistent with FORTRAN IV mathematical subroutines.

Transformations are also stored in CON as specified on the TRANSFORMATION cards. The TERMS card then specifies which of the locations in CON finally contain the values needed to construct the \(X'X\) and \(X'y\) matrices. After all the transformations have been performed on an observation, the contents of the relative locations of the CON array specified on the TERMS card are moved back to X in consecutive locations beginning with location 01.

Note especially that CONSTANTS data are stored in CON from location 01 through KONNO. Thus, if a transformation specifies that a result is to be placed in any of these locations, the result will replace the constant, so that further operations on subsequent transformations would use the new value stored instead of the constant value to which it was initialized. Care should be taken, therefore, that the results of the transformations be stored in relative locations greater than KONNO.

Each transformation code is made up of four subfields of two card columns each, with the following interpretation:
Thus, subfield 1 always references the X array, and subfields 3 and 4 reference the CON array. The result of every transformation is a term which is stored in the designated location of the CON array, with the added feature that, if the term is stored in relative location 61 or beyond, it is also stored in the parallel location in the X array. This is illustrated by the arrows in figure 3. This feature allows successive transformations to be performed more easily.

The OP (operation codes) are tabulated in table III. The transformation with OP = 00 is simply an identity transformation. This transfers data from X to CON so that when terms are selected there is a value available in CON that can be moved back to X.

When there are no transformations, NEWRAP assumes the first NOTERM values on a data card are the independent variables and the last NODEP values are the depend-

<table>
<thead>
<tr>
<th>Subfield</th>
<th>Interpretation</th>
<th>Relative location in X</th>
<th>Relative location in CON</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Operand</td>
<td>Relative location</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Operation (OP)</td>
<td>Arithmetic operation</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Operator</td>
<td>Relative location</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Result</td>
<td>Relative location</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. - Map of X and CON arrays. Data transferred into any location of CON array beyond location 60 are immediately duplicated in same relative location in X array.
ent variables. When transformations are used, this convention need not hold for the raw input data but instead holds for the terms on the TERMS card. Thus, the first NOTERM values input on the TERMS card indicate the locations of the CON array which correspond to the independent variables and the last NODEP values indicate the locations of the CON array which correspond to the dependent variables. If the analyst desires to force certain terms to remain in the model regardless of their significance, these terms must be the first terms of the model. Then if the input for NTKEEP of the PROBLEM SIZE card is not zero, the first NTKEEP terms of the model will be retained.

The complete sequence of MODEL cards and the formats used are summarized as follows:

(a) **MODEL SIZE** (NTRANS, KONNO)(214): NTRANS specifies the number of transformations required and KONNO the number of CONSTANTS involved. NTRANS may not be greater than 100 and KONNO not greater than 60. If NTRANS = 0, the following three sets are skipped and the program goes directly to input set 5.

(b) **TERMS** (4012): One or more cards as necessary, using two-column fields to denote the relative locations of the CON array containing the final values for the terms to be used and the order in which they enter into the model equation. The number of terms used is specified on the PROBLEM SIZE card.

(c) **TRANSFORMATIONS** (4012): As many cards as necessary containing up to 10 transformation instructions per card. Each transformation instruction is composed of four two-column fields. See table III for the list of available transformations.

(d) **CONSTANTS** (5E15.7): As many cards as necessary containing the number of CONSTANTS as specified by KONNO. Up to 60 CONSTANTS may be specified. If KONNO = 0, these cards are not expected by the program.

(5) **REJECTION** (DELETE, P)(L1, F3.3): If DELETE is set to T, the backward rejection option is used and the desired level of confidence is given by P. The P value is written without a decimal point so that a 95 percent confidence level is indicated by a 950, a 99.9 percent level as 999, and so forth.

(6) **REPLICATION** (REPS)(L1): If REPS is F, the program skips to set 7 and assumes there is no replicated data. If REPS is set to T, then more cards are read in 2014 format specifying:

IREP in the first field of the first card indicates the number of replicate sets.

NARAY in the second field of the first card and the remaining fields of this and succeeding cards consists of an array containing the number of observations in each of the replicate sets.

Note that it is not safe for the program to assume that all the data points for an experiment with the same levels of the independent variables are true replicates. Thus the user must explicitly specify the truly replicated sets. NEWRAP does check that all
independent terms within a replicate set are the same. If not, the program stops. A nonreplicated data point is considered to be a group of size 1. Note that the data in table I are grouped to clearly indicate the replicated data points.

(7) FORMAT (INPUT, FMT)(I2, 13A6): INPUT specifies the unit number on which the input data is stored; and FMT supplies the format for reading it.

Note that, if a weighting factor other than 1.0 is to be used, its value will be read with each data point, and the format must allow for this.

The example from Draper and Smith (ref. 5) uses a weighting value of 1.0 for all data. The format is (12X, 5F6.0) since there are four independent and one dependent variable to be read. If a weighting value other than 1.0 is used, it must appear with every data point as the last value on the card. In such a case, the format could, for example, be (12X, 5F6.0, F10.3).

(8) DATA: Each observation consisting of the given z's and y's read by the execution of one READ statement. Thus, there will be at least one card for each observation. As mentioned previously, if the transformation option is not used, the program expects the first variables to be the independent variables, in the order in which they enter the model, followed by the dependent variables and then the weighting value if IFWT = . F. Otherwise, if transformations are used, the independent and dependent variables may be entered in any convenient order, because the TERMS card(s) will be needed to specify the order in which the values will enter the model equation. However, if IFWT = . F., the weighting value is still the last value supplied with each observation.

(9) PREDICTIONS (PREDCT)(L1):

If, with the program LOGIC (input set 3) card, a computation of residuals is requested by a T in card column 4, then predicted values of the dependent variables are computed for all the input values of the independent variables. In addition to these predicted values, predictions at other values of the independent variables might be desired. In PREDICTIONS input, one card with one column is used to indicate if these other predictions are desired (T or F). If this is F, a new case is started and the new case should start with input set 1 cards. If it is T, the following cards are read: One card with one four-column field specifying the number of predictions desired. This is followed by cards with the values of the independent variables at which predictions are desired. Only the final regression model is used, but the number of independent and dependent variables originally supplied on the PROBLEM SIZE data cards are read. All transformations indicated on the MODEL cards are performed. Then the proper terms are chosen by the program to correspond to the final model. Since the dependent variables are not needed in this part of the program, the numerical values for the dependent variables are dummy values and should be in the appropriate range so that when subroutines required for the transformations (e.g., ALOG, SQRT) use these values, abnormal exits will not occur.
Illustrative Problem Requiring Transformations

As an example of the MODEL cards usage consider the following. Suppose the model we are required to construct is

$$\log_{10}(y + 273.15) = b_0 + b_1 z_1 + b_2 z_2 + b_3 z_3 + b_4 z_1 z_2 + b_5 z_1 z_3 + b_6 z_2 z_3 + b_7 z_1^2 + b_8 z_2^2 + b_9 z_3^2$$

Thus, in terms of equation (2) we have

$$x_1 = z_1$$
$$x_2 = z_2$$
$$x_3 = z_3$$
$$x_4 = z_1 z_2$$
$$x_5 = z_1 z_3$$
$$x_6 = z_2 z_3$$
$$x_7 = z_1^2$$
$$x_8 = z_2^2$$
$$x_9 = z_3^2$$

$$y = \log_{10}(y + 273.15)$$

Table IV shows a sequence of transformations which could be used to construct this model equation. Figure 4 shows how the MODEL cards describing this equation would appear on a FORTRAN data sheet. Figure 5 shows the X and CON array contents both before and after the transformations are performed upon one observation and the X array after the appropriate terms have been selected according to the TERMS card data.
### TABLE IV. - SEQUENY OF TRANSFORMATIONS FOR EXAMPLE

<table>
<thead>
<tr>
<th>Transformation number</th>
<th>Operand</th>
<th>Operation</th>
<th>Operator</th>
<th>Result</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01</td>
<td>00</td>
<td>00</td>
<td>11</td>
<td>$x_1 - \text{CON}(11)$</td>
</tr>
<tr>
<td>2</td>
<td>01</td>
<td>00</td>
<td>00</td>
<td>61</td>
<td>$x_1 - X(61), \text{CON}(61)$</td>
</tr>
<tr>
<td>3</td>
<td>02</td>
<td>00</td>
<td>00</td>
<td>12</td>
<td>$x_2 - \text{CON}(12)$</td>
</tr>
<tr>
<td>4</td>
<td>02</td>
<td>00</td>
<td>00</td>
<td>62</td>
<td>$x_2 - X(62), \text{CON}(62)$</td>
</tr>
<tr>
<td>5</td>
<td>03</td>
<td>00</td>
<td>00</td>
<td>13</td>
<td>$x_3 - \text{CON}(13)$</td>
</tr>
<tr>
<td>6</td>
<td>03</td>
<td>00</td>
<td>00</td>
<td>63</td>
<td>$x_3 - X(63), \text{CON}(63)$</td>
</tr>
<tr>
<td>7</td>
<td>61</td>
<td>02</td>
<td>61</td>
<td>17</td>
<td>$x_1^2 - \text{CON}(17)$</td>
</tr>
<tr>
<td>8</td>
<td>62</td>
<td>02</td>
<td>62</td>
<td>18</td>
<td>$x_2^2 - \text{CON}(18)$</td>
</tr>
<tr>
<td>9</td>
<td>63</td>
<td>02</td>
<td>63</td>
<td>19</td>
<td>$x_3^2 - \text{CON}(19)$</td>
</tr>
<tr>
<td>10</td>
<td>61</td>
<td>02</td>
<td>62</td>
<td>14</td>
<td>$x_1x_2 - \text{CON}(14)$</td>
</tr>
<tr>
<td>11</td>
<td>61</td>
<td>02</td>
<td>63</td>
<td>15</td>
<td>$x_1x_3 - \text{CON}(15)$</td>
</tr>
<tr>
<td>12</td>
<td>62</td>
<td>02</td>
<td>63</td>
<td>16</td>
<td>$x_2x_3 - \text{CON}(16)$</td>
</tr>
<tr>
<td>13</td>
<td>04</td>
<td>01</td>
<td>01</td>
<td>98</td>
<td>$y + 273.15 - X(98), \text{CON}(98)$</td>
</tr>
<tr>
<td>14</td>
<td>98</td>
<td>07</td>
<td>00</td>
<td>20</td>
<td>$\log_{10}(y + 273.15) - \text{CON}(20)$</td>
</tr>
</tbody>
</table>

Figure 4. - An example of MODEL cards.
Figure 5. - Arrays X and CON before and after transformations and terms selection for the example.
**SAMPLE WRAP PROBLEM**

**DATA IS FROM WRAP AND SMITH APPLIED REGRESSION ANALYSIS (REFERENCE 4 OF WRAP REPORT)**

**CHAPTER 7**

**INITIAL MODEL EQUATION IS**

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon \]

\( Y \): CHAMBER PRESSURE

\( X_1 \): TEMPERATURE OF CYCLE

\( X_2 \): VIBRATION LEVEL

\( X_3 \): DROP DURATION

\( X_4 \): STATIC FIRE

**RESIDUALS ARE BEING REQUESTED TO BE PLOTTED (FOR CASEPLT PROGRAM)**

**FOR RESIDUAL PLOTTING ANALYSES**

**THERE IS A NEED TO ESTIMATE**

**THERE ARE 18 REPLICATE SETS**

<table>
<thead>
<tr>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>3</th>
<th>3</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
</table>

**SAMPLE WRAP PROBLEM**

**EQUATION, OBSERVATION = 1**

<table>
<thead>
<tr>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>3</th>
<th>3</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
</table>

**REPLICATE SET 1**

**TERMS OF THE EQUATION, OBSERVATION = 2**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>1.400000</th>
</tr>
</thead>
</table>

**REPLICATE SET 2**

**TERMS OF THE EQUATION, OBSERVATION = 3**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>0</th>
<th>150.0000</th>
<th>26.30000</th>
</tr>
</thead>
</table>

**REPLICATE SET 3**

**TERMS OF THE EQUATION, OBSERVATION = 4**

<table>
<thead>
<tr>
<th>0</th>
<th>165.0000</th>
<th>-65.0000</th>
<th>29.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 4**

**TERMS OF THE EQUATION, OBSERVATION = 5**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>150.0000</th>
<th>32.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 5**

**TERMS OF THE EQUATION, OBSERVATION = 6**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>-75.0000</th>
<th>-65.0000</th>
<th>150.0000</th>
<th>26.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 6**

**TERMS OF THE EQUATION, OBSERVATION = 7**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>175.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>8.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 7**

**TERMS OF THE EQUATION, OBSERVATION = 8**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>-75.0000</th>
<th>-65.0000</th>
<th>150.0000</th>
<th>28.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 8**

**TERMS OF THE EQUATION, OBSERVATION = 9**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>175.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>11.50000</th>
</tr>
</thead>
</table>

**REPLICATE SET 9**

**TERMS OF THE EQUATION, OBSERVATION = 10**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>-65.0000</th>
<th>-65.0000</th>
<th>1.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 10**

**TERMS OF THE EQUATION, OBSERVATION = 11**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>165.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>22.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 11**

**TERMS OF THE EQUATION, OBSERVATION = 12**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>-75.0000</th>
<th>-65.0000</th>
<th>150.0000</th>
<th>3.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 12**

**TERMS OF THE EQUATION, OBSERVATION = 13**

<table>
<thead>
<tr>
<th>0</th>
<th>175.0000</th>
<th>165.0000</th>
<th>150.0000</th>
<th>22.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 13**

**TERMS OF THE EQUATION, OBSERVATION = 14**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>-65.0000</th>
<th>3.70000</th>
</tr>
</thead>
</table>

**THERE IS A NEED TO ESTIMATE**

**THERE ARE 18 REPLICATE SETS**

**SAMPLE WRAP PROBLEM**

**EQUATION, OBSERVATION = 1**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>-65.0000</th>
<th>1.400000</th>
</tr>
</thead>
</table>

**REPLICATE SET 1**

**TERMS OF THE EQUATION, OBSERVATION = 2**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>0</th>
<th>150.0000</th>
<th>26.30000</th>
</tr>
</thead>
</table>

**REPLICATE SET 2**

**TERMS OF THE EQUATION, OBSERVATION = 3**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>-65.0000</th>
<th>150.0000</th>
<th>29.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 3**

**TERMS OF THE EQUATION, OBSERVATION = 4**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>165.0000</th>
<th>-65.0000</th>
<th>9.700000</th>
</tr>
</thead>
</table>

**REPLICATE SET 4**

**TERMS OF THE EQUATION, OBSERVATION = 5**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>150.0000</th>
<th>32.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 5**

**TERMS OF THE EQUATION, OBSERVATION = 6**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>-75.0000</th>
<th>150.0000</th>
<th>26.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 6**

**TERMS OF THE EQUATION, OBSERVATION = 7**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>175.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>8.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 7**

**TERMS OF THE EQUATION, OBSERVATION = 8**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>-75.0000</th>
<th>-65.0000</th>
<th>150.0000</th>
<th>28.40000</th>
</tr>
</thead>
</table>

**REPLICATE SET 8**

**TERMS OF THE EQUATION, OBSERVATION = 9**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>175.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>11.50000</th>
</tr>
</thead>
</table>

**REPLICATE SET 9**

**TERMS OF THE EQUATION, OBSERVATION = 10**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>-65.0000</th>
<th>-65.0000</th>
<th>1.30000</th>
</tr>
</thead>
</table>

**REPLICATE SET 10**

**TERMS OF THE EQUATION, OBSERVATION = 11**

<table>
<thead>
<tr>
<th>175.0000</th>
<th>165.0000</th>
<th>0</th>
<th>-65.0000</th>
<th>22.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 11**

**TERMS OF THE EQUATION, OBSERVATION = 12**

<table>
<thead>
<tr>
<th>-75.0000</th>
<th>-65.0000</th>
<th>-65.0000</th>
<th>0.400000</th>
</tr>
</thead>
</table>

**REPLICATE SET 12**

**TERMS OF THE EQUATION, OBSERVATION = 13**

<table>
<thead>
<tr>
<th>0</th>
<th>175.0000</th>
<th>165.0000</th>
<th>150.0000</th>
<th>22.90000</th>
</tr>
</thead>
</table>

**REPLICATE SET 13**

**TERMS OF THE EQUATION, OBSERVATION = 14**

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>-65.0000</th>
<th>3.70000</th>
</tr>
</thead>
</table>
**REPLICATE SET 15** 
**TERMS OF THE EQUATION, OBSERVATION = 15**

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>SSQ</th>
<th>SUM</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.006686</td>
<td>76.400000</td>
<td>25.400000</td>
</tr>
</tbody>
</table>

**TERMS OF THE EQUATION, OBSERVATION = 16**

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>SSQ</th>
<th>SUM</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.880044</td>
<td>55.200000</td>
<td>27.600000</td>
</tr>
</tbody>
</table>

**TERMS OF THE EQUATION, OBSERVATION = 17**

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>SSQ</th>
<th>SUM</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8000137E-01</td>
<td>23.200000</td>
<td>11.600000</td>
</tr>
</tbody>
</table>

**MEANS OF INJP AND DEP VARIABLES**

<table>
<thead>
<tr>
<th>X</th>
<th>TRANSPOSE X MATRIX WHERE X IS DEVIATION FROM MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROW 1</td>
<td>105000.0</td>
</tr>
<tr>
<td>ROW 2</td>
<td>62500.00</td>
</tr>
<tr>
<td>ROW 3</td>
<td>26250.00</td>
</tr>
<tr>
<td>ROW 4</td>
<td>-26875.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Y TRANSPOSE Y MATRIX WHERE X AND Y ARE DEVIATIONS FROM MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROW 1</td>
</tr>
<tr>
<td>ROW 2</td>
</tr>
<tr>
<td>ROW 3</td>
</tr>
<tr>
<td>ROW 4</td>
</tr>
</tbody>
</table>

**CORRELATION COEFFICIENTS**

<table>
<thead>
<tr>
<th>X</th>
<th>TRANSPOSE X MATRIX WHERE X IS DEVIATION FROM MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROW 1</td>
<td>1.000000</td>
</tr>
<tr>
<td>ROW 2</td>
<td>0.537706</td>
</tr>
<tr>
<td>ROW 3</td>
<td>0.231943</td>
</tr>
<tr>
<td>ROW 4</td>
<td>-0.157485</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>THE FOLLOWING ARE EIGENVALUES OF X TRANSPOSE X MATRIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>489039.32</td>
</tr>
</tbody>
</table>

**THIS IS THE MODAL MATRIX OR MATRIX OF EIGENVECTORS. EIGENVECTORS ARE WRITTEN IN COLUMNS LEFT TO RIGHT IN SAME ORDER AS EIGENVALUES**

| ROW 1   | 0.117713 | 0.231943 | 0.537706 | 0.231943 |
| ROW 2   | 0.674835 | 0.224353 | 0.212138 | -0.671970 |
| ROW 3   | 0.358164 | 0.666753 | 0.500097 | 0.423079 |
| ROW 4   | -0.157485 | 0.682691 | 0.177159 | -0.343240 |
SAMPLE: \texttt{WRAP PROBLEM}

Each column contains one dependent term

**Constant Term** (D1)

\begin{align*}
1 & 1.01423 \\
2 & 0.751127e-02 \\
3 & 0.111256e-01 \\
4 & 0.424585e-02 \\
5 & 0.101374
\end{align*}

**ANOVA of Regression on Dependent Variable**

\begin{tabular}{|c|c|c|c|}
\hline
Source & Sums of Squares & Degrees of Freedom & Mean Squares \\
\hline
Regression & 2492.56296 & 4 & 623.141.79 \\
Residual & 24.03694 & 23 & 1.0107.998 \\
Total & 2711.59998 & 23 & \\
\hline
\end{tabular}

- $R^2 = 0.908159$, \( R = 0.952974 \)

- Standard Error of Estimate: 3.620185

**ANOVA of Lack of Fit**

\begin{tabular}{|c|c|c|c|}
\hline
Source & Sums of Squares & Degrees of Freedom & Mean Squares \\
\hline
Lack of Fit & 249.03668 & 19 & 1.3107.698 \\
Residual & 244.03694 & 23 & 1.0107.998 \\
Total & 2711.59998 & 23 & \\
\hline
\end{tabular}

**SUMS OF SQUARES DUE TO EACH VARIABLE IF IT WERE LAST TO ENTER REGRESSION**

\begin{align*}
1 & 5.207.672 \\
2 & 1.442.704 \\
3 & 2.021.596 \\
4 & 1783.952 \\
\end{align*}

**Standard Deviation of Regression Coefficients (Derived from Diagonal Elements of \( \text{EX TRANSPOSE X INVERSE MATRIX} \))**

\begin{align*}
0 & 0.354068 \\
1 & 0.454576e-02 \\
2 & 0.204526e-02 \\
3 & 0.455526e-02 \\
4 & 0.320988e-02 \\
\end{align*}

**EX TRANSPOSE X INVERSE MATRIX**

\begin{align*}
\text{ROW 1} & 0.011236e-04 \\
\text{ROW 2} & 0.320436e-05 \\
\text{ROW 3} & 0.220456e-06 \\
\text{ROW 4} & 0.615436e-06 \\
\end{align*}

**Sample 'Wrap Problem'**

Calculate T Statistics

The T Statistics can be used to test the set regression coefficients \( b(i) \).

\begin{align*}
1 & 1.67982 \\
2 & 2.78262 \\
3 & 3.10788 \\
4 & 3.10788
\end{align*}

**Under Null Hypothesis The Interval \(-t, t\), Where \( t \) is Given Above, has Approx Probability Listed Below**

\begin{align*}
1 & 0.256 \\
2 & 0.008 \\
3 & 0.008 \\
4 & 0.008
\end{align*}

The Desired Value of Probability is 95.0 percent

The term at 31 is being deleted.
SAMPLE INVAP PUBLI M
EACH COLUMN CONTAINS THE COEFFICIENTS FOR ONE DEPENDENT TERM
CONSTANT TERM (.00)

\[ a_1 = 1.17449 \]
\[ a_2 = 0.751732E-02 \]
\[ a_3 = 2.17E-01 \]
\[ a_4 = 0.416394 \]

ANOVA OF REGRESSION ON DEPENDENT VARIABLE Y

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SUMS OF SQUARES</th>
<th>DEGREES OF FREEDOM</th>
<th>MEAN SQUARES</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>2463.34143</td>
<td>3</td>
<td>820.11388</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>251.258162</td>
<td>20</td>
<td>12.5624079</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2714.59938</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

\[ R^2 = \frac{SS(\text{REGRESSION})}{SS(\text{TOTAL})} = 0.907339 \]
\[ R = 0.952544 \]

ANOVA OF LACK OF FIT

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SUMS OF SQUARES</th>
<th>DEGREES OF FREEDOM</th>
<th>MEAN SQUARES</th>
</tr>
</thead>
<tbody>
<tr>
<td>LACK OF FIT</td>
<td>240.184813</td>
<td>14</td>
<td>17.1560571</td>
</tr>
<tr>
<td>REPLICAION</td>
<td>1.1073571</td>
<td>6</td>
<td>1.8455995</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>251.258162</td>
<td>20</td>
<td>12.5624079</td>
</tr>
</tbody>
</table>

\[ F = \frac{MS(\text{LACK OF FIT})}{MS(\text{REPLICATION})} = 9.296 \]

SUMS OF SQUARES DUE TO EACH VARIABLE IF IT WERE LAST TO ENTER REGRESSION

| 1 \[ 0.053833 \]
| 2 \[ 27.08253 \]
| 4 \[ 1891.095 \]

STANDARD Deviation OF REGRESSION COEFFICIENTS (Derived from Diagonal Elements of \((X \text{ transpose} \times X)\text{inverse} \text{matrix})

| 1 \[ 0.353984 \]
| 2 \[ 0.455239E-02 \]
| 3 \[ 0.370935-07 \]
| 4 \[ 0.121752E-02 \]

\((X \text{ transpose} \times X)\text{inverse} \text{matrix})

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a_1 0.111809E-04</td>
</tr>
<tr>
<td>2</td>
<td>a_2 -0.364602E-05</td>
</tr>
<tr>
<td>3</td>
<td>a_3 -0.713401E-05</td>
</tr>
<tr>
<td>4</td>
<td>a_4 0.362557E-05</td>
</tr>
</tbody>
</table>

SAMPLE N\times P HUEBJ

CAlculated T statistics

THE T statistics can be used to test the YET REGRESSION COEFFICIENTS U.T.I.

| 1.66794 |
| 3.830724 |
| 31.56767 |

UNDER null hypothesis the interval (1-T,t) WHEN T IS GIVEN ABOVE, HAS APPROXIMATE probability listed below.

MINUS 5%4 INDICATES PROB EXCEEDS .9494

| 1 \[ 1.854 \]
| 2 \[ 3.991 \]
| 4 \[ 9.994 \]

THE NEEDED VALUE OF PROBABILITY IS .05, 9494
THE TERM AT .01 IS BEING USED
SAMPLE NEW PAP PROBLEM
EACH COLUMN CONTAINS THE COEFFICIENTS FOR ONE DEPENDENT TERM
CONSTANT TERM (B₀)
11.7 6.1
REGRESSION COEFFICIENTS (B₁,...,Bₖ)
2 0.159851 0.1
4 0.102354

ANOVA OF REGRESSION ON DEPENDENT VARIABLE

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SUMS OF SQUARES</th>
<th>DEGREES OF FREEDOM</th>
<th>MEAN SQUARES</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>2492.28062</td>
<td>2</td>
<td>1247.64200</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>256.320102</td>
<td>21</td>
<td>12.2053314</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2748.60074</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

R SQUARED = SSREG / SSTR = 0.809565
R = 0.905435

R SQUARED = SSREG / SSTR = 0.809565
R = 0.905435

MSSREG = 1247.64200
MSE = 12.2053314

F = MSSREG / MSE = 101.2672

ANOVA OF LACK OF FIT

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SUMS OF SQUARES</th>
<th>DEGREES OF FREEDOM</th>
<th>MEAN SQUARES</th>
</tr>
</thead>
<tbody>
<tr>
<td>LACK OF FIT</td>
<td>245.238604</td>
<td>15</td>
<td>16.3492401</td>
</tr>
<tr>
<td>REPLIATION</td>
<td>11.9735971</td>
<td>6</td>
<td>1.84555951</td>
</tr>
<tr>
<td>RESIDUAL</td>
<td>256.311962</td>
<td>21</td>
<td>12.2053314</td>
</tr>
<tr>
<td>F = MSSOF / MSE = 8.859</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SUMS OF SQUARES DUE TO EACH VARIABLE IF IT WERE LAST TO ENTER REGRESSION
2 41.2672
4 1071.546

STANDARD DEVIATION OF REGRESSION COEFFICIENTS (DERIVED FROM DIAGONAL ELEMENTS OF (X TRANSPOSE X) INVERSE MATRIX)
0 0.336495
2 0.324983E-02
4 0.361448E-02

(X TRANSPOSE X) INVERSE MATRIX

<table>
<thead>
<tr>
<th>ROWNUM</th>
<th>0.585090E-05</th>
<th>0.595829E-05</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROWNUM</td>
<td>0.328383E-05</td>
<td>0.559879E-05</td>
</tr>
</tbody>
</table>

SAMPLE NEW PAP PROBLEM
CALCULATED T STATISTICS
THE T STATISTICS CAN BE USED TO TEST THE REGRESSION COEFFICIENTS (B₁)
1.421864
3.844403

UNDER NULL HYPOTHESIS THE INTERVAL [-T,T] WHERE T IS GIVEN ABOVE, HAS APPROX PROBABILITY LISTED BELOW.
MINUS SIGN INDICATES PROP EXCEEDS .999.
7 0.997
4 <.999

THE DESIRED VALUE OF PROBABILITY IS 95.0 PERCENT
SAMPLE NEW PAP PROBLEM
FOR EACH DEPENDENT TERM AND OBSERVATION IS PRINTED
OBSERVED (RESP: Y OBSERVED)
CALCULATED RESPONSE (Y CALC)
RESIDUAL (Y395 - YCALC=YDIF)
STUDENTIZED RESIDUAL (Z)

| Y OBSERVED | 1.4000 |
| Y CALC     | 1.6650 |
| Y DIF      | -0.2650 |
| STUDENTIZED| -2.9909 |

| Y OBSERVED | 26.300 |
| Y CALC     | 27.063 |
| Y DIF      | -0.763 |
| STUDENTIZED| -0.3619 |

| Y OBSERVED | 29.400 |
| Y CALC     | 27.063 |
| Y DIF      | 2.3367 |
| STUDENTIZED| 1.7200 |
SAMPLE NEWRAP PROBLEM

<table>
<thead>
<tr>
<th>Y OBSERVED</th>
<th>Y CALC</th>
<th>Y DIF</th>
<th>STUDENTIZED</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.400</td>
<td>27.083</td>
<td>-5.683</td>
<td>-4.169</td>
</tr>
<tr>
<td>22.900</td>
<td>29.793</td>
<td>-6.893</td>
<td>-2.56%</td>
</tr>
<tr>
<td>20.200</td>
<td>25.093</td>
<td>-4.893</td>
<td>-1.83%</td>
</tr>
<tr>
<td>26.500</td>
<td>25.093</td>
<td>1.407</td>
<td>-0.55%</td>
</tr>
<tr>
<td>23.400</td>
<td>25.093</td>
<td>-1.693</td>
<td>-0.66%</td>
</tr>
<tr>
<td>26.500</td>
<td>25.093</td>
<td>1.407</td>
<td>-0.55%</td>
</tr>
<tr>
<td>5.8000</td>
<td>7.7850</td>
<td>-1.985</td>
<td>-1.46%</td>
</tr>
<tr>
<td>7.4000</td>
<td>7.7850</td>
<td>-0.385</td>
<td>-0.27%</td>
</tr>
<tr>
<td>5.8000</td>
<td>7.7850</td>
<td>-1.985</td>
<td>-1.46%</td>
</tr>
</tbody>
</table>

SAMPLE NWRAP PROBLEM

<table>
<thead>
<tr>
<th>Y OBSERVED</th>
<th>Y CALC</th>
<th>Y DIF</th>
<th>STUDENTIZED</th>
</tr>
</thead>
<tbody>
<tr>
<td>28.600</td>
<td>25.093</td>
<td>3.507</td>
<td>2.134%</td>
</tr>
</tbody>
</table>
NEWRAP DOCUMENTATION AND LISTINGS

The contents of this section include a flow chart of the program, a listing of the routines used in NEWRAP and their major functions, the call structure of the program, a dictionary of the program, and the listing.

General Mathematical and Logical Flow of Program

The flow of operation in NEWRAP is illustrated in figure 6.
I Perform
I transformations
Replications?
Yes
Calculate replication error
Replications?
No
Calculate $X'X$, $X'y$, $C$$1$
Add one term at a time?
Yes
j = 0
j = j + 1$
3
Calculate $(X'X)^{-1}_j$
Compute coefficients
Compute ANOVA data $R^2$, $R$, $\bar{y}$
Compute t's
No
START
Yes
START
No
Want to delete terms?
Yes
Choose term, among those not being forced to remain in model, with minimum significance
No
Below critical level?
Yes
Delete proper row and column from $X'X$$1$
No
Compute residuals and chi-squared statistic?
Yes
Compute residuals, compute number of cells, compute chi-squared statistic
No
Want to predict variables?
Yes
Given $X_0$
Compute $X_0'(X'X)^{-1}_j X_0$
Compute $s^2 = y - x_0 S^2$
More points to predict?
No
START

Figure 6. - Flow chart for NEWRAP.
## Routines and Their Major Functions

<table>
<thead>
<tr>
<th>FORTRAN name</th>
<th>Function of routine</th>
</tr>
</thead>
<tbody>
<tr>
<td>BORD</td>
<td>Inverts symmetric matrix of order $n$ by adding bordering column to already inverted matrix of order $n - 1$</td>
</tr>
<tr>
<td>EIGEN</td>
<td>Computes eigenvalues and eigenvectors of input symmetric matrix</td>
</tr>
<tr>
<td>HIST</td>
<td>Prints histogram of residuals</td>
</tr>
<tr>
<td>INVXTX</td>
<td>Inverts symmetric matrix</td>
</tr>
<tr>
<td>LOC</td>
<td>When given row and column indices of symmetric matrix element, it computes location this element would have if only lower triangular part were stored as vector.</td>
</tr>
<tr>
<td>MATINV</td>
<td>Controls inversion process; computes regression coefficients; computes eigenvalues and eigenvectors of $X'X$ if requested</td>
</tr>
<tr>
<td>MFIX</td>
<td>Prints $X'X$ and computes and prints $C$</td>
</tr>
<tr>
<td>NEWRAP</td>
<td>Executes overall problem control; computes replication error; controls deletion of variables when given results of t-test</td>
</tr>
<tr>
<td>OUTPLT</td>
<td>Computes residuals at observed points and plots them. Compute chi-squared statistic</td>
</tr>
<tr>
<td>PREDCT</td>
<td>Computes predicted values, variances, and standard deviations of regression line and further observations at specified points</td>
</tr>
<tr>
<td>RECT</td>
<td>Writes rectangular matrix</td>
</tr>
<tr>
<td>RSTATS</td>
<td>Computes regression statistics and writes regression and lack-of-fit analysis of variance tables</td>
</tr>
<tr>
<td>SUMUPS</td>
<td>Constructs $X'X$ and $X'y$ matrices one observation at a time, in double precision</td>
</tr>
<tr>
<td>TRAN</td>
<td>Performs transformations</td>
</tr>
<tr>
<td>TRIANG</td>
<td>Writes lower triangular part of symmetric matrix</td>
</tr>
<tr>
<td>TTEST</td>
<td>Computes t-statistics and their significance levels; determines which variable should be deleted</td>
</tr>
</tbody>
</table>

### Call Structure of Program

The call structure of the program is illustrated in figure 7.
### Dictionary of Program

<table>
<thead>
<tr>
<th>FORTRAN name</th>
<th>Mathematical symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTDEV</td>
<td>$e_i$</td>
<td>Error in observation $i$; difference between observed and predicted response</td>
</tr>
<tr>
<td>B</td>
<td>$b$</td>
<td>Regression coefficients other than the constant</td>
</tr>
<tr>
<td>BO</td>
<td>$b_0$</td>
<td>Constant regression coefficient</td>
</tr>
<tr>
<td>BZERO</td>
<td></td>
<td>Logical variable set to T if constant $b_0$ coefficient should be in regression model</td>
</tr>
<tr>
<td>CHISQ</td>
<td>$\chi^2$</td>
<td>Chi-squared statistic</td>
</tr>
<tr>
<td>CON</td>
<td></td>
<td>Constants used in transformations, and results of transformations</td>
</tr>
<tr>
<td>DELETE</td>
<td></td>
<td>Logical variable set to T when deletion of terms is desired</td>
</tr>
<tr>
<td>DUMMY</td>
<td></td>
<td>Extra array used for plotting data</td>
</tr>
<tr>
<td>ECONMY</td>
<td></td>
<td>Logical variable indicating suppress printout of $X'X$, $X'X$ deviations, and $C$ if $T$</td>
</tr>
<tr>
<td>ERRMS</td>
<td>$\sigma^2$</td>
<td>Estimate of $\sigma^2$ used in hypothesis tests</td>
</tr>
<tr>
<td>FMT</td>
<td></td>
<td>Variable input format</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>FMTTRI</td>
<td>Format for printing matrix</td>
<td></td>
</tr>
<tr>
<td>IDENT</td>
<td>First identification printed at top of each page</td>
<td></td>
</tr>
<tr>
<td>IDOUT</td>
<td>Original sequence number of each term relating reduced models to original model</td>
<td></td>
</tr>
<tr>
<td>IFCHI</td>
<td>Logical variable set to T if residual computations and plots are desired</td>
<td></td>
</tr>
<tr>
<td>IFSSR</td>
<td>Logical variable set to T if sequential regressions are desired</td>
<td></td>
</tr>
<tr>
<td>IFTT</td>
<td>Logical variable set to T if t-statistics are desired</td>
<td></td>
</tr>
<tr>
<td>IFWT</td>
<td>Logical variable set to T if all weights of observations are 1.0</td>
<td></td>
</tr>
<tr>
<td>INPUT</td>
<td>Input logical tape unit number for data</td>
<td></td>
</tr>
<tr>
<td>INPUT5</td>
<td>Set equal to 5 to denote input device is card reader</td>
<td></td>
</tr>
<tr>
<td>INTER</td>
<td>Tape unit where input data is stored for use in OUTPLT</td>
<td></td>
</tr>
<tr>
<td>IOUT</td>
<td>Sequence number of term among those remaining which is to be deleted</td>
<td></td>
</tr>
<tr>
<td>JCOL</td>
<td>Total number of independent and dependent terms in regression model</td>
<td></td>
</tr>
<tr>
<td>KONNO</td>
<td>Number of constants originally supplied for transformations</td>
<td></td>
</tr>
<tr>
<td>LENGTH</td>
<td>Number of locations in matrix storage area currently needed</td>
<td></td>
</tr>
<tr>
<td>LIST</td>
<td>Set equal to 6 to denote output device is printer</td>
<td></td>
</tr>
<tr>
<td>NARAY</td>
<td>Number of replications per replicate set</td>
<td></td>
</tr>
<tr>
<td>NCON</td>
<td>Array containing addresses in CON array for use in transformations</td>
<td></td>
</tr>
<tr>
<td>NERROR</td>
<td>Degrees of freedom for error mean square estimate</td>
<td></td>
</tr>
<tr>
<td>NLOF</td>
<td>Degrees of freedom for estimating variance due to lack of fit</td>
<td></td>
</tr>
<tr>
<td>NODEP</td>
<td>Number of dependent variables</td>
<td></td>
</tr>
<tr>
<td>NOOB</td>
<td>Number of observations</td>
<td></td>
</tr>
</tbody>
</table>

N = J - NPDEG - D
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTERM</td>
<td>Number of terms in current regression model</td>
</tr>
<tr>
<td>NOVAR</td>
<td>Number of independent variables to be read</td>
</tr>
<tr>
<td>NPDEG</td>
<td>Pooled degrees of freedom for replication error</td>
</tr>
<tr>
<td>NREG</td>
<td>Degrees of freedom for determining variance due to regression</td>
</tr>
<tr>
<td>NRES</td>
<td>Degrees of freedom for estimation of residual variance</td>
</tr>
<tr>
<td>NTERM</td>
<td>Array containing locations of terms in CON array that should be in regression model</td>
</tr>
<tr>
<td>NTOT</td>
<td>Total degrees of freedom</td>
</tr>
<tr>
<td>NTRAN</td>
<td>Array containing transformation codes for use in performing transformations</td>
</tr>
<tr>
<td>NTRANS</td>
<td>Number of transformations to perform</td>
</tr>
<tr>
<td>NWHERE</td>
<td>Location in X array of first dependent variable; used in prediction routine to adjust for deleted terms</td>
</tr>
<tr>
<td>NXCOD</td>
<td>Array containing addresses of variables (or terms with address &gt;60) for use in transformations</td>
</tr>
<tr>
<td>P</td>
<td>Probability that interval (-t, t) must have before a term is considered to be significant</td>
</tr>
<tr>
<td>PNCH</td>
<td>Logical variable set to T if residuals are to be punched</td>
</tr>
<tr>
<td>POOLED</td>
<td>Array containing pooled sums of squares from replications for each dependent term</td>
</tr>
<tr>
<td>PREDCT</td>
<td>Logical variable set to T if prediction option is desired</td>
</tr>
<tr>
<td>REPS</td>
<td>Logical variable set to T if there are replicate sets in data</td>
</tr>
<tr>
<td>REPVAR</td>
<td>Array containing replication variance of each dependent term</td>
</tr>
<tr>
<td>RESMS</td>
<td>Array containing residual mean square or variance of each dependent term</td>
</tr>
<tr>
<td>RNLOF</td>
<td>Reciprocal of degrees of freedom for lack of fit</td>
</tr>
<tr>
<td>RNREG</td>
<td>Reciprocal of degrees of freedom for regression</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>RNRES</td>
<td>Reciprocal of degrees of freedom for residual</td>
</tr>
<tr>
<td>RWT</td>
<td>Reciprocal of total weight</td>
</tr>
<tr>
<td>SATRTD</td>
<td>Logical variable indicating that there are no degrees of freedom for residual if T</td>
</tr>
<tr>
<td>STORYX</td>
<td>Logical variable set to T if eigenvectors and eigenvalues of X'X are to be computed and printed</td>
</tr>
<tr>
<td>SUMX</td>
<td>( \sum x, \sum y ) Array containing sums of independent and dependent terms</td>
</tr>
<tr>
<td>SUMX2</td>
<td>( \sum x^2, \sum y^2 ) Array containing sum of squared independent and dependent terms</td>
</tr>
<tr>
<td>SUMXX</td>
<td>X'X Sums of squares and crossproducts matrix, and variance-covariance matrix of independent terms</td>
</tr>
<tr>
<td>SUMXXI</td>
<td>((X'X)^{-1}) Inverse of variance-covariance matrix of independent variables</td>
</tr>
<tr>
<td>SUMXY</td>
<td>X'y Array containing sums of crossproducts of independent terms with dependent terms</td>
</tr>
<tr>
<td>TOTWT</td>
<td>( w_i ) Sum of weight of observations</td>
</tr>
<tr>
<td>X</td>
<td>Before transformations are performed, this contains the variables as read in. After transformations are performed, appropriate data from CON array are placed here according to information on TERMS cards.</td>
</tr>
<tr>
<td>XCHK</td>
<td>Array used in checking if all values of independent terms are the same within a replicate set</td>
</tr>
<tr>
<td>ZEAN</td>
<td>E(X), ( \mu ) Expected or mean value (or X)</td>
</tr>
</tbody>
</table>

**Program Listing**

```
$IBFTC BLDV

BLOCK DATA 1
COMM001 FN0MTS/ FMT(13),FMTTRI(14) 2
COMM002 BIG/B(60,9),SUMXY(60,9),SUMXX(1830),SUMXXI(1830) 3
X,DUMMY(2300) 4
DOUBLE PRECISION B,SUMXY, SUMXX, SUMXXI 5
COMM003 MED/B0(9),SUMX(69),SUMX2(69),SUMY2(9),LEAN(69), 6
X CON(99), ERRMS(9), IDENT(13), IDOUT(60), NCON(200), NTERM(69), 7
```
C THIS IS NEWRAP, MAIN PROGRAM FOR REGRESSION ANALYSIS PROVIDING
C INTERVAL EVALUATION OF RESULTS.
C**********************************************************************
C
COMMON /FRMTS/ FMT(13),FMTTRI(14)
COMMON/BIG/B(160,9),SUMXY(60,9),SUMXX(1830),SUMXXI(1830)
DOUBLE PRECISION S,B,SUMXY,SUMXX,SUMXXI
COMMON/MED/B(80,9),SUMX(69),SUMXY(69),SUMXX(69),SUMXX1(9),ZEAN(69),
X CONF(99),ERRMS(9),IDENT(13),IDOUT(60),NCONF(200),INTERM(69),
X NTRAN(100),NXCOD(100),POOLED(9),REPVAR(9),RESMS(9),X(99)
DOUBLE PRECISION B,SUMXY,SUMXX,SUMXX1
COMMON/SALL/ BYPASS,BZERO,DELETE,FIRST,IFCHI,IFSSR,
X IFIT, IFLT, INPUT, INPUT5, INTER,
X ISTRAT, JCOL, KONNO, LENGTH, LIST,
X NERROR, NODEP, NOOB, NOTERM,
X NOVAR, NPDEG, NRES, NTRANS, NWHERE,
X P, PREDCT, REPS, RWT,
X STORY, STORYC, STORYX, TOTWT, WEIGHT,
X ERRFxd, ECONMY, IOUT, ICOL
LOGICAL ECONMY
DOUBLE PRECISION RWT,TOTWT,WEIGHT
DATA INTER/3/,INPUT5/5/, LIST/6/
DATA (FMTTRI(I),I=1,4)/6H(SH RO, 6HW 15129 6HX*(8Gl, 6H5.6)) /
COMMON/MAXPLT
C MAXPLT SHOULD BE THE NUMBER OF SINGLE LENGTH WORDS IN COMMON/BIG/
C BEGINNING AT THE FIRST LOCATION OF SUMXY
DATA MAXPLT/10700/
END

$IBFTC NEWRAP

C
C**********************************************************************
C
COMMON /FRMTS/ FMT(13),FMTTRI(14)
COMMON/BIG/B(160,9),SUMXY(60,9),SUMXX(1830),SUMXXI(1830)
DOUBLE PRECISION S,B,SUMXY,SUMXX,SUMXXI
COMMON/MED/B(80,9),SUMX(69),SUMXY(69),SUMXX(69),SUMXX1(9),ZEAN(69),
X CONF(99),ERRMS(9),IDENT(13),IDOUT(60),NCONF(200),INTERM(69),
X NTRAN(100),NXCOD(100),POOLED(9),REPVAR(9),RESMS(9),X(99)
DOUBLE PRECISION B,SUMXY,SUMXX,SUMXX1
COMMON/SALL/ BYPASS,BZERO,DELETE,FIRST,IFCHI,IFSSR,
X IFIT, IFLT, INPUT, INPUT5, INTER,
X ISTRAT, JCOL, KONNO, LENGTH, LIST,
X NERROR, NODEP, NOOB, NOTERM,
X NOVAR, NPDEG, NRES, NTRANS, NWHERE,
X P, PREDCT, REPS, RWT,
X STORY, STORYC, STORYX, TOTWT, WEIGHT,
X ERRFxd, ECONMY, IOUT, ICOL
LOGICAL ECONMY
DOUBLE PRECISION RWT,TOTWT,WEIGHT
DATA INTER/3/,INPUT5/5/, LIST/6/
DATA (FMTTRI(I),I=1,4)/6H(SH RO, 6HW 15129 6HX*(8Gl, 6H5.6)) /
COMMON/MAXPLT
C MAXPLT SHOULD BE THE NUMBER OF SINGLE LENGTH WORDS IN COMMON/BIG/
C BEGINNING AT THE FIRST LOCATION OF SUMXY
DATA MAXPLT/10700/
END
DO 102 J=1,225

102 B0(J)=0.000

C*****************************************************************************
C READ IDENTIFICATION CARD AND OPTIONS CARD
C*****************************************************************************
C READ(INPUT5,110) I,IDENT
WRITE(LIST,111) IDENT
FIRST=.TRUE.
ERRFX=.FALSE.
113 IF(I) 120,120
120 READ(INPUT5,300) FMT
WRITE(LIST,301) FMT
I=I-1
GO TO 113
115 READ(INPUT5,300) FMT
WRITE(LIST,301) FMT
I=I-1
GO TO 113
120 READ(INPUT5,1282) NOVAR,NODEP,NOTERM,NOOB,NTKEEP
WRITE(LIST,1283) NOVAR,NODEP,NOTERM,NOOB
IF(NTKEEP.NE.0) WRITE(LIST,1307) NTKEEP
READ(INPUT5,117) BZERO, IFTT, IFWT, IFCHI, STORYX, IFSSR,
X ECONMY, ISTRAT, PNCX
WRITE(LIST,118) BZERO, IFTT, IFWT, IFCHI, STORYX, IFSSR, ECONMY,
X ISTRAT, PNCX
C THESE ARE INITIALIZATIONS MADE BEFORE EACH SET OF DATA
C ICOL DETERMINES THE NUMBER OF VARIABLES READ PER OBSERVATION
C JCOL IS THE NUMBER OF TERMS IN THE TOTAL REGRESSION EQUATION
C LENGTH IS THE NUMBER OF STORES NEEDED FOR THE MATRICES
LENGTH= NOTERM*(NOTERM+1)/2
ICOL=NOVAR+NODEP
JCOL = NOTERM +NODEP
NWHERE = NOTERM
REWIND INTER
DO 140 J=1,60
IDOUT(J) = J
140 CONTINUE
DO 145 J=1,100
NXCOD(J) = J
NTRAN(J) = 0
145 CONTINUE
140 NTERM(J) = J
DO 145 J=1,100
NXCOD(J) = J
NTRAN(J) = 0
145 CONTINUE
C IF(BZERO) WRITE(LIST,190)
IF(NTERM.NE.BZERO) WRITE(LIST,170)
C*****************************************************************************
C*****************************************************************************
READ(INPUT5,282) NTRANS,KONNO
IF(NTRANS.EQ.0) GO TO 255
220 READ (INPUT5,230) (NTERM(K), K=1,JCOL)
WRITE(LIST,235) (NTERM(K), K=1,JCOL)
READ(INPUT5,230) (NXCOD(I), NTRAN(I), NCON(2*I-1), NCON(2*I), I=1,NTRANS)
WRITE(LIST,240) (NXCOD(I), NTRAN(I), NCON(2*I-1), NCON(2*I), I=1,NTRANS)
IF(KONNO) 255,255,250
WRITE(LIST,262) (CON(I), I=1,KONNO)
250 READ (INPUT5,260) (CON(I), I=1,KONNO)
WRITE(LIST,262) (CON(I), I=1,KONNO)
C*****************************************************************************
C*****************************************************************************
255 READ(INPUT5,257) DELETE,P
IF(DELETE) IFTT=.TRUE.
C IF THERE ARE REPLICATED POINTS READ IN THE NUMBER OF POINTS AND
C THE NUMBER OF REPLICATIONS. SINGLE DATA POINTS ARE DATA POINTS
C REPLICATED ONCE. IMPLIED HERE.
READ(INPUT5,257) REPS
XS SAVE=.FALSE.
265 IF(.NOT.REPS) GO TO 290
READ(INPUT5,282) IREP,NARAY(I),I=1,IREP
WRITE(LIST,284) IREP
WRITE(LIST,283)(NARAY(I),I=1,IREP)
NPD C=0
IREP=1
GARAY(I)
XS SAVE=.TRUE.
DO 315 I=1,NODEP
POOLED(I)=0.0
S(I)=1.0
315 SSQ(I)=0.0
C READ VARIABLE FORMAT FOR DATA
290 READ(INPUT5,110) INPUT,FMT
WRITE(LIST,111) FMT
310 TOTWT=0.00
WEIGHT=1.000
WRITE(LIST,301) IDENT
C READ IN INPUT VARIABLES
DO 492 J=1,NOOR
330 IF(.NOT.IFHT) GO TO 350
340 REAL) (INPUT,FMT) (X(I),I=1,ICOL)
GO TO 360
350 READ (INPUT,FMT)(X(I),I=1,ICOL), WEIGHT
360 CONTINUE
IF(.EDC4MY) WRITE(LIST,381) J, (X(I),I=1,ICOL)
381 FORMAT(1H I4,9G14.6/(5X,9G14.6))
IF(INTERFQ.EQ.0) GO TO 450
IF(.EDC4MY) GO TO 390
WRITE(LIST,370) WEIGHT,J
WRITE (LIST,38) (X(I),I=1,ICOL)
390 CALL TRANS
420 DO =35 K=1,JCOL
I=TERM(K)
X(K) = CON(I)
430 CONTINUE
450 CONTINUE
IF(.EDC4MY) GO TO 4609
WRITE(LIST,460) J
461 WRITE (LIST,389)(X(I),I=1,JCOL)
4609 CONTINUE
IF(.CHI) WRITE(INTER) (X(I),I=1,69),WEIGHT
IF(.NOT.XSAVE) GO TO 4611
DO 4610 K=1,NOTERM
4610 XCHK(K)=X(K)
XSAVE=.FALSE.
4611 CONTINUE
C
C COMPtTE THE ERROR VARIANCE FROM REPLICATED DATA
IFI(.N)-.REPS) G0 TO 480
IGOTO = 1
IFI(NAAY(I REP),GT.1) IGOTO=2
IFI(NAAY(I REP),LE.1) WRITE(6,462) IREP
DO 475 I=1,NODEP
IFI(I-1) 4629,4629,464
4629 DO 463 K=1,NOTERM
IFI(X(K).NE.XCHK(K)) GO TO 2001
463 CONTINUE
464 CONTINUE
KBAR=NOTERM+1
S(I)=S(I)+X(KBAR)
SSQ(I)=SSQ(I)+X(KBAR)**2
IFI(IC).LE.1) 475,465,465
465 GO TO (468,466,IGOTO
466 ZEAN(I)=S(I)/FLOAT(NARAY(I REP))
SSQ(I)=SSQ(I)-ZEAN(I)*S(I)
P0OLED(I)=P0OLED(I)+SSQ(I)
WRITE LIST,467) I,SSQ(I),S(I),ZEOA(I)
468 IF(I.LT.NODEP) GO TO 469
NPDEG=NPDEG+NARAY(I REP)-1
IREP=IREP+1
IC = IC + NARAY(I REP)
WRITE LIST,467)
469 S(I)=0.0
SSQ(I)=0.0
XSAVE=.TRUE.
475 CONTINUE
C******************************************************************************
C CALCULATE SUMS, SUMS OF SQUARES AND SUMS OF CROSS PRODUCTS.
C******************************************************************************
480 CALL SUMUP
490 CONTINUE
C 490 CONTINUE IS THE END OF THE LOOP FOR READING DATA CARDS
IFI(.N)-.REPS) G0 TO 496
DO 493 I=1,NODEP
REPVAR(I)=P0OLED(I)/FLOAT(NPDEG)
493 CONTINUE
496 CONTINUE
C******************************************************************************
C ALL DATA HAS BEEN READ IN AND THE XTRANSPOSEX AND XTRANSPOSEY
C MATRIX HAVE BEEN CALCULATED.
C NOW WRITE THE MATRICES
CALL XFI
WRITE INTER
GO TO 640
C******************************************************************************
C THIS CODING DELETES THE DATA FROM THE SUMXX MATRIX
C CORRESPONDING TO THE INDEPENDENT TERM DELETED
6500 CONTINUE
IR=IODT-1
IC=NOTERM-10UT
IFI (IC.EQ.0) GO TO 6700
INOCH= IOUT+IR/2
INew = INOCH
IOL(I)=INEX+IOUT
IRC=0
IBC=0
ITC=0

DO 6630 I=IOLD,LENGTH
INEW=INEW+1
IOLD=IOLD+1
IF(ITC.GT.0) GO TO 6540
IRC=IRC+1
IF(IRC.GT.IR) GO TO 6530
SUMXX(INEW)=SUMXX(IOLD)
GO TO 6600
6530 IRC=IRC+1
IOLD=IOLD+1
IRC=J
6540 ITC=ITC-1
SUMXX(INEW)=SUMXX(IOLD)
6600 CONTINUE
6700 LENGTH=LENGTH-NOTERM
NOTERM=NOTERM-1
JCOL=VOTERM+NODEP
C
C*****************************************************************************
C INVERT THE SUMXX MATRIX AND COMPUTE REGRESSION COEFS
C AND SJS OF SQUARES DUE TO REGRESSION IN THE MATRIX INVERSION
C ROUTINE
C 640 CONTINUE
C CALL MATINV
C FIRST=.FALSE.
C
C*****************************************************************************
C WRITE(XXI INVERSE. THIS MATRIX TIMES ERROR MEAN SQUARE (ERRMS)
C IS THE VARIANCE-COVARIANCE MATRIX OF REGRESSION COEFFICIENTS.
C IF(ECONMY) GO TO 970
C WRITE(LIST,700)
C CALL TRIANG(X,SUMXI,NOTERM,8,FMTTRI,2)
C
C*****************************************************************************
C IF A VARIABLE HAS BEEN DELETED ADJUST COUNTERS AND RECOMPUTE THE
C REGRESSION. IF NO VARIABLE HAS BEEN DELETED CONTROL WILL PASS
C FROM THE TTEST ROUTINE TO THE CHI-SQUARE OPTION.
C 970 CONTINUE
C IF(.NOT.IFFT) GO TO 1020
C 980 WRITE (LIST,301)IDENT
C CALL TTEST($1020,NTKEEP)
C IF(NODEP=1) 985,990,985
C 985 WRITE(LIST,986) NODEP
C NODEP=1
C 990 J=JCOL-1
C DO 995 K=IOUT,J
C NTERM(K)=NTERM(K+1)
C ZEAN(K)=ZEAN(K+1)
C SUMX(K)=SUMX(K+1)
C SUMX2(K)=SUMX2(K+1)
C IDOUT(K)=IDOUT(K+1)
C SUMXY(K,1)=SUMXY(K+1,1)
C 995 CONTINUE
C IF(NOTERM.EQ.1) GO TO 1000
C GO TO 6500
C 1000 WRITE(LIST,1005)
C NOTERM=0
C GO TO 1035
C
C*****************************************************************************
C
1020 IF(.NOT.IFCHI1) GO TO 1035
1030 WRITE(LIST,301) IDENT
     CALL JUTPLT(PNCH)
C
C*****************************************************************************
1035 READ(INPUT5,117) PREDICT
     IF(.NOT.PREDICT) GO TO 100
     CALL PREDIC
1040 GO TO 100
C
C*****************************************************************************
2001 WRITE(LIST,1306)
     STOP
C
C*****************************************************************************
8001 FORMAT(1H1)
8002 FORMAT(1H2)
  110 FORMAT (I2,13A6)
  111 FORMAT (1HI,13A6,A2)
  117 FORMAT(7I1,1I,1I)
  118 FORMAT(1H,7I1,1I,1I)
  170 FORMAT(33H THERE IS NO BO TERM IN THE MODEL)
  190 FORMAT(26H THERE IS A BO TO ESTIMATE )
  230 FORMAT(40I2)
  235 FORMAT(1H,NTERM(K) = /(1H 30I4))
  240 FORMAT(25H THE TRANSFORMATIONS ARE /(1H 5(4I4,5X)))
  257 FORMAT(11I1, F3.3)
  260 FORMAT(5E15.7)
  262 FORMAT(19H THE CONSTANTS ARE /(1H 8G15.7))
  282 FORMAT(20I4)
  283 FORMAT(1H 20I4)
  284 FORMAT(1H THERE ARE I5,16H REPLICATE SETS )
  300 FORMAT(13A6,1A2)
  301 FORMAT(1H 13A6,A2)
  370 FORMAT(1H0,29H OBSERVED VARIABLES, WEIGHT = G14.6,6X,15H OBSERVATION
        1 = ,15)
  380 FORMAT(1H 9G14.6)
  460 FORMAT(1H 37 TERMS OF THE EQUATION, OBSERVATION = ,15)
  462 FORMAT(18H,K** REPlicate SET I5,3X,100(1H*))
  4671 FORMAT(1H 125(1H*))
  467 FORMAT(14H DEP. VAR. I6,8H SSQ=G14.7,8H SUM=G14.7,8H M 
        XEAN= G14.7)
  540 FORMAT(1H 8G14.7)
  560 FORMAT(21H2X TRANPOSE X MATRIX )
  670 FORMAT(25H2CORRELATION COEFFICIENTS )
  700 FORMAT(32H2X TRANPOSE X) INVERSE MATRIX )
  986 FORMAT(39H THE NUMBER OF DEPENDENT VARIABLES WAS I3,83H IT IS BE 
        XING SET TO ONE AND THE REJECTION OPTION EXERCISED ON DEPENDENT VAR 
        XIABLE 1 )
1005 FORMAT(39H THERE IS NO EVIDENCE OF A REGRESSION, / 
        X 74H USE THE MEAN RESPONSE FOR THE BEST ESTIMATE OF THE DEPEND 
        XENT VARIABLE(S))
  1282 FORMAT(314,15,14)
  1283 FORMAT(1H 314,15)
  1306 FORMAT(40H REPLICATE SETS ARE NOT GROUPED PROPERLY )
  1307 FORMAT(11H THE FIRST I2,64H TERMS OF THE MODEL WILL BE RETAINED R 
        XEGARDLESS OF SIGNIFICANCE )
END
SUBROUTINE MATINV

C SUBROUTINE MATINV

C*******************************************************************************
C PURPOSE
C 1) COMPUTE EIGENVALUES AND EIGENVECTORS OF (X-TRANSPOSE X)
C MATRIX IF REQUESTED. (STORYX=.TRUE.)
C 2) COMPUTE (X TRANSPOSE X) INVERSE
C 3) COMPUTE REGRESSION COEFFICIENTS
C 4) COMPUTE OTHER REGRESSION STATISTICS
C
C SUBROUTINES NEEDED
C BORD
C LOC
C EIGEN
C INVXTX
C RECT
C RSTATS
C TRIANG
C
C REMARKS
C THE EIGENVALUES ARE COMPUTED AS AN AID IN DETERMINING THE
C CONDITION OF THE SYSTEM OF EQUATIONS FOR THE REGRESSION
C COEFFICIENTS. EXAMINATION OF THEM AND THEIR ASSOCIATED
C EIGENVECTORS MAY SHOW THAT CERTAIN SETS OF INDEPENDENT
C VARIABLES ARE HIGHLY CORRELATED AND NOT EASILY LIABLE TO
C INDEPENDENT STUDY.

C*******************************************************************************
COMMON/BIG/B(60,9),SUMXY(60,9),SUMXX(1830),SUMXXI(1830)
COMMON/MED/M(69),SUMX(69),SUMX2(69),SUMY2(9),ZEAN(69),
COMMON/SMALL/R(69),SUMX2(9),SUMYZ(9),X(9),Y(9),Z(9),
COMMON/FMTS/FMT(13),FMTTRI(14)
COMMON/BYPASS,BZERO,DELETE,FIRST,IFCHI,IFSSR,
COMMON/IFIT,IFWT,INPUT,INTER,
COMMON/ISTRAT,JCOL,KONNO,LIST,
COMMON/NERROR,NODEP,NOOB,NOTERM,
COMMON/NOVAR,NPDEG,NRES,NTRANS,NWHERE,
COMMON/P,PREDCT,REPS,RWT,
COMMON/STORYI,STORYC,STORYX,TOTWT,WEIGHT,
COMMON/EKRFKD,ECONMY,IOUT,ICOL

INTEGER A

DOUBLE PRECISION B, SUMXY, SUMXX, SUMXXI,
DOUBLE PRECISION R, SUMX, SUMX2, SUMY2, ZEAN,
DOUBLE PRECISION L, SUMX2, SUMXX, SUMXXI

DATA IXTX(I), I=1,3 / 6HX TRANS, 6HPOSE, 6HX /
10  SUMXXI(I)= 1.0/SUMX2(I)
    GO TO 350
C
12  IF(.,NJ,T, STORYX) GO TO 30
    DO 14 I=1,LENGTH
    A(I)=SUMXX(I)
    14  CONTINUE
16  CALL EIGEN(A,C      ,IORDER,0)
    WRITE(LIST,17)(XTX(I),I=1,3)
    J=0
    DO 18 I=1,IORDER
    J=J+1
18  A(I)=A(J)
    WRITE(LIST,19) (A(I),I=1,IORDER)
    WRITE(LIST,20)
    CALL RECT(IORDER,IORDER,IORDER,IORDER,C,X ,FMTTRI,1)
30  DO 35 I=1,LENGTH
35  SUMXXI(I)=SUMXXII)
C
C********************************************************************
C NO SJBMODELS TO ANALYZE SO INVERT A DIRECTLY BY GAUSS
49  IF(IFSRR) GO TO 50
    CALL INVXTX(SUMXXI,NTERM,D,1.0)
    GO TO 60
C
C********************************************************************
C SUBMODELS HAVE BEEN REQUESTED SO WE USE BORDERING
50  IORDER=0
55  IORDER=IORDER +1
    CALL BORD(IORDER, SUMXXI)
60  CONTINUE
C
C********************************************************************
C COMPUTE COEFFICIENTS AND PRINT THEM
350 DO 373 J=1,NODEP
    DO 373 K=1,IORDER
    B(K,J)=0.000
    DO 373 L=1,IORDER
    CALL LOC(L,K,IR)
    B(K,J) = B(K,J) + SUMXXI(IR) * SUMXY(L,J)
370  CONTINUE
C
    WRITE(LIST,380) IDENT
    WRITE(LIST,382)
    IF(.,NJ,T,BZERO) GO TO 400
    DO 390 J=1,NODEP
    SUM=0.000
    KBAR= NOTERM + J
    DO 385 K=1,IORDER
    SUM = SUM + B(K,J) * ZEAN(K)
385  CONTINUE
    BO(J) = ZEAN(KBAR) - SUM
390  CONTINUE
    WRITE(LIST,395)
    WRITE(LIST,397) (BO(K),K=1,NODEP)
400 WRITE(LIST,410)
    DO 430 J=1,IORDER
    WRITE(LIST,432) IDOUT(J), (B(J,K),K=1,NODEP)
430  CONTINUE
C
C******************************************************************************
C COMPUTE REGRESSION STATISTICS IN RSTATS
C
CALL RSTATS(IORDER)

C******************************************************************************
C IF IORDER IS LESS THAN NOTERM WE HAVE USED THE BORDERING OPTION
C AND MUST GO BACK TO FINISH.
C
IF(IORDER-NOTERM) 55,500,500

500 STORYX=.FALSE.
IFSSR=.FALSE.
RETURN

17 FORMAT(34H2THE FOLLOWING ARE EIGENVALUES OF 2A6,A1, 7H MATRIX)
19 FORMAT(1H 8G16.7)
20 FORMAT(132H2THIS IS THE MODAL MATRIX OR MATRIX OF EIGENVECTORS. EIGENVECTORS ARE WRITTEN IN COLUMNS LEFT TO RIGHT IN SAME ORDER AS EIGENVALUES)
380 FORMAT(1H1,13A6,1A2)
382 FORMAT( 6111 EACH COLUMN CONTAINS THE COEFFICIENTS FOR ONE DEPENDENT VARIABLE)
395 FORMAT(2OH CONSTANT TERM)
397 FORMAT(4X,9G14.6)
410 FORMAT(36H REGRESSION COEFFICIENTS B1,...,BK)
432 FORMAT(1H 13,9G14.6)
END

$IBFC TT-STX

C******************************************************************************
C SUBROJTINE TTTEST
C
C PURPOSE
C COMPUTE THE T-STATISTICS FOR EACH REGRESSION TERM AND ITS TWO TAIRED SIGNIFICANCE LEVEL. THEN DETERMINE THE TERM WITH THE LEAST SIGNIFICANCE AND RETURN THIS INFORMATION TO NEWRAP
C
C******************************************************************************
C SUBROJTINE TTTEST(*,NTKEEP)
C
C******************************************************************************
COMMON /FRMTS/ FMT(13),FMTTRI(14)
COMMON/BIG/B(60,9),SUMXY(60,9),SUMXX(1830),SUMXX(1830)
X,DUMMY(1)
DOUBLE PRECISION B,SUMXY,SUMXX,SUMXI
COMMON/MED/B0(9),SUMX(69),SUMX2(69),SUMY(1830),ZEN(69),
X,CON(99),ERRMS(9),IDENT(1),IDOUT(60),NCON(200),NTERM(69),
X,NTRAN(100),NXCOND(100),POOLED(9),REPVAR(9),RESMS(9),X(99)
DOUBLE PRECISION BO,SUMX,SUMXX2,SUMY2,ZEAN
COMMON/SMALL/ BYPASS,BZERO,DELETE,FIRST,IFCHI,IFSSR,
X,IFWT,INPUT,INPUTS,INTER,
X,ISTRAT,JCOL,KONNO,LENGTH,LIS
X,ERROR,NOUEP,NOOB,NTERM,
X,VOVAR,NDVDEG,NRES,NTCONS,NWHERE,
X,P,PREDICT,REPS,RWT,
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>X STORY, STORYC, STORYX, TOTWT, WEIGHT</td>
<td>29</td>
<td>X ERRFXD, ECONMY, IOUT, ICOL</td>
<td>30</td>
<td>LOGICAL ECONMY</td>
</tr>
<tr>
<td>LOGICAL SATRDO</td>
<td>32</td>
<td>LOGICAL BYPASS, BZERO, DELETE, IFCHI, XIFSSR, IFTI, IFWT, REPS, PREDCT, XSTORY, STORYX, STORYI, FIRST, ERRFXD</td>
<td>33</td>
<td>DOUBLE PRECISION RWT, TOTWT, WEIGHT</td>
</tr>
</tbody>
</table>
C T(I,J,J) IS THE T-STATISTIC AT THE TABULATED DEGREES OF FREEDOM
C (I) AND AT THE TABULATED PROBABILITY LEVELS (JJ).
C II=31 IS FOR 40 DEGREES
C II=32 IS FOR 60
C II=33 IS FOR 120
C II=34 IS FOR INFINITY
C
C CALCULATE T STATISTICS
C
C DATA (T(I,J,J)*JJ=13)
C
C THE T-STATISTIC AT THE TABULATED DEGREES OF FREEDOM
C
C T(I,J,J) IS THE T-STATISTIC AT THE TABULATED DEGREES OF FREEDOM
C (I) AND AT THE TABULATED PROBABILITY LEVELS (JJ).
C II=31 IS FOR 40 DEGREES
C II=32 IS FOR 60
C II=33 IS FOR 120
C II=34 IS FOR INFINITY
C
C JJ PROBABILITY LEVEL   *   JJ PROBABILITY LEVEL
C 1   0.10   *   8   0.80
C 2   0.20   *   9   0.90
C 3   0.30   *   10  0.95
C 4   0.40   *   11  0.98
C 5   0.50   *   12  0.99
C 6   0.60   *   13  0.999
C 7   0.70
C
C**********************************************************************************************
C CALCULATE T STATISTICS
C
C 220 WRITE (LIST,230)
C 230 FORMAT(1HO,23HCALCULATED T STATISTICS /75H THE T STATISTICS CAN BE
C 1 USED TO TEST THE NET REGRESSION COEFFICIENTS B(I). )
C DO 260 J=1,NOTERM
C DO 240 K=1,NODEP
C TT(J,K)=ABS(B(J,K)/DEVB(J,K))
C 240 CONTINUE
C 250 CONTINUE
C 260 CONTINUE
C
C**********************************************************************************************
C SEARCH THE TABLE OF TABULATED DEGREES OF FREEDOM
C
C**********************************************************************************************
MAKEJ=FALSE
IF(NDEG-30)<90,290,300
290 II=NDEG
GO TO 400
300 IF(NDEG-40)<310,320,330
310 FINV=1.0/40.0
FM1INV=1.0/30.0
MAKEJ=TRUE.
320 II=31
GO TO 400
330 IF(NDEG-60)<340,350,360
340 FINV=1.0/60.0
FM1INV=1.0/40.0
MAKEJ=TRUE.
350 II=32
GO TO 400
360 IF(NDEG-120)<370,380,390
370 FINV=1.0/120.0
FM1INV=1.0/60.0
MAKEJ=TRUE.
380 II=33
GO TO 400
390 II=34
FINV=0.0
FM1INV=1.0/120.0
MAKEJ=TRUE.
C
C 400 WRITE(LIST,410)
410 FORMAT(104H UNDER NULL HYPOTHESIS THE INTERVAL (-T,T) WHERE T IS G
XIVEN ABOVE, HAS APPROX PROBABILITY LISTED BELOW. /42H MINUS S
XIGN INDICATES PROB EXCEEDS .999. )
415 IF(.NOT.MAKENU) GO TO 430
420 FNDEG=NDEG
425 DO 440 JJ=1,13
426 T(35,JJ)=T(II,JJ)+((1.0/FNDEG-FINV)/(FM1INV-FINV))*T(II-1,JJ)
430 CONTINUE
435 II=35
440 CONTINUE
445 DO 540 J=1,NTERM
450 DO 540 K=1,NDEP
455 DO 440 JJ=1,13
460 JJ=JJ
465 IF(T(II,JJ)-T(J,K))440,450,460
470 CONTINUE
475 PROR(J,K)=-0.999
480 GO TO 540
490 PROB(J,K)=PLEVEL(JJ)
500 GO TO 540
510 IF(JJ,LE,9) GO TO 470
515 JJ1=JJ-2
520 JJ2=JJ-1
525 JJ3=JJ
530 GO TO 490
535 IF(JJ,LE,4) GO TO 480
540 JJ1=JJ-1
545 JJ2=JJ
550 JJ3=JJ+1
555 GO TO 490
560 JJ1=JJ
565 JJ2=JJ+1
570 JJ3=JJ+2
580
PERFORM A THREE-POINT LAGRANGE INTERPOLATION

\[ X = ALOG(TT(J, K)) \]
\[ X1 = ALOG(T(II, JJ1)) \]
\[ X2 = ALOG(T(II, JJ2)) \]
\[ X3 = ALOG(T(II, JJ3)) \]

IF \((TT(J, K) \leq 1.0)\) GO TO 500
\[ Y1 = ALOG(1.0 - PLEVEL(JJ1)) \]
\[ Y2 = ALOG(1.0 - PLEVEL(JJ2)) \]
\[ Y3 = ALOG(1.0 - PLEVEL(JJ3)) \]
GO TO 510

\[ Y1 = ALOG(PLEVEL(JJ1)) \]
\[ Y2 = ALOG(PLEVEL(JJ2)) \]
\[ Y3 = ALOG(PLEVEL(JJ3)) \]

\[ PROB(J, K) = \frac{(X - X2)(X - X3)Y1}{((X2 - X1)(X2 - X3) + ((X - X1)(X - X3))} + \frac{(X - X1)(X - X2)Y2}{((X2 - X1)(X2 - X3) + ((X - X1)(X - X3))} \]

IF \((TT(J, K) > 1.0)\) GO TO 520

\[ PROB(J, K) = \exp(\text{PROB}(J, K)) \]
GO TO 540

\[ PROB(J, K) = 1.0 - \exp(\text{PROB}(J, K)) \]

CONTINUE

WRITE THE PROBABILITIES (1.0 - ALPHA)

\[ \text{WRITE}(\text{LIST}(550)) \text{IDOUT}(J), (\text{PROB}(J, K), K = 1, \text{NUDEP}) \]

\[ \text{FORMAT}(1 \ H \ [3, 2 \ (8 \ X, F6.3)) \]

CONTINUE

LIST THE DESIRED VALUE OF PROBABILITY (PWant)

\[ \text{PERCENT} = \text{PWant} \times 100.0 \]
\[ \text{WRITE}(\text{LIST}(580)) \ \text{PERCENT} \]

\[ \text{FORMAT}(1 \ H 0, 36 \ H \ \text{THE DESIRED VALUE OF PROBABILITY IS}, F5.1, 8 \ \text{PERCENT} \)

DELETE THE TERM WITH THE LOWEST COMPUTED PROBABILITY IF THAT PROBABILITY IS LESS THAN THE DESIRED (PWant)

\[ \text{IF}(\text{NTKEEP} \times \text{DELET}) \text{GO TO 660} \]
\[ \text{IF}(\text{NTKEEP} \times \text{NOTERM}) \text{GO TO 660} \]
\[ \text{OUT} = J \]

\[ \text{AMIN} = \text{PWant} \]
\[ \text{JLO} = \text{MAX}(1, \text{NTKEEP}) \]
\[ \text{DO} 620 J = \text{JLO}, \text{NOTERM} \]
\[ \text{IF}(\text{ABS}(\text{PROB}(J, 1)) - \text{PWant}) \text{GO TO 600, 620, 620} \]
\[ \text{IF}(\text{ABS}(\text{PROB}(J, 1)) - \text{AMIN}) \text{GO TO 610, 620, 620} \]
\[ \text{AMIN} = \text{ABS}(\text{PROB}(J, 1)) \]
\[ \text{OUT} = J \]

CONTINUE

\[ \text{IF}(\text{IDOUT}) \text{GO TO 660, 660, 630} \]

\[ \text{WRITE}(\text{LIST}(580)) \text{IDOUT}(\text{OUT}) \]
\[ \text{FORMAT}(1 \ H 10 X, 11 \ H \ \text{THE TERM X(12,}, 18 \ H) \text{IS BEING DELETED}) \]
GO TO 670

ALL VARIABLES REMAINING HAVE BEEN CONCLUDED SIGNIFICANT

RETURN

RETURN

END
SUBROUTINE KSTATS

PURPOSE

1) Compute and print the analysis of variance tables on regression and lack-of-fit if appropriate.
2) Compute and print R-squared and standard error of estimate.
3) Compute and print sums of squares due to each variable if it were last to enter regression.
4) Compute and print the standard deviations of each regression coefficient.

SUBROUTINE RSTATS(IORDER)

COMMON/BIG/8(60,9),SUMXY(60,9),SUMXX(1830),SUMXII(1830)
  B,SUMXY,SUMXX,SUMXI
COMMON/MED/8(60,9),SUMXY(60,9),SUMX2(69),SUM2(9),ZEAN(69),
  CON(99),ERRMS(9),IDENT(13),IDOUT(60),NCON(200),NTerm(69),
  NTRAN(100),NPOLED(100),POoled(9),REPVAR(9),RESMS(9),X(99)
DOUBLE PRECISION BO, SUMX, SUMX2, SUMY2, ZEAN
COMMON /FRMT/ FMT(13)
 COMMON /SMALL/ BYPASS,BZERO,DELETE, FIRST, IFCHI, IFSSR,
  IFIT, IFWT, INPUTS, INTERVAL, INPUTS
  ISTRAT, JCOL, KONNO, LENGTH, LIST,
  ERROR, NODEP, NODB, NOTERM,
  NOVAR, NPDEG, NRES, NTRANS, NWHERE,
  P, PREDICT, REPS, RWT,
  STORY, STORY, STORYX, TOTWT, WEIGHT,
  ERRFXD, ECQMY, IOUT, ICOL
LOGICAL ECONMY
LOGICAL SATRTD
LOGICAL BYPASS, BZERO, DELETE, IFCHI, IFSSR,
  FIRST, IFIT, IFWT, REPS, PREDCT,
  STORY, STORYX, STORY,
DOUBLE PRECISION RWT, TOTWT, WEIGHT

DIMENSION SSQREG(9), SSQRES(9), REGMS(9),
  XLOF(9), XLOFMS(9), RATI0(9), RSQD(9), R(9),
  SSQSTL(9), DEVB(9)
EQUIVLENC (DUMMY(10), SSQRES), (DUMMY(19), REGMS), (DUMMY(37), XLOF)
  (DUMMY(46), XLOFMS), (DUMMY(55), FRATIO), (DUMMY(64), RSQD),
  (DUMMY(73), R), (DUMMY(82), SSQSTL), (DUMMY(91), DEVB),
  (DUMMY(100), SSQREG)
DOUBLE PRECISION SSQREG

COMPUTE DEGREES OF FREEDOM AND RECIPROCALS

NREG= IORDER
NTOT= IFIX(TOTWT)-1
IF(NTOT,BZERO) NTOT= NTOT+1
NRES= VTOT-NREG
NLUP= VRES- NPDEG
RNREG= 1.0/FLOAT(NREG)
IF(NRES,EQ.0) GO TO 980
RNREG= 1.0/FLOAT(NRES)
SATRTD= .FALSE.
IF(I3F.EQ.0) GO TO 90
RNL0F=1.0/FLOAT(NLOF)
GO TO 100
90 SATRD=.TRUE.*
100 CONTINUE
NXTFRM=10ORDER
RNDB=RMT

C******************************************************************************
C COMPJTE RESIDUAL SUM OF SQUARES, RESIDUAL VARIANCE, VARIANCE
C FROM REPLICATIONS IF APPROPRIATE, AND THE F-RATIO OF MEAN SQUARE
C LACK-OF-FIT AND MEAN SQUARE RESIDUALS.
C DO 213 J=1,NODEP
SSQRES(J)=O.O
DO 203 I=1,NXTERM
SSQRE(J)=SSQRE(J) + B(I,J)*SUMXY(I,J)
CONT I
203 CONTINUE
SSQRE(J)=SUMY2(J)-SSQREG(J)
REGMS(J)= SSQREG(J)* RNREG
RESMS(J)= SSQRES(J)*RNRES
RSQD(J)=SSQREG(J)/SUMY2(J)
R(J)=SRT(RSQD(J))
IF((.NOT. REPS).OR. SATRD) GO TO 210
XLOF(J)=SSQRES(J)-POOLD(J)
XLOFMS(J)= XLOF(J)*RNL0F
FRATI(J)=XLOFMS(J)/REPVAR(J)
210 CONTINUE

C******************************************************************************
C DETERMINE WHICH ESTIMATE OF SIGMA SQUARED SHOULD BE USED IN
C HYPOTHESIS TESTS. PUT THE PROPER ONE IN ERRMS AND SET ERKFXD
C TO TRUE IF THE PRESENT VALUE IS TO BE USED FOR ALL FOLLOWING
C TESTS AND T-STATISTICS.
IOUT=ISTRAT
IF(ERKFXD) GO TO 250
IF(ISTRAT .NE. 3) GO TO 214
211 DO 213 J=1,NODEP
ERRMS(J)= RESMS(J)
NEROR= NRES
IOUT=3
GO TO 250
214 IF(ISTRAT .NE. 1) GO TO 218
IF(.NOT. REPS) GO TO 211
DO 215 J=1,NODEP
215 ERRMS(J)= REPVAR(J)
NEROR= NPDEC
ERRFXD=.TRUE.*
IOUT=1
GO TO 250
218 IF(FIRST. .AND. (IORDER.EQ. NOTERM)) GO TO 220
GO TO 211
220 ERRFXD=.TRUE.*
DO 222 J=1,NODEP
222 ERRMS(J)= RESMS(J)
NEROR= NRES
ISTRAT=2
IOUT=2

C******************************************************************************
C WRITE ANOVA TABLES
250 DO 500 J=1,NODEP
IF(ERRMS(J) .EQ. 0.0) ERRMS(J)=1.0E-30
WRITE(LIST,1001) J

1001 FORMAT(2X,I3,3X,15A11,1X,5F4.1,1X,A2)
500 CONTINUE
WRITE(LIST,1002)
WRITE(LIST,1003) SSQREG(J), NREG, REGMS(J)
WRITE(LIST,1004) SSQRES(J), NRES, RESMS(J)
WRITE(LIST,1005)
WRITE(LIST,1006) SUMY2(J),NTOT
WRITE(LIST,1007)
WRITE(LIST,1500) RSQD(J), R(J)
STD=SQRT(RESMS(J))
WRITE(LIST,1600) STD
WRITE(LIST,1700) IOUT,ERRMS(J),NERROR
F=REGMS(J)/ERRMS(J)
WRITE(LIST,1750)F,NREG,NERROR
IF(.NOT.REPS).OR.SATRTD GO TO 500
WRITE(LIST,2001)
WRITE(LIST,1002)
WRITE(LIST,2005) XLOF(J), NLOF, XLOFMS(J)
WRITE(LIST,2006) POOLED(J), NPDEG, REPVAR(J)
WRITE(LIST,1004) SSQRES(J), NRES, RESMS(J)
WRITE(LIST,1005)
WRITE(LIST,2008) FRATIO(J)
WRITE(LIST,1007)
500 CONTINUE
C
C*******************************************************************************
C COMPUTE CONTRIBUTION OF EACH INDEPENDENT VARIABLE TO REG SUM
C OF SQUARES AS IF IT WERE LAST TO ENTER
WRITE(LIST,370)
IR=0
DO 8635 K=1,NXTERM
IR= IR+K
DO 8632 J=1,NODEP
8632 SSQSLST(J)= SUMXXI(IR)
8635 CONTINUE
C
C*******************************************************************************
C COMPUTE STANDARD DEVIATION OF REGRESSION COEFFICIENTS
WRITE(LIST,375)
IF(.NOT.BZERO) GO TO 959
DO 910 J=1,NXTERM
R(J)=0.0
DO 910 I=1,NXTERM
CALL LOC(I,J,IR)
R(J)=R(J)+ZEAN(I)*SUMXXI(IR)
910 CONTINUE
XXT=0.0
DO 920 J=1,NXTERM
920 XXT=XT+ZEAN(J)*R(J)
DO 930 K=1,NODEP
930 DEVB(1,K)=SQRT(ERRMS(K)*(RNOOB+XXT))
K=0
WRITE(LIST,380) K,(DEVB(1,J),J=1,NODEP)
959 IR=1
DO 970 J=1,NXTERM
970 CONTINUE
970 CONTINUE
WRITE(LIST,380) IOUT(J),(DEVB(J,K),K=1,NODEP)
C
RETURN
C
C FORMATS
1001 FORMAT(42H4ANOVA OF REGRESSION ON DEPENDENT VARIABLE 15)
1002 FORMAT(1H 79(1H*)/79H SOURCE SUMS OF SQUARES DEG
XREES OF FREEDOM MEAN SQUARES /1H 79(1H*)
1003 FORMAT(17H REGRESSION G20.8, 5X, I10, 5X, G20.8)
1004 FORMAT(17H RESIDUAL G20.8, 5X, I10, 5X, G20.8)
1005 FORMAT(1H 79(1H*))
1006 FORMAT(17H TOTAL G20.8, 5X, I10)
1007 FORMAT(1H 79(1H*))
2001 FORMAT(1H/1H/22H ANOVA OF LACK OF FIT )
2005 FORMAT(17H LACK OF FIT G20.8, 5X, I10, 5X, G20.8)
2006 FORMAT(17H REPLICATION G20.8, 5X, I10, 5X, G20.8)
2008 FORMAT(28H F = MS(LOF)/MS(REPS) = F10.3 )
370 FORMAT(74H1 SUMS OF SQUARES DUE TO EACH VARIABLE IF IT WERE LAST T
X0 ENTER REGRESSION )
375 FORMAT(115H2 STANDARD DEVIATION OF REGRESSION COEFFICIENTS (DERIVE
X0 FROM DIAGONAL ELEMENTS OF (X TRANSPOSE X)INVERSE MATRIX ))
380 FORMAT(1H I13,9G14.6)
1500 FORMAT(40H R SQUARED = SSQ(REG) / SSQ(TOT) = F8.6,
X 5X, 4HR = F7.6)
1600 FORMAT(34H STANDARD ERROR OF ESTIMATE G14.6)
1700 FORMAT(24H USING POOLING STRATEGY I2,25H THE ERROR MEAN SQUARE =
X G14.7, 26H WITH DEGREES OF FREEDOM = I6)
1750 FORMAT(5X,19HF = MS(REG)/MS(ERR) = F6.2, 5X, 13HCOMPARE TO F(12,1H,13,1
XH))
980 WRITE(LIS1981
981 FORMAT(4H ZERO RESIDUAL DEGREES OF FREEDOM. STOP.
END

LIBFTC RECT

SUBROUTINE RECT(IROW, JJCOL, IMAX, JMAX, A, B, FMT, 11)
DIMENSION A(IMAX, JMAX), FMT(14), XOUT(8)
DOUBLE PRECISION B, DXOUT
DIMENSION B(IMAX, JMAX), DXOUT(8)
COMMON/SMA1L/DUM(15), LIST
DATA J8/8/
LOGICAL OUT
OUT = .FALSE.
JTIMES=0
JCOL=JJCOL
5 JNXT=JCOL-J8
IF(JNXT) 10,20,30
10 JP=JCOL
GO TO 40
20 JP=J8
GO TO 40
30 JCOL=JNXT
J8=J8
GO TO 50
40 OUT=.TRUE.
50 DO 10 I=1, IROW
GO TO (55,75), II
55 CONTINUE
DO 60 J=1, JP
JJ=JTIMES +J
60 CONTINUE

981
SUBROUTINE PREDIC

PURPOSE
1) READ INPUT LEVELS OF INDEPENDENT VARIABLES AND COMPUTE
   A PREDICTED RESPONSE FROM THE ESTIMATED REGRESSION EQUATION.
2) COMPUTE VARIANCE AND STANDARD DEVIATION OF THE PREDICTED
   MEAN VALUE AND A SINGLE FURTHER OBSERVATION.

SUBROUTINES NEEDED
TRANS
LOC

REMARKS
VALUES FOR DEPENDENT VARIABLES ARE NOT NECESSARY FOR THE
PREDICTING OF VALUES. HOWEVER, A DUMMY VALUE MAY NEED TO
BE SUPPLIED IF A ZERO (BLANK) INPUT VALUE WILL CAUSE AN
IMPOSSIBLE OPERATION TO BE ATTEMPTED DURING THE
TRANSFORMATIONS.

******************************************************************************

SUBROUTINE PREDIC
COMMON /FRMTS/, FMT(13), FMTTRI(14)
COMMON/BIG/B(60,9), SUMXY(60,9), SUMXX(1830), SUMXXI(1830)
DOUBLE PRECISION B, SUMXX, SUMXXI
COMMON/MED/80(9), SUMX(69), SUMXX(69), SUMY(9), ZEAN(69),
X CON(99), ERRMS(9), IDENT(13), IDOUT(60), NCOL(200), NTERM(69),
X NTRAN(100), NXCND(100), POOLED(9), REPVAR(9), RESMS(9), X(99)
DOUBLE PRECISION 80, SUMX, SUMXX, SUMXXI
COMMON/SMALL/, BYPASS, BZERO, DELETE, FIRST, IFCH1, IFSR,
X IFTT, IFWT, INPUT, INPUT5, INTER,
X ISTAT, JCOL, KUNNO, LENGTH, LIST,
X NERROR, NODEP, NOOB, NOTERM,
X NOVAR, NPMDEG, NRES, NTRANS, WHERE,
X P, PREDICT, REPS, RMT,
X STORI, STORC, STORYX, FI WT, WEIGHT,
X ERFXO, ECUUMY, IOUT, ICOL
LOGICAL ECOUMY
LOGICAL BYPASS, BZERO, DELETE, IFCH1, IFSR,
X IFSSR, IFTT, IFWT, REPS, PREDICT,
X STORIC, STORCX, STORYI, FIRST, ERFXO

60 XOUT(J) = A(I, JJ)
WRITE(List, FMT) 1, (XOUT(K), K=1, JP)
GO TO 100
75 DO 70 J=1, JP
JJ = JTIMES + J
80 WRITE(List, FMT) 1, (XOUT(J), J=1, JP)
CONTINUE
100 IF(OUT) RETURN
WRITE(List, 110)
110 FORMAT(1H /1H
 JTIMES = JTIMES + JP
GO TO 5
END

**IHTC PRO-DIX**
DOUBLE PRECISION WEIGHT, RWT, TOTWT
DIMENSION YCALC(9), V(60), VARM(9), SEEM(9),
X, VARP(9), SEEP(9)
DOUBLE PRECISION XXT, V
EQUIVALENCE (YCALC(I), SUMXX(I)), (V(I), SUMXX(150)),
(X, (VARM(I), SUMXX(71)), (SEEM(I), SUMXX(80)), (VARP(I), SUMXX(89))
X, (SEEP(I), SUMXX(98))
EQUIVALENCE (RNOOB, RWT)

C
C*****************************************************************************
C
C IF( NOTERM .EQ. 0 ) RETURN
WRITE(LIST, 3)
READ(INPUT5, 5) NPRED
C
DO 503 KK = 1, NPRED
105 READ(INPUT5, FMT)(X(I), I = 1, ICOL)
WRITE(LIST, 110)(X(I), I = 1, ICOL)
125 CALL TRANS
DO 132 J = 1, JCOL
I = NOTERM(K)
XXT = CON(I)
130 CONTINUE
WRITE(LIST, 135)(X(I), I = 1, NOTERM)
C
C COMPUTE PREDICTED RESPONSE
140 DO 153 K = 1, NODEP
YCALC(K) = BO(K)
IF( NOTERM .EQ. 0 ) YCALC(K) = .0
DO 152 J = 1, NOTERM
YCALC(K) = YCALC(K) + B(J, K) * X(J)
150 CONTINUE
C
C COMPUTE VARIANCE AND STANDARD DEVIATION OF REGRESSION LINE
C AND VARIANCE AND STANDARD DEVIATION OF PREDICTED VALUE
C AT THE POINT XO
C
DO 253 K = 1, NODEP
V(K) = 0.000
DO 252 J = 1, NOTERM
CALL LOC(J, K, IR)
V(K) = V(K) + (X(J) - ZEAN(J)) * SUMXX(I, IR)
250 CONTINUE
XXT = 0.000
DO 273 K = 1, NODEP
XXT = XXT + (X(K) - ZEAN(K)) * V(K)
275 CONTINUE
XRNODB = RNOOB
IF( NOTERM .EQ. 0 ) XRNODB = 0.0
DO 300 K = 1, NODEP
VARM(K) = ERRMS(K) * (XRNODB + XXT)
SEEM(K) = SQRT(VARM(K))
VARP(K) = ERRMS(K) + VARM(K)
SEEP(K) = SQRT(VARP(K))
300 CONTINUE
WRITE(LIST, 310)(YCALC(K), K = 1, NODEP)
WRITE(LIST, 320)(VARM(K), K = 1, NODEP)
WRITE(LIST, 320)(SEEM(K), K = 1, NODEP)
WRITE(LIST, 320)(VARP(K), K = 1, NODEP)
WRITE(LIST, 320)(SEEP(K), K = 1, NODEP)
C
 SUBROUTINE SUMUPS

C PURPOSE
  1) CALCULATE (X TRANSPOSE X) AND (X TRANSPOSE Y) MATRICES ON
      OBSERVATION AT A TIME.
  2) COMPUTE TOTAL OF THE WEIGHTS
   ** BOTH CALCULATIONS ARE IN DOUBLE PRECISION

C SUBROUTINES NEEDED
  LOC

C******************************************************************************

SUBROUTINE SUMUP
COMMON/BJG/B(60,9),SUMXY(60,9),SUMXX(1830),SUMXI(1830)
DOUBLE PRECISION B,SUMXY,SUMXX
COMMON/MED/BO(9),SUMX(69),SUMZ(69),SUMY2(9),ZER(69),
  X CON(99),ERRMS(9),IDENT(13),IDOUT(60),NCON(200),NTERM(69),
  X NTRAN(100),NXC0D(100),POLED(9),REPVAR(9),RESMS(9),X(99)
DOUBLE PRECISION DUB1,DUB2
COMMON/SMALL/ BYPASS,BZERO,DELETE, FIRST, IFCHI, IFSSK,
  X IFTT, IFWT, INPUT, INPUT5, INTER,
  X ISTRAT, JCOL, KONNO, LENGTH, LIST,
  X NERROR, NDDEP, NOOB, NOTERM,
  X NODAR, NPDAG, NRES, NTRANS, NWHERE,
  X P, PREDCT, REPS, RWT,
  X STORYI, STORYG, STORYX, TOFWT, WEIGHT,
  X ERRFXD, ECONMY, IOUT, ICOL
LOGICAL ECONMY
LOGICAL BYPASS, BZERO, DELETE, IFCHI,
  X IFSSR, IFTT, IFWT, REPS, PREDCT,
XSTORYG, STORYX, STORYI, FIRST, ERRFXD
DOUBLE PRECISION RWT,TOFWT,WEIGHT
DOUBLE PRECISION DUB1,DUB2

C******************************************************************************

DO 113 I=1,JCOL
  SUMX(I)=SUMX(I)+X(I)*WEIGHT
CONTINUE
IR=0
DO 100 K=1,NOTERM
DUB1=X(K)
DO 90 J=1,NODEP
KBAR=J+NOTERM
DUB2=X(KBAR)
SUMXY(K,J)=SUMXY(K,J)+DUB1*DUB2*WEIGHT
90 CONTINUE
DO 50 I=1,K
IR=IR+L
DUBZ=K(I)
SUMXX(IR)=SUMXX(IR)+DUB1*DUB2*WEIGHT
50 CONTINUE
DO 15 J=1,NODEP
KBAR=NOTERM + J
DUB1=X(KBAR)
SUMY2(J)=SUMY2(J)+DUB1*DUB2*WEIGHT
15 CONTINUE
TOTWT=TOTWT+WEIGHT
RETURN
END

SUBROJTINE BORD

PURPOSE
TO COMPLETE THE INVERSION OF A SYMMETRIC POSITIVE DEFINITE
MATRIX A OF ORDER N GIVEN THAT THE UPPER LEFT SUB-
MATRIX OF ORDER N-1 HAS ALREADY BEEN INVERTED.

SUBROJTINES NEEDED
LOC

REMARKS
ONLY THE UPPER TRIANGULAR PART OF A IS STORED AS A
VECTOR IN THE ORDER A(1,1),A(1,2),A(2,2),A(1,3),...ETC
SUBROJTINE BORD(IORDER,A)

DIMENSION BETA(60),A(1)
DOUBLE PRECISION A,ALPHA,RALPHA ,BETA

ALPHA= 0.000
NM1= IORDER-1
IF(NM1) 100,100,200
100 A(1) = 1.0/A(1)
GO TO 600
200 M=NM1*(NM1+1)/2
LEN = M + IORDER
DO 400 I=1,NM1

39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
BETA(I) = 0.000
MI = M+I
DO 350 J=1,NM1
CALL LOG(I,J,II)
MJ = M+J
BETA(I) = BETA(I) - A(II) * A(MJ)
350 CONTINUE
ALPHA = ALPHA + A(MI) * BETA(I)
400 CONTINUE

C
C
ALPHA = ALPHA + A(LEN)
RALPHA = 1.000 / ALPHA
A(LEN) = RALPHA
C
DO 500 I=1,NM1
DO 500 J=1,1
CALL LOG(I,J,II)
A(II) = A(II) + BETA(I) * BETA(J) * RALPHA
500 CONTINUE
C
DO 550 J=1,NM1
MJ = M+J
A(MJ) = BETA(J) * RALPHA
550 CONTINUE
C
600 CONTINUE
RETURN
END

$IBFTC MFIXXX

SUBROUTINE MFIX

C******************************************************************************
COMMON /FRMTS/ FMT(13), FMTTRI(14)
COMMON/BIG/B(60,9), SUMXY(60,9), SUMXX(1830), SUMXI(1830)
DOUBLE PRECISION B, SUMXY, SUMXX, SUMXI
COMMON/MED/BO(9), SUMX(69), SUMX2(69), SUMY(9), ZEAN(69),
X,NCON(100), NTERM(99), NCON200, NTERM69,
X, NTRAN(100), NXCOD(100), POOLED(9), REPVAR(9), RESMS(9), X(99)
DOUBLE PRECISION BO, SUMX, SUMX2, SUMY, ZEAN
COMMON/SMALL/ BYPASS, BZERO, DELETE, FIRST, IFCHI, IFSSR,
X, IFTT, IFWT, INPUT, INPUT5, INTER,
X, ISTRAT, JCOL, KONNO, LENGTH, LIST,
X, NERROR, NODEP, NOOB, NOTERM,
X, NVOAR, NDEG, NRES, NTRANS, NWHERE,
X, P, PREDCT, REPS, RWT,
X, STORY, STORYX, STORYX, TOTWT, WEIGHT,
X, ERRFXD, ECONMY, IOUT, ICOL
LOGICAL ECONMY
DOUBLE PRECISION RWT, TOTWT, WEIGHT
LOGICAL BYPASS, BZERO, DELETE, IFCHI,
XIFSSR, IFTT, IFWT, REPS, PREDCT,
XSTORY, STORYX, STORYI, FIRST, ERRFXD
C******************************************************************************
C
IF(ECJNMY) GO TO 500
IF(NOT.BZERO) GO TO 500
WRITE(LIST,530)
WRITE (LIST,540) (SUMX(I),I=1,JCOL)
WRITE(LIST,560)
CALL TRIANG(X, SUMXX, NOTERM, B, FMTTRI, 2)
WRITE(LIST,565)
CALL RECT(NOTERM, NODEP, 60, 9, X, SUMXY, FMTTRI, 2)
C***********************************************************************
C COMPUTE AND PRINT MEANS. COMPUTE AND PRINT THE (X TRANSPOSE X)
C MATRIX IN TERMS OF DEVIATIONS FROM MEAN. THE DEVIATIONS FORM
C OF (X T X) IS THE VARIANCE-COVARIANCE MATRIX OF THE
C INDEPENDENT VARIABLES.
C***********************************************************************
500 CONTINUE
RWT=1.000/TUTWT
DO 570 I=1,JCOL
570 ZEAN(I)=SUMX(I)*RWT
WRITE(LIST,580)
WRITE(LIST,540) (ZEAN(I),I=1,JCOL)
IR = J
DO 600 J=1,NOTERM
IR=IR + J
IF(.NOT.BZERO) GO TO 601
SUMX2(J)=SUMXX(IR)-SUMX(J)**2*RWT
GO TO 600
601 SUMX2(J)=SUMXX(IR)
600 CONTINUE
602 CONTINUE
IR=1
DO 620 J=1,NOTERM
DO 618 K=1,NODEP
IF(.NOT.BZERO) GO TO 618
KBAR=NOTERM+K
SUMXY(J,K)=SUMXY(J,K)-SUMX(J)*SUMX(KBAR)*RWT
618 CONTINUE
619 DO 620 K=1,J
IF(.NOT.BZERO) GO TO 619
SUMXX(IR)=SUMXX(IR)-SUMX(K)*SUMX(J)*RWT
619 SUMXX(IR)=SUMXX(IR)/DSQRT(SUMX2(J)*SUMX2(K))
619 CONTINUE
620 CONTINUE
IF(.NOT.BZERO) GO TO 6220
DO 6210 J=1,NODEP
K=NOTERM+J
SUMY2(J)=SUMY2(J)-SUMX(K)**2*RWT
6210 CONTINUE
6220 CONTINUE
C***********************************************************************
IF(ECJNMY) GO TO 622
IF(.NOT.BZERO) GO TO 621
WRITE(LIST,625)
CALL TRIANG(X, SUMXX, NOTERM, B, FMTTRI, 2)
WRITE(LIST,630)
CALL RECT(NOTERM, NODEP, 6C, 9, X, SUMXY, FMTTRI, 2)
621 WRITE(LIST,670)
CALL TRIANG(X, SUMXXI, NOTERM, B, FMTTRI, 2)
622 CONTINUE
C***********************************************************************
530 FORMAT(1HO,32H SUMS OF INDEP AND DEP VARIABLES )
FORIYAT (1H 8615.7)
565 FORMAT (21H2X TRANSPOSE X MATRIX )
580 FORMAT (33H MEANS OF INDEP AND DEP VARIABLES )
625 FORMAT (53H2X TRANSPOSE X MATRIX WHERE X IS DEVIATION FROM MEAN )
630 FORMAT (60H2X TRANSPOSE Y MATRIX WHERE X AND Y ARE DEVIATIONS FROM XMEAN )
670 FORMAT (25H2CORRELATION COEFFICIENTS )
END

$IBFTC LOCXXX

SUBROUTINE LOC(I,J,IR)
IX= I
JX= J
20 IF(IX-JX) >= 22.24924
22 IRX= IX + (JX*JX-JX)/2
GO TO 36
24 IRX= JX + (IX*IX - IX)/2
36 IR= IX
RETURN
END

$IBFTC XTRANS

SUBROUTINE TRANS
C******************************************************************************
C COMMON/BIG/B(60,9),SUMXY(60,9),SUMXX(1830),SUMXXI(1830)
DOUBLE PRECISION B,SUMXY,SUMXX,SUMXI
COMMON/MED/B0(9),SUMX(69),SUMXX(69),SUMZ(69),ZMEN(69),
X CINV(99),EKRMS(9),IDENT(13),IDOUT(60),NCON(22),NTFRM(69),
X NTRAN(100),NXCD1(100),POOLED(9),REPVAR(9),REMS(9),X(99)
DOUBLE PRECISION BO,SUMX,SUMXX,SUMXI,LEN
COMMON/SMALL/ BYPASS, BZERO, DELTE, FIRST, IFCHI, IFSSR,
X IFTT, IFWT, INPUT, INPUT5, interfer,
X ISTRAT, JCOL, KUNNO, LENGTH, LIST,
X VERROR, NODEP, NVOB, NOTERM,
X WNUM, NPDEG, NRES, NTTRANS, NWHERE,
X P, PREDCT, REPS, RWT,
X STORI, STORY, STORY, STORY, STORY, TOWT, WEIGHT,
X ERRFXO, ECONMY, IDUT, ICOL
LOGICAL ECONMY
LOGICAL BYPASS, BZERO, DELTE, IFCHI,
X IFSSR, IFTT, IFWT, REPS, PREDCT,
X STORY, STORY, STORY, STORY, FIRST, FRFXO
DOUBLE PRECISION RWT, TOTWT, WEIGHT
C******************************************************************************
C THIS SUBROUTINE PERFORMS TRANSFORMATIONS IF THIS OPTION IS
C REQUESTED.
C******************************************************************************

70
TRANFORMATION SET NUMBER.

CONSTANT NUMBER TO USE.

DERIVED CONSTANT.

NUMBER OF TRANSFORMATION REQUESTED.

VARIABLE NUMBER

DO 300 K=1,NTRANS
I=NCON(2*K-1)
IF(I).EQ.100,110,10

100 CONS=0.0
GO TO 120

110 CONS=CON(I)

120 I=NXCID(K)

Y=X(I)

MTRAN = NTRAN(K)

IF(MTRAN.LE.0) MTRAN=32

140 GO TO(150,160,170,180,190,200,210,220,230,240,250,260,270,280,290,
X300,310,320,330,340,350,360,370,380,390,400,410,420,430,440,
X 442,450),MTRAN

150 CONS=Y+CONS
GO TO 460

160 CONS=Y*CONS
GO TO 460

170 CONS=CONS/Y
GO TO 460

180 CONS=EXP(Y)
GO TO 460

190 CONS=Y**CONS
GO TO 460

200 CONS=ALOG(Y)
GO TO 460

210 CONS=ALOG10(Y)
GO TO 460

220 CONS=SIN(Y)
GO TO 460

230 CONS=COS(Y)
GO TO 460

240 CONS=SIN(3.14159265*(CONS*Y) )
GO TO 460

250 CONS=COS(3.14159265*(CONS*Y) )
GO TO 460

260 CONS=1.0/Y
GO TO 460

270 CONS=EXP(CONS/Y)
GO TO 460

280 CONS=EXP(CONS/(Y*Y))
GO TO 460

290 CONS=SQRT(Y)
GO TO 460

300 CONS=1.0/SQRT(Y)
GO TO 460

310 CONS=CONS**Y
GO TO 460

320 CONS=10.0**Y
GO TO 460

330 CONS=SH(Y)
GO TO 460

340 CONS=COSH(Y)
GO TO 460

350 CONS=(1.0-COS(Y))/2.0
GO TO 460
360 CONS=ATAN(Y)
   GO TO 460
370 CONS=ATAN2(Y,CONS)
   GO TO 460
380 CONS=Y*Y
   GO TO 460
390 CONS=Y*Y*Y
   GO TO 460
400 CONS=ARSIN(SQRT(Y))
   GO TO 460
410 CONS=2.0*3.14159265*Y
   GO TO 460
420 CONS=1.0/(2.0*3.14159265*Y)
   GO TO 460
430 CONS=ERF(Y)
   GO TO 460
440 CONS=SAMMA(Y)
   GO TO 460
442 CONS=Y/CONS
   GO TO 460
450 CONS=Y
   GO TO 460
460 I=NCOV(2*K)
480 CON(I)=CONS
   IF(I-53)
      500,500,490
490 XI(I)=CONS
500 CONTINUE
   RETURN
   END

$IBFTC OUTPLX

C*******************************************************************************
C
C*******************************************************************************

C SUBROUTINE OUTPLT(PNCH)
COMMON/BIG/B(60,9),SUMXY(60,9),SUMXX(1830),SUMXI(1830)
DOUBLE PRECISION B, SUMXY, SUMXX, SUMXI
COMMON/MED/B0(9),SUMX(69),SUMX2(69),SUMY2(9),ZEN(69),
X CON(99),ERRMS(9),IDENT(13),IDOUT(60),NCON(200),NTERM(69),
X NTRAN(100),NXCOD(100),POOLED(9),REPVAR(9),RESMS(9),X(99)
DOUBLE PRECISION B0, SUMX, SUMX2, SUMY2, ZEN
COMMON/SMALL/ BYPASS, BZERO, DELETE, FIRST, IFCHI, IFSSR,
X IFTT, IFWT, INPUT, INPUT5, INTER,
X ISTRAT, JCOL, KUNNO, LENGTH, LIST,
X NERROR, NODEP, NOOB, NOTERM,
X NVAR, NPDEG, NRES, NTRANS, NWHEREF,
X P, PREDCT, REPS, RWT,
X STORYI, STORYC, STORYX, TOTWT, WEIGHT,
X ERRFXD, ECONMY, IOUT, ICOL
LOGICAL ECONMY
DOUBLE PRECISION RWT, TOTWT, WEIGHT
LOGICAL BYPASS, BZERO, DELETE, IFCHI,
XIFSSR, IFTT, IFWT, REPS, PREDCT,
XSTORYI, STORYX, STORYC, FIRST, ERRFXD
COMMON/MAX/MAXPLT
C
INTEGER CELLS, PLUS1
DIMENSION BOUND(45), CELLBD(21), OBS(20, 9), RCT(212), X STDERR(9), VAR(9), YCALC(9), YDIFR(9), XZ(9)

EQUIVALENCE (VAR, RESMS)

ODATA BOUND(67448907, 67448922, 67448930), OBS(20, 9), RCT(212), X STDERR(9), VAR(9), YCALC(9), YDIFR(9), XZ(9)

LOGICAL SAVRES

DIMENSION RESPLT(1)
DIMENSION SKEM(9), SKUR(9)
EQUIVALENCE (RESPLT, SUMXY)

DIMENSION ICHAR(9), HCHAR(9)

DATA(ICHAR(I) = I = 1, 9) / 6HRFSID, 6HALS, 6FD, 6HR DEP, 6HVAR, 1H

DATA(YCHAR(I) = I = 1, 9) / 6HRFSID, 6HALS, 6FD, 6HR DEP, 6HVAR, 1H

LOGICAL PNC

JCOL=VOTERM+ NODEP
NUVAR=VOTERM+1
BYPASS=. FALSE.
KOUNT= 0
SAVRES=. FALSE.
ITPLT=XOOB*(NWHERE+2*NODEP)

IF(ITPLT.GE.MAXPLT) SAVRES=. TRUE.

CALL LRLEGN(IDENT(10), 24, 0, 1, 4, 5, 1, 0)

IF(NO3B-20) 110, 120, 120

GO TO 125

120 CELLS=XOOB/5
CELLS=MINO(CELLS, 20)
I= MOD(CELLS, 2)
IF(I.EQ.0) CELLS=CELLS + 1
FCHELLS= FLOAT(CELLS)
PLUS1= CELLS + 1
MINUS1= CELLS - 1

NDEGC= CECCS-3
IR= CELLS/2-1
IC=IR*(IR-1)/2
IS=IR+2
DO 122 J=1, IR
JC=IC+1
IBC=IS-J
IRC=IS+J

CELLBD(I8C)= BOUND(1C)
CELLBD(I8C)= BOUND(1C)
CONTINUE
CELLBD(1)=-1.0E+37
CELLBD(PLUS1) =1.0E37
CELLBD(IS )=0.0
DO 124 K=1,NODEP
DO 124 I=1,CELLS
OBS(I,K)=0.0
124 CONTINUE

DO 130 K=1,NODEP
SKW(K)=0.0
SKUR(K)=0.0
STDER(K)= SQRT(ERRMS(K))
130 CONTINUE
WRITE(LIST,135)

DO 430 J=1,NOOB
READ(INTER) (X(I),I=1,69),WEIGHT
IF(.NOT.SAVRES) GO TO 141
INOPLT=NWHERE
DO 140 I=1,INOPLT
K=(I-1)*NOOB+J
140 RESPLT(K)=X(I)
141 CONTINUE
DO 142 I=1,NOTERM
K = IOOUT(I)
X(I)= X(K)
142 CONTINUE
KBAR=NWHERE
DO 143 I=1,NODEP
IC= NOTERM+ I
KBAR=KBAR+1
X(IC)= X(KBAR)
143 CONTINUE

DO 165 K= 1,NODEP
YCALK(K)= BO(K)
IF(.NOT.BZERO) YCALC(K)= 0.0
KBAR= K+NOTERM
DO 150 I=1,NOTERM
YCALK(K) = YCALC(K) + B(I,K)*X(I)
150 CONTINUE
ACTDEV= X(KBAR)- YCALC(K)
YDIFR(K) = ACTDEV
Z(K) = ACTDEV/STDER(K)
A=ACTDEV**3
SKW(K)=SKW(K)+A
SKUR(K)=SKUR(K)+A*ACTDEV
160 CONTINUE
IF(.NOT.SAVRES) GO TO 179
K=INOPLT*NOOB+J
KBAR=KBAR+NODEP
DO 175 I=1,NODEP
ITC=(I-1)*NOOB
ISC=K+ITC
IS=KBAR+ITC
RESPLT(ISC)=YCALK(I)
RESPLT(IS)=Z(I)
175 CONTINUE
CONTINUE
WRITE(LIST,180) (X(K),K=NUVAR,NCOL)
WRITE(LIST,190) (YCALC(K),K=1,NODEP)
WRITE(LIST,200) (YDIIFR(K),K=1,NODEP)
IF(PNZH) PUNCH 5250,J,(YDIIFR(K),YCALC(K),K=1,NODEP)
WRITE(LIST,210) (Z(K),K=1,NODEP)
IF(BYPASS) GO TO 410
C
DO 250 K=1,NODEP
DO 230 I=1,PLUS1
IF(Z(K)=CELLBD(I)) 220,220,230
220 OBS(I-1,K)=OBS(I-1,K)+1.0
GO TO 250
230 CONTINUE
250 CONTINUE
C
410 KOUNT = KOUNT +1
IF(KOUNT.LT.10) GO TO 430
WRITE(LIST,270) IDENT
KOUNT=0
430 CONTINUE
C
IF(.NOT.SAVRES) GO TO 439
ITC=NWHERE+NODEP
DO 435 IRC=1,NODEP
IC=(ITC+IRC-1)*NOOB+1
DO 434 K=1,NWHERE
IS=(K-1)*NOOB+1
ICHAR(5)=IRC
ICHAR(9)=K
CALL LRNCVT(ICHAR(5),1,ICHAR(5),1,6,0)
CALL LRNCVT(ICHAR(9),1,ICHAR(9),1,6,0)
CALL LRTEG(ICHAR,54)
CALL LRPLT(RESPLT(IS),RESPLT(IC),NOOB)
434 CONTINUE
IS=(NWHERE+IRC-1)*NOOB+1
ICHAR(9)=IRC
MCHAR(5)=IRC
CALL LRNCVT(MCHAR(5),1,MCHAR(5),1,6,0)
CALL LRTEG(MCHAR,54)
CALL LRPLT(RESPLT(IS),RESPLT(IC),NOOB)
435 CONTINUE
439 CONTINUE
C
IF(BYPASS) RETURN
DO 650 K=1,NODEP
SKEW(K)=SKEW(K)**2/(FLOAT(NOOB)**2*ERRMS(K)**3)
SKUR(K)=SKUR(K)/(FLOAT(NOOB)*ERRMS(K)**2)
CHISQ=0.0
DO 640 I=1,CELLS
RCT(I)=OBS(I,K)
CHISQ=CHISQ+RCT(I)**2
640 CONTINUE
CHISQ=FCCELLS*CHISQ/FLOAT(NOOB)-FLOAT(NOOB)
WRITE(LIST,280) NDEGCH,CHISQ,SKEW(K),SKUR(K)
CALL -LIST(K,RCT,CELLS)
650 CONTINUE
RETURN
135 FORMAT(15H FOR EACH DEPENDENT TERM AND OBSERVATION IS PRINTED
X /31H OBSERVED RESPONSE (Y OBSERVED)
X /29H CALCULATED RESPONSE (Y CALC)
CRSPLT PROGRAM

CRSPLT accepts a subset of the data used for a NEWRAP problem. It can be used as a preregression analysis program to help formulate model equations to be analyzed with NEWRAP or it may be used as a postregression program by using punched output from NEWRAP to obtain more complex residual plots than direct use of NEWRAP allows.

When used as a preregression program, it will compute an \( X'X \) and \( C \) matrix including all the terms (independent and dependent) if requested. It can also compute eigenvalues and eigenvectors of the submatrix of \( X'X \) corresponding to the independent variables.

When used as a postregression program, the punched output of residuals and predicted values from NEWRAP can be plotted against new functions of the independent variables.

The input is much the same as for NEWRAP. The seven sets of input are as follows:

1. **IDENTIFICATION** (I, IDENT)(I2, 13A6): IDENT is Hollerith data used to identify the problem. I indicates the number of additional cards to be read for identification (columns 1 to 78).
2. **PROBLEM SIZE** (NOVAR, NODEP, NOTERM, NOOB)(314, I5)
   - NOVAR: Number of input independent variables
   - NODEP: Number of input dependent variables
   - NOTERM: Number of terms in model equation
   - NOOB: Number of observations
3. **TRANSFORMATIONS**: This input is the same as the transformations of NEWRAP except for the upper limit of 150 transformations.
4. **FORMAT** (INPUT, FMT)(I2, 13A6): INPUT specifies the unit number the input data is stored on and FMT indicates the format for reading it.
(5) PLOTTING REQUESTS

NOPLTS (I4)
(MPLT, IYPLT)

One card supplies NOPLTS, the number of plots desired. The following cards supply pairs of integers indicating which pairs of terms to plot. The format is 4012 (i.e., 20 plots per card). The first number of the pair (IXPLT) specifies the sequence number of the term to be used as the abscissa. The second number (IYPLT) specifies the sequence number of the term to be used as the ordinate. As an example, the following sequence of transformations and subsequent plotting requests would cause $X_1^2$ to be plotted against $X_1$ as well as against $X_3$:

```
MODEL SIZE 0003
TERMS 616263
TRANSFORMATIONS 0100061 01026162 01026263
CONSTANTS blank card
NOPLTS 0002
(MPLT, IYPLT) 0102 0302
```

(6) MATRIX REQUESTS (XTXC, EIGENC)(2L1): If XTXC is F, no matrix calculations are executed. If it is T, then an $X'X, X'X$ deviation and a correlation matrix of all the NOTERM + NODEP terms appearing on the TERMS card are computed. If EIGENC is T, then the eigenvalues and eigenvectors of the sub-matrix corresponding to the independent terms (the first NOTERM terms) are calculated.

(7) DATA: Same usage as in NEWRAP.

An illustrative set of input is given, followed by the corresponding sample of output and a main program listing. The subprograms TRIANG, RECT, and EIGEN are required and are the same as in NEWRAP.

15 SAMPLE CRSFLT PROBLEM
DATA IS FROM DRAPER AND SMITH
APPLIED REGRESSION ANALYSIS (REFERENCE 4 OF NEWRAP REPORT)
CHAPTER 7

INITIAL MODEL EQ WAS
$Y = \text{CHAMBER PRESSURE}$
$X_1 = \text{TEMPERATURE OF CYCLE}$
$X_2 = \text{VIBRATION LEVEL}$
$X_3 = \text{DROP(Shock)}$
$X_4 = \text{STATIC FIRE}$
$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + \text{ERR}$
THE FOLLOWING TERMS ARE BEING CREATED FOR RESIDUAL PLOTS

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>X2</td>
<td>X3</td>
<td>X4</td>
<td>X1*X2</td>
<td>X1*X3</td>
<td>X1*X4</td>
<td>X2*X3</td>
<td>X2*X4</td>
<td>X3*X4</td>
<td>Y(PRED)</td>
<td>UNIT NO.</td>
</tr>
</tbody>
</table>

RESIDUALS

UNIT NO. 1  -75  0  -65  1.4
  1 -0.36550470E+01  C  5550470E+01 175  C  150  26.3
  2 -0.76331992E+00  C  27633199E+02 7  C  0  -65  150  29.4
  3 0.23366811E+01  C  27633199E+02 8  C  0  165  -65  9.7
  4 0.46449530E+01  C  50550470E+01 0  C  0  0  150  32.9
  5 0.58366811E+01  C  27633199E+02 0  -75  -75  0  150  26.4
  6 0.50644449E+00  C  25893355E+02 11  175  175  C  0  65  8.4
  7 0.61503899E+00  C  77849610E+01 14  -75  -75  -65  150  28.4
  8 0.35676645E+01  C  25893355E+02 15  175  175  165  -65  11.5
  9 0.37150390E+01  C  77849610E+01 18  0  65  -65  -65  1.3
 10 -0.37550470E+01  C  5550470E+01 19  0  165  150  21.4
  11 -0.56633189E+01  C  27633199E+02 20  0  -75  -65  -65  0.4
  12 -0.34850838E+01  C  38850838E+01 21  0  175  165  150  22.9
  13 -0.63932328E+01  C  29793232E+02 24  0  0  -65  3.7
  14 -0.1355470E+01  C  5550470E+01 3  0  -75  0  150  26.5
  15 0.6664439E+00  C  25893355E+02 5  0  -75  0  150  23.4
  16 -0.24933555E+01  C  25893355E+02 16  0  -75  0  150  26.5
  17 0.60644439E+00  C  25893355E+02 17  0  175  0  -65  5.8
  18 -0.1984961E+01  C  7784961E+01 18  0  175  0  -65  7.4
  19 -0.3845610E+00  C  7784961E+01 19  0  175  0  -65  5.8
  20 -0.1355470E+01  C  25893355E+02 12  0  -75  -65  -150  28.8
  21 0.29066643E+01  C  25893355E+02 22  0  -75  -65  -150  26.4
  22 0.5664439E+00  C  25893355E+02 13  0  175  -65  -65  11.8
  23 0.40150389E+01  C  7784961E+01 23  0  175  165  -65  11.4
$IRFTC CARSPLX

COMMON/AL/ X(99),C(99),SUMX(70),SUMXX(2485),A(72,7C)
X,XI,DATA(1200C)
COMMON/AL2/XMEAN(70),XSTD(70),SUMX2(7C),NTRANS,NCON(3C0),
X,NTERM(71),NTRAN (15C),NXCUD(15C)
DIMENSION IDENT(13),FMT(13),FMTTR(14),CORR(1),FMTSGL(3)
DATA(FMTTR(1),I=1,13)/6H(1H 16.6H,8G15.2H6) /
DATA(FMTSGL(1),I=1,3)/6H(1H 16.6H,8G15.2H6) /
LOGICAL XTC,EIGENC
EQUIVALENCE (CORR,A)
DATA ICHAR(216) VS /
COMMON IXPLT(4CC),IXPLT(4CC)
DIMENSION ICHAR(3)

C**************************************************************
1C READ(5,110) IDENT
   WRITE(6,111) IDENT
   DO 100 J=1,19863
      X(J)=0.0
   100 CONTINUE
11C IF(I)120,120,115
115 READ(5,300) FMT
   WRITE(6,301) FMT
   I=1-1
   GO TO 113
12C READ(5,112) NVAR,NUDEP,NTERM,NJOB
   WRITE(6,105) NVAR,NUDEP,NTERM,NJOB
   MVTERM=MVTERM+NUDEP
   IF(L.GT.1200C) GO TO 100C
   READ (5,282) NTRANS,KONNU
   IF(NTRANS)255,255,22C
   READ(5,239)(TERM(K),K=1,MTERM)
   WRITE(6,235)(TERM(K),K=1,MTERM)
   READ(5,230)(NXCUD(I),NTRAN(I),NCON(2*I-1),NCON(2*I),I=1,NTRANS)
   WRITE(6,240)(NXCUD(I),NTRAN(I),NCON(2*I-1),NCON(2*I),I=1,NTRANS)
   IF(NCONNO)255,255,25C
25C READ(5,260) (CON(I),I=1,KONNU)
   WRITE(6,262)(CGN(I),I=1,KCONNO)
   READ(5,3000) IUIF,FMT
   WRITE(6,3001) IUIF,FMT
   READ(5,282) NPLTS
   IF(NPLTS.LE.0) GO TO 32C
   IF(NPLTS.GT.33C) GO TO 22CC
   READ(5,230)(IXPLT(I),IXPLT(I),I=1,NPLTS)
   WRITE(6,5000)(IXPLT(I),IXPLT(I),I=1,NPLTS)
   32C CONTINUE
   READ(5,6000) XTC,EIGENC
   NORD=NOVAR+NUDEP
C
C**************************************************************

KNOB = 1.0
/FLOAT(KNOB)
DO 690 J=1,KNOB
   READ(IUIF,FMT) (X(I),I=1,NORD)
   IF(NTRANS)45C,45,340
690 CALL TRANS
   DO 430 K=1,MTERM
      I= NTERM(K)
      X(K)=CGN(I)
430 CONTINUE

79
CONTINUE
IF (.NOT. TXTC) GO TO 620
IR=0
DO 610 I=1,MTERM
SUMX(1) = SUMX(1) + X(I)
DO 600 I=1,I
IR=IR+1
SUMXX(IR) = SUMXX(IR) + X(I)*X(I)
660 CONTINUE
610 CONTINUE
620 CONTINUE
IA = J-NOOB
DO 650 I=1,MTERM
IX = IA + I*NOOB
XDATA(IX) = X(I)
650 CONTINUE
IA = J-NOOB
DO 690 I=1,MTERM
IX = IA + I*NOOB
XDATA(IX) = X(I)
690 CONTINUE
IF (.NOT. TXTC) GO TO 720
WRITE(6,1010)
CALL RECT(NOOB,MTERM,NOOB,MTERM,XDATA,FMTSGL)
WRITE(6,1020)
CALL TRIANG(SUMXX,MTERM,8,FMMTRI)
DO 710 I=1,MTERM
710 SUMXX(IR) = SUMXX(IR) - SUMX(IR)*SUMX(IR)*RNOOB
CORR(IR) = SUMXX(IR)/ SQRT(SUMXX(IR)*SUMXX(IR))
IR=IR+1
720 DO 718 I=1,AOTERM
718 SUMXX(I) = SUMXX(I)*RNOOB
I=1
DO 710 J=1,MTERM
710 K=1,J
SUMX(K) = SUMX(K) + XDATA(I)
DO 718 I=1,AOTERM
718 CORR(K) = CORR(K) / SQRT(SUMXX(I)*SUMXX(K))
WRITF(6,1050)
CALL EIGEN(SUMXX,AOTERM)
WRITF(6,1060)
CALL LREL(1,1,AOTERM,FMMTRI)
720 DO 900 K=1,NOPLTS
900 CALL LRLEGN(IDENT,54,0,1,5,0,5,0)
CALL LRLEGN(IDENT,1C,24,0,1,4,5,1,0)
DO 900 K=1,NOPLTS
ICHAR(1) = IXPAT(K)
ICHAR(3) = IYPAT(K)
IS1 = (IXPAT(K)-1)*NOOB+1
IS2 = (IYPAT(K)-1)*NOOB+1
CALL LRCNMV(ICHAR(1),1,ICHAR(1),1,6,0)
CALL LRCNMV(ICHAR(3),1,ICHAR(3),1,6,0)
CALL LRLEG(ICHAR,18)
CALL LRPLOT(XDATA(IS1),XDATA(IS2),NOOB)
900 CONTINUE
GO TO 10
IF (OR WRITE(6,45)L
  STOP
 55 FORMAT(62H THE REQUIRED NUMBER OF LOCATIONS EXCEEDS THE 12000 AVAIL
XLABLE IF))
STOP
70 IF WRITE(6,2005) NOPLOTS
STOP
70 5 FORMAT(25H MAX NO. OF PLOTS IS 30018)
110 FORMAT(12,13A6)
111 FORMAT(1H1,13A6)
112 FORMAT(314,15)
400 FORMAT(2L1)
500 FORMAT(1HK/((15(1X,12,1X,12,2X))))
300 FORMAT(13A6.A2)
301 FORMAT(1H 13A6,.A2)
305 FORMAT(1H 316)
282 FORMAT 2(14)
230 FORMAT(4012)
235 FORMAT(11H TERMS ARE / (1H 3014))
24 FORMAT(25H THE TRANSFORMATIONS ARE / (1H 5(4I4,5X)))
26 fmt (5E15.7)
262 FORMAT(19H THE CONSTANTS ARE / ((1H 8G15.7))
161 FORMAT(16H1 THE DATA MATRIX )
102 FORMAT(26H2 THE X TRANSPOSE X MATRIX )
1075 FORMAT(32H X TRANSPOSE X DEVIATIONS MATRIX )
1030 FORMAT(3H2 THE CORRELATION MATRIX )
1046 FORMAT(46H2 FOLLOWING ARE EIGENVALUES OF X TRANS X MATRIX)
1056 FORMAT(1H 8G15.7)
1066 FORMAT(53H K F IGENVECTORS BY COLUMNS IN SAME ORDER AS EIGENVALUES)
3060 FORMAT(12,13A6)
3061 FORMAT(1H 1E,2X,13A6)
END

**SUBROUTINE TRANS**

C******************************************************************************
C COMMON/H1/ X(99),CUN(99),SUMX(7C),SUMXX(2485),A(7C,70)
X,DATA(12000)
COMMON/H2/XMEAN(70),XSTD(7C),SUMX2(7C),NTRANS,NCON(300),
X TERM(70),NTRAN (15C),NXCOD(150)
C******************************************************************************
C THIS SUBROUTINE PERFORMS TRANSFORMATIONS IF THIS OPTION IS
C REQUESTED.
C******************************************************************************

K TRANSFORMATION SET NUMBER.
C NCON(7*K-1) CONSTANT NUMBER TO USE.
C NCON7*K) DERIVED CONSTANT.
C NTRANS (K) NUMBER OF TRANSFORMATION REQUESTED.
C NXCOD(K) VARIABLE NUMBER.
AC 30 K=1,TRANS
 I=NCON(2*K-1)
 IF(I)100,100,110
23.
24.
25.
26.
27.
28.
29.
30.
31.
32.
33.
34.
35.
36.
37.
38.
39.
40.
41.
42.
43.
44.
45.
46.
47.
48.
49.
50.
51.
52.
53.
54.
55.
56.
57.
58.
59.
60.
61.
62.
63.
64.
65.
66.
67.
68.
69.
70.
71.
72.
73.
74.
75.
76.
77.
78.
79.
80.
81.
82.
83.
84.
SAMPLE CRSPLOT PROBLEM
DATA IS FROM DRAPER AND SMITH
APPLIED REGRESSION ANALYSIS (REFERENCE 4 OF NEWRAP REPORT)
CHAPTER 7
INITIAL MODEL EU 155
Y = CHAMBER PRESSURE
X1 = TEMPERATURE OF CYCLE
X2 = VIBRATION LEVEL
X3 = DROPPED SHOCK
X4 = STATIC FIRE
Y = X0 + AX1 + AX2 + BX3 + CX4 + ERR
THE FOLLOWING TERMS ARE BEING CREATED FOR RESIDUAL PLOTS
1 X1 2 X2 3 X3 4 X4
5 X1X2 6 X1X3 7 X1X4 8 X2X3
9 X2X4 10 X3X4 11 X2X4
12 YIPRED
UNIT NO.

THF DATA MATRIX
1 -75.60000
2 175.0000
3 35.0000
4 165.0000
5 150.0000
6 -65.0000
7 175.0000
8 -35.0000
9 175.0000
10 65.0000
11 165.0000
12 0
13 175.0000
14 0
15 -75.0000
16 0
17 175.0000
18 0
19 -75.0000
20 0
21 175.0000
22 0
23 175.0000
24 0

1 4875.0000
2 2450.0000
3 0
4 -1125.0000
5 0
6 -1125.0000
7 -1125.0000
8 -1125.0000

1 5650.47
2 27.6332
3 0
4 5.65047
5 27.4532
6 25.8335
7 1.6949
8 0

1 5.65047
2 27.6332
3 0
4 5.65047
5 27.4532
6 25.8335
7 1.6949
8 0
### The X Transpose Matrix

<table>
<thead>
<tr>
<th></th>
<th>104750.0</th>
<th>2000.00</th>
<th>33750.00</th>
<th>-11375.00</th>
<th>-11375.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-11375.00</td>
<td>7.749461</td>
<td>15.00000</td>
<td>3.715380</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-11375.00</td>
<td>5.259047</td>
<td>18.00000</td>
<td>-3.715380</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-11375.00</td>
<td>27.03322</td>
<td>19.00000</td>
<td>-5.643319</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-11375.00</td>
<td>3.885604</td>
<td>23.00000</td>
<td>-3.487214</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-11375.00</td>
<td>29.759235</td>
<td>21.00000</td>
<td>-6.809323</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>-11375.00</td>
<td>5.759047</td>
<td>24.000000</td>
<td>-1.355477</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-11375.00</td>
<td>25.893356</td>
<td>3.000000</td>
<td>6.666666</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>-11375.00</td>
<td>25.893356</td>
<td>5.000000</td>
<td>-2.493356</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-11375.00</td>
<td>25.893356</td>
<td>16.00000</td>
<td>0.666666</td>
<td></td>
</tr>
</tbody>
</table>

### The Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>1.000000</th>
<th>0.314372</th>
<th>0.314372</th>
<th>0.314372</th>
<th>0.314372</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000000</td>
<td>1.0000000</td>
<td>0.19592800</td>
<td>-0.355580</td>
<td>0.19592800</td>
<td></td>
</tr>
<tr>
<td>0.314372</td>
<td>0.19592800</td>
<td>1.0000000</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td></td>
</tr>
<tr>
<td>0.314372</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td>1.0000000</td>
<td>0.19592800</td>
<td></td>
</tr>
<tr>
<td>0.314372</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td>1.0000000</td>
<td></td>
</tr>
<tr>
<td>0.314372</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td>0.19592800</td>
<td>1.0000000</td>
</tr>
</tbody>
</table>
In optimum-seeking experimentation involving many independent variables, quadratic response surfaces are often used. A development of the most important aspect of the design and analysis of response surface experiments can be found in Davies (ref. 9) and Box and Hunter (ref. 10). A discussion of the interpretation of a quadratic surface fitted to a large experiment is given in reference 11.

The general form of a quadratic surface is given by

\[ y = b_0 + b'X + X'BX + \epsilon \]

\[ = b_0 + (b_1, b_2, \ldots, b_p) \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \]

\[ + (x_1, \ldots, x_p) \begin{pmatrix} b_{11} & \frac{1}{2}b_{12} & \cdots & \frac{1}{2}b_{1p} \\ \frac{1}{2}b_{12} & b_{22} & \cdots & \frac{1}{2}b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & b_{pp} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} + \epsilon \]  

(26)

The analysis of such an equation is simplified by two calculations: (1) the calculation of the stationary point of the surface and (2) a transformation of axes to new independent variables which changes the \( B \) matrix to a diagonal matrix.

The stationary point of the surface is the solution \( X_s \) to

\[ \frac{\partial y}{\partial x_i} = 0 \quad i = 1, p \]

The transformation of the axes is given by the computation of the orthogonal matrix \( P \) which reduces \( B \) to a diagonal matrix; that is,
The \( \lambda_i \) are the eigenvalues of \( B \). The new variables are given by

\[
Z = P'X
\]

where \( P \) is the matrix whose columns are the eigenvectors of \( B \). If two successive transformations of the variables are made as follows:

\[
W = X - X_s
\]

\[
Z = P'W
\]

then equation (26) becomes

\[
y - y_s = \lambda_1 Z_1^2 + \lambda_2 Z_2^2 + \ldots + \lambda_p Z_p^2
\]

(27)

where

\[
y_s = b_0 + b'X_s + X_s'BX_s
\]

(28)

From examination of equation (27), some general conclusions can be drawn concerning the attainment of a maximized response. For example, consider just two of the possible results.

(1) Suppose all the \( \lambda_i \leq 0 \) and \( X_s \) is near or in the region of \( X \) at which the experiments were performed. Then clearly any deviation of \( Z \) from \( Z = 0 \) will decrease the response. Thus \( Z = 0 \) (or equivalently \( X = X_s \)) is a maximum and is the combination of independent variables the experimenter seeks.

(2) Some \( \lambda_i < 0 \) and some \( \lambda_i > 0 \), and \( X_s \) is close to the region of experimentation. Thus \( X_s \) represents what is sometimes called a saddlepoint. Moving in some directions will cause a decrease in \( y \) and moving in other directions will cause an increase in \( y \). Thus the experimenter could move from \( X_s \) in the direction that corresponds to the direction of the \( Z \) which has the largest positive coefficient in equation (27). This will increase \( y \) most rapidly from the value of \( y_s \).
The INPUT (Sample input is shown in fig. 8) is as follows:

1. Identification: One card, all 80 columns.
2. The number of independent variations JFAC. (14) (JFAC must be less than or equal to 15)
3. Variable Format. One card, all 80 columns.
4. \( b_0 \) (according to FORMAT in item (3)).
5. \( b_1, \ldots, b_{JFAC} \) (according to format in item (3)).
6. \( b_{1,1} \)

\[
\begin{array}{cccc}
  b_{1,2} & b_{2,2} \\
  b_{1,3} & b_{2,3} & b_{3,3} \\
  \vdots & \vdots & \vdots & \vdots \\
  b_{1,JFAC} & b_{2,JFAC} & \ldots & b_{JFAC,JFAC}
\end{array}
\]
One READ statement for each line, according to the format in item (3). As an example, consider the following estimated equation:

\[
y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_{11} x_1^2 + b_{12} x_1 x_2 + b_{22} x_2^2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{33} x_3^2 + b_{14} x_1 x_4 + b_{24} x_2 x_4 + b_{34} x_3 x_4 + b_{44} x_4^2 + b_{15} x_1 x_5 + b_{25} x_2 x_5 + b_{35} x_3 x_5 + b_{45} x_4 x_5 + b_{55} x_5^2
\]

\[
y = 147.2686 - 8.989120 x_1 - 6.817975 x_2 - 14.60964 x_3 + 9.248688 x_4 + 14.19698 x_5 + 1.805520 x_1^2 - 2.126719 x_1 x_2 - 1.475730 x_2^2 - 7.314218 x_1 x_3 - 16.10219 x_2 x_3 + 2.274270 x_3^2 + 1.310780 x_1 x_4 - 2.477188 x_2 x_4 - 1.289688 x_3 x_4 + 2.236770 x_4^2 + 2.664464 x_1 x_5 + 1.805520 x_2 x_5 + 2.389934 x_3 x_5 + 6.014934 x_4 x_5 - 12.58198 x_5^2
\]

A canonical reduction of this is given in the sample output. The listing of the main program is supplied. The subprograms TRIANG, DGELG, EIGEN, and RECT are required. TRIANG and RECT are as in NEWRAP. DGELG and EIGEN are the double precision general linear equation and eigenvalue routines from reference 12.

```
$IBFTC CRFDJC
DIMENSION BL(15), BS(105), BSAVE(15), BSAVE(105), FMT(14), XIN(225)
DOUBLE PRECISION BL, BS, BSAVE, BSAVE, XIN, ZERO, YS
DIMENSION SBL(L5), SBS(225), FMTTRI(14)
DATA(FMTTRI(II, I=1, 4), 6H15H RO, 6HW 15, 2, 6HX, (8G1, 6H5, 6)) /

999 READ (5, 1001) FMT
WRITE(5, 1002) FMT
READ(5, 1003) JFAC
WRITE(6, 1004) JFAC
```
REAL(5.1001) FMT
READ(5,FMT) BZERO
WRITE(6,1005) BZERO
READ(5,FMT) (BL(I),I=1,JFAC)
IE=0
DO 5 I=1,JFAC
IE=IE+1
IS=IE-I+1
READ(5,FMT) (BS(K),K=IS,IE)
DO 4 <= IS,IE
BSAVE(K)=BS(K)
4 SBS(K)= SNGL(BS(K))
BLSAVE(I)=BL(I)
SBL(I)= SNGL(BL(I))
BL(I)= -BL(I)
5 CONTINUE
LENGTH = JFAC*(JFAC+1)/2
WRITE(5,1006)
WRITE(5,1007) (SBL(I),I=1,JFAC)
WRITE(5,1008)
CALL TRIANG(SBS,JFAC,B,FMTTRI)
IJ=1
XIN(I)= BS(I)*2.000
DO 50 I=2,JFAC
II= I-1
IIK=I
IIJ= JFAC*II
DO 40 J=1,II
IJ= IJ+1
BSAVE(IJ)=0.5000*BS(IJ)
XIN(IIK)=BS(IJ)
IIK=IIK+JFAC
IIJ= IIJ+1
XIN(IIJ)= BS(IJ)
40 CONTINUE
IIJ=IIJ+1
IJ= IJ+1
XIN(IIJ)=2.000*BS(IJ)
50 CONTINUE
EPS=1.0E-10
CALL JDEIG(BL,XIN,JFAC,EPS,IER)
IF(IER.NE.0) WRITE(6,1009) IER
WRITE(6,1010)
WRITE(6,1007) (BL(I),I=1,JFAC)
IJ=0
YS= BZERO
DO 153 I=1,JFAC
YS=YS+BL(I)*BSAVE(I)
DO 140 J=1,II
IJ= IJ+1
140 YS= YS+ BL(I)*BL(J)*BS(IIJ)
150 CONTINUE
WRITE(5,1011) YS
C
IJ=0
DO 163 L=1,LENGTH
IJ=IJ+1
160 SBS(IJ)= SNGL(BSAVE(IJ))
CALL EIGEN(BSAVE,XIN,JFAC,0)
IJ=0
DO 203 I=1,JFAC
IJ=IJ+1
200 BL(I)= BSAVE(IJ)
WRITE(5,1012)
WRITE(5,1007) (BL(I),I=1,JFAC)
WRITE(5,1013)
JJ=JFAC*JFAC
DO 210 I=1,JJ

SAMPLE CANONICAL REDUCTION PROBLEM

THE COEFFS ARE BZERO          147.268600

\begin{align*}
\text{LINEAR} & \\
\text{SECOND ORDER} & \\
\text{ROW 1} & 1.805520 & \\
\text{ROW 2} & -2.126719 & -1.475730 & \\
\text{ROW 3} & -7.314218 & -16.10219 & 2.274270 & \\
\text{ROW 4} & 1.307800 & -2.477188 & -1.289688 & 2.236770 & \\
\text{ROW 5} & 2.664464 & 1.827434 & 2.389934 & 6.014934 & -12.58198 & \\
\end{align*}

THE STATIONARY POINT IS       3.05811526     -1.96462563     0.12068067     -3.89057544     -0.93711068E-01

THE VALUE OF YS IS            120.589275

\begin{align*}
\text{EIGENVALUES} & \\
\text{EIGENVECTORS} & \\
\text{ROW 1} & -0.300108 & 0.519711 & 0.750815 & 0.234709 & -0.138316 & \\
\text{ROW 2} & -0.549544 & -0.275150 & -0.293631 & 0.690293 & -0.244071 & \\
\text{ROW 3} & 0.779700 & 0.809764E-02 & 0.804860E-01 & 0.577909 & -0.270383 & \\
\text{ROW 4} & -0.219489E-02 & 0.790049 & -0.582318 & -0.22827E-02 & -0.191618 & \\
\text{ROW 5} & 0.105615E-02 & 0.173062 & -0.669725E-01 & 0.364045 & 0.912707 & \\
\end{align*}

Lewis Research Center,  
National Aeronautics and Space Administration,  
Cleveland, Ohio, January 25, 1972,  
132-80.
APPENDIX - BORROWED ROUTINES

Some of the routines used in the programs were taken from the literature. Both INVXTX and TRIANG are by Webb and Galley (ref. 13), and EIGEN and HIST are from the IBM programmer's manual (ref. 12), as are DGELG and the double precision version of EIGEN.

Listing of INVXTX and TRIANG are given here, as follows:

```
SUBROUTINE TRIANG(A,B,NN,NKOL,FORMAT,II)
DIMENSION FORMAT(II)
DIMENSION A(1), B(1)
DOUBLE PRECISION B

1 FORMAT (1H1)
2 FORMAT (1H/1H/1H)
3 FORMAT (1H/1H/1H)
COMMON/SMALL/DUM(15),LIST
N = NN
NCOL = NKOL
KLUMPS = N/NCOL
C

KEEPT = 0
K1 = 1
K2 = NCOL - 1
K3 = NCOL
IF (KLUMPS .EQ. 0) GO TO 120
C
DO 90 KLUMP=1,KLUMPS
ITR1 = KEEPT
I = -1
ILO = (KLUMP-1)*NCOL + ITR1 + 1
DO 30 K=K1,K2
I = I + 1
ITR1 = ITR1 + K - 1
ILO = ILO + K - 1
IHI = ILO + 1
GO TO (26,28),II
26 WRITE(LIST,FORMAT) K,(A(J),J=ILO,IHI)
GO TO 30
28 WRITE(LIST,FORMAT) K,(B(J),J=ILO,IHI)
30 CONTINUE
KEEPT = ITR1 + K2
DO 60 K=K3,N
ITR1 = ITR1 + K - 1
ILO = ILO + K - 1
IHI = ILO + NCOL - 1
GO TO (56,58),II
56 WRITE(LIST,FORMAT) K,(A(J),J=ILO,IHI)
GO TO 60
58 WRITE(LIST,FORMAT) K,(B(J),J=ILO,IHI)
60 CONTINUE
K1 = K1 + NCOL
K2 = K2 + NCOL
K3 = K3 + NCOL
90 WRITE(LIST,3)
```
SUBROUTINE INVXX(A, NN, D, FACT)
C    ASSUMES THE MATRIX A IS SYMMETRIC AND POSITIVE DEFINITE, AND ONLY
C    THE UPPER TRIANGLE IS STORED AS A ONE-DIMENSIONAL ARRAY IN THE
C    ORDER A(1,1), A(1,2), A(2,2), A(1,3), A(2,3), A(3,3), ..., A(N,N).
C    N IS THE ORDER N OF THE INPUT MATRIX A.
C    D IS (ON EXIT) THE DETERMINANT OF A, DIVided BY FACTOR**NN.
C
DIMENSION A(1)
DOUBLE PRECISION A, PV, F
N = NN
ITR1 = 0
DO 145 K=1,N
C
ITR1 = ITR1+K-1
KP1 = K+1
KM1 = K-1
KK = ITR1+K
PV = 1.0D0/A(KK)
C
ITR2 = 0
IF (K-1) 150,80,50
C
DO 60 J=1,KM1
ITR2 = ITR2+J-1
KJ = ITR1+J
F = A(KJ)*PV
DO 60 I=1,J
IJ = ITR2+I
IK = ITR1 + I
60 A(IJ) = A(IJ) + A(IK)*F
C
IF (K-N) 70,120,150
C
REduce REST OF TRIANGLE, RIGHT OF PIVOTAL COLUMN
70 ITR2 = ITR1
80  DO 110  J=KP1,N
   ITR3 = ITR1
   ITR2 = ITR2+J-1
   KJ = ITR2+K
   F = A(KJ)*PV
   DO 100  I=1,J
   IF  (I-K) 90,100,95
90  IJ = ITR2+I
   IK = ITR1 + I
   A(IJ) = A(IJ) - A(IK)*F
   GO TO 100
95  IJ = ITR2 + I
   ITR3 = ITR3 + I - 1
   IK = ITR3 + K
   A(IJ) = A(IJ) - A(IK)*F
100 CONTINUE
110 CONTINUE

C   DIVIDE PIVOTAL ROW-COLUMN BY PIVOT, INCLUDING APPROPRIATE SIGNS
120  ITR2 = ITR1
   DO 140  I=1,N
   IF  (I-K) 125,130,135
125  IK = ITR1+I
   A(IK) = -A(IK)*PV
   GO TO 140
130  A(KK) = PV
   GO TO 140
135  ITR2 = ITR2+I-1
   KI = ITR2+K
   A(KI) = A(KI)*PV
140 CONTINUE

C   (REPLACE PIVOT BY RECIPROCAL)
145 CONTINUE
C   RETURN
150 RETURN
END
REFERENCES


"The aeronautical and space activities of the United States shall be conducted so as to contribute... to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

— National Aeronautics and Space Act of 1958

**NASA Scientific and Technical Publications**

**Technical Reports:** Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

**Technical Notes:** Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

**Technical Memoranda:** Information receiving limited distribution because of preliminary data, security classification, or other reasons.

**Contractor Reports:** Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

**Technical Translations:** Information published in a foreign language considered to merit NASA distribution in English.

**Special Publications:** Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

**Technology Utilization Publications:** Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

**Scientific and Technical Information Office**

**National Aeronautics and Space Administration**

Washington, D.C. 20546