Analyzing a Regression Model with a General Positive Definite Covariance Matrix with The SAS System

Gerardo L. Hurtado, SAS Institute Inc., Cary, NC
April 10, 1990

1 Abstract
This article discusses and proposes a procedure for the analysis of the univariate linear regression model with known general positive definite covariance matrix with SAS/STAT software of the SAS System. Estimation of parameters, hypothesis testing, estimation under constraints and collinearity and influence diagnostics are reviewed. An example is given to illustrate the procedure.

KEY WORDS: Linear Regression Model, Collinearity Diagnostics, Influence Diagnostics, Positive Definite Covariance Matrix.

2 The Model
Consider the model equation
\[ y = X\beta + \epsilon \] (1)
where \( X \) is a \( n \times p \) known matrix of rank \( p \leq n \), \( \beta \) a vector of unknown parameters, and \( \epsilon \) a random vector with \( E(\epsilon) = 0 \) and \( \text{pd} \) variance-covariance matrix \( V = \sigma^2H \) for some \( H \) \( \text{pd} \) and \( \sigma^2 > 0 \). In this article we assume that \( H \) is known and \( \sigma^2 \) is an unknown parameter. The common assumption of the regression model is to consider \( H = I \), but in this article we relax this assumption. Many of the standard results of this article can be found in books such as Searle (1971), Judge et al. (1985) or Theil (1971).

The analysis of the model (1) is usually carried out performing a transformation of the data to produce an error vector with covariance matrix which is a scalar multiple of the identity matrix. We have at least two ways to do this with SAS/STAT software.

One way is by premultiplying (1) by \( H^{-1/2} \)
\[ z = Q\beta + w, \] (3)
where \( \Gamma \) is an orthogonal matrix containing the eigenvectors of \( H \) and \( A \) is a diagonal matrix containing the eigenvalues. \( A^{-1/2} \) is obtained by taking the \(-1/2\) power of every diagonal element of \( A \). We obtain:
\[ s = Q\beta + w, \] (4)
where \( s = H^{-1/2}y \), \( Q = A^{-1/2}X \) and \( w = H^{-1/2}\epsilon \). Note that multiplicities of the eigenvalues do not change the transformation to obtain (3).

A second way to transform the data is by premultiplying (1) by \( A^{-1/2} \). That is, we define \( s = A^{-1/2}y \), \( Q = A^{-1/2}X \) and \( w = A^{-1/2}\epsilon \).

Since the second transformation is simpler, and since multiplicities of the eigenvalues of \( H \) only causes the rows of \( s \), \( Q \) and \( w \) to be switched around for these rows corresponding to the multiplicities, the results using SAS/STAT software will be presented in terms of this redefinition of \( s \), \( Q \) and \( w \). The non-unique eigenvectors corresponding to the multiplicities do not change the results of this article.

Note that the transformations do not change the parameter space.

3 Estimation of Parameters and Hypothesis Testing
Since \( \text{var}(w) = \sigma^2I \), most of the results of the standard regression model hold under (3), in particular the estimation of \( \beta \) can be accomplished using the least squares criterion applied to (3), which leads to the Generalised Least Squares Estimator (GLSE) with solution
\[ \tilde{\beta} = (Q'Q)^{-1}Q's = (X'H^{-1}X)^{-1}X'H^{-1}y, \] (5)
and variance
\[ \text{var}(\tilde{\beta}) = \sigma^2(Q'Q)^{-1} = \sigma^2(X'H^{-1}X)^{-1}. \] (6)
\( \tilde{\beta} \) is in fact the BLUE (Best Linear Unbiased Estimator) of \( \beta \).

Also
\[ \hat{\sigma}^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n - p}, \] (7)
which is the biased estimator of \( \sigma^2 \). Under the additional assumption that \( y \) is normally distributed, it can be shown that the MLE (Maximum Likelihood Estimators) are (4) and (7) and \( \hat{\sigma}^2 \) can be obtained with SAS/STAT by producing the spectral decomposition of \( H \) (eigenvalues and eigenvectors decomposition) using the PRINCOMP procedure and applying the REG procedure to (3) with the NOINT option, after transforming the original data.

The REG procedure also computes the \( R^2 \) for the model as
\[ R^2 = \frac{s'Q(Q'Q)^{-1}Q's}{s'Q}, \] (8)
which is a number between 0 and 1, that can be interpreted as the proportion of the weighted variation of \( y \) explained by the model. Note that since
\[ F = \frac{n - p}{p} \frac{R^2}{1 - R^2}, \] (9)
the \( R^2 \) is monotonically related to the \( F \) statistic used to test the hypothesis that all the regression coefficients are zero.

The predicted values are computed as \( X\hat{\beta} \), and the residuals as \( y - X\hat{\beta} \), which have covariance matrices
\[ \sigma^2X(Q'Q)^{-1}X' \], (10)
and
\[ \sigma^2[H - X(X'Q'Q)^{-1}X'], \] (11)
respectively. Under the assumption that the disturbance component of future values are uncorrelated with the \( n \) disturbances of the sample, both with the same covariance matrix, the predictor error variance is given by
\[ \sigma^2[H + X(X'Q'Q)^{-1}X']. \] (12)
If the above assumption does not hold, that is, if for future observations \( y \), we have that
\[
y = X_0 \beta + \epsilon, \tag{13}
\]
with
\[
\text{var} \left( \begin{array}{c} \epsilon \\ \epsilon_s \\
\end{array} \right) = \sigma^2 \left( \begin{array}{cc} I & W' \\ W & H \\
\end{array} \right), \tag{14}
\]
then the Best Linear Unbiased Predictor (BLUP) of \( y \), is
\[
X_0 \hat{\beta} + W' H^{-1} (y - X_0 \hat{\beta}), \tag{15}
\]
with covariance matrix of the predictor error
\[
\text{var}(y_s, y - X \hat{\beta}) = \sigma^2 \left[ H - W' H^{-1} W + (X_s - W' H^{-1} X) (Q' Q)^{-1} (X_s - W' H^{-1} X)' \right]. \tag{16}
\]

Estimates of \( (10), (11) \) and \( (12) \) can not be obtained from the output of REG applied to (3) or (1). The example shows the algorithm used to compute these values in a DATA step.

To test the hypothesis \( K' \beta = m \), where \( K \) is a known \( p \times s \) full rank matrix, \( s \leq p \), \( m \) is a known vector, and assuming that \( y \) is normally distributed, we observe that the quadratic form
\[
(\hat{\beta} - \beta)' K (K' Q Q^{-1} K^{-1} K' (\hat{\beta} - \beta)) / \sigma_1^2
\]
is distributed as a \( \chi^2(s) \). Therefore, if \( K' \beta = m \) is true the statistic
\[
(\hat{\beta} - m)' K (K' Q Q^{-1} K^{-1} (K' \hat{\beta} - m)) / \sigma_1^2
\]
is distributed as an \( F(s, n - p) \) and can be used as a basis for hypothesis testing and interval estimation. REG with the NOINT option applied to (3) compute this statistic for the entire model by default, that is when \( K = I \) and \( m = 0 \), and using the TEST statement we can test other hypothesis (see the example).

4 Estimation under Constraints

Under linear consistent constraints of the form \( K' \beta = m \), where \( K \) is a known \( p \times s \) full rank matrix, \( \beta \) is estimated in model (1) using the same transformation that produced model (3). This is so because the transformation from model (1) to (3) does not carry a transformation of the parameter space. Therefore, the BLUE \( \beta \) under \( K' \beta = m \) is
\[
\hat{\beta} = \hat{\beta} - (Q' Q)^{-1} K (Q' Q)^{-1} K' (\hat{\beta} - m), \tag{19}
\]
with covariance matrix given by
\[
\text{cov}(\hat{\beta}) = A (Q' Q)^{-1} A', \tag{20}
\]
where \( A = I - (Q' Q)^{-1} K (Q' Q)^{-1} K' \), and \( \sigma_1^2 \) is an unbiased estimator of \( \sigma^2_1 \) under the given restrictions, that is
\[
\hat{\sigma}_1^2 = \frac{(n - p) \hat{\sigma}_r^2}{n - p + s},
\]
\[
(\hat{K}' - m)' (K' Q Q^{-1} K^{-1} (K' \hat{\beta} - m)) / n - p + s, \tag{21}
\]
where \( s \) is the rank of \( K \).

The Lagrange multiplier is given by
\[
(K' Q Q^{-1} K^{-1} (K' \hat{\beta} - m), \tag{22}
\]
with covariance matrix
\[
BB^t \hat{\sigma}_1^2, \tag{23}
\]
where \( B = [K' (Q' Q)^{-1} K]^{-1} K' (Q' Q)^{-1} Q' \).

Values of \( (16), (20), (21), (22), \) and \( (23) \) are produced by REG with the NOINT option when the procedure is applied to (3) using the RESTRICT statement.

5 Collinearity Diagnostics

In this section we study the collinearity diagnostics produced by the REG procedure as applied to our model. We follow Belsley, Kuh, and Welsch (1980) closely. Our interest lies on the extent that collinearity among the columns of \( Q \) results in a matrix \( Q' Q \) whose ill conditioning causes statistics such as the GLSE \( \beta \) and the covariance matrix \( \sigma^2 \) to be numerically unstable. Therefore the diagnostics are expressed in terms of the transformed data \( Q \) instead of the original \( X \).

5.1 The Variance Inflation Factor (VIF)

The diagonal elements of the inverse correlation matrix of the parameters \( \beta \) are called the VIF. Their diagnostic value is explained by the fact that for the \( i^{th} \) component of \( \beta \)
\[
\text{var}(\hat{\beta}_{(i)}) = \frac{\sigma_1^2 VIF_i}{q_i q_i}, \tag{24}
\]
and the relation
\[
VIF_i = \frac{1}{1 - R_i^2}, \tag{25}
\]
where \( R_i \) is the multiple correlation coefficient of the column \( q_i \) of \( Q \) regressed on its remaining columns.

The option VIF of REG applied to (3) with the NOINT option produces \( [25] \), and the option TOL produces \( 1 - R_i^2 \).

5.2 The Condition Indexes

The Singular Value Decomposition (SVD) of the matrix \( Q \) yields
\[
Q = S D T', \tag{26}
\]
where \( S \) is a \( n \times p \) matrix, \( D = \{d_i\} \) is a \( p \times p \) diagonal matrix containing the singular values and \( T = \{t_{ij}\} \) is a \( p \times p \) matrix such that \( S' S = T'T = T = I \).

The condition indexes are defined as the ratio
\[
\eta_i = \frac{d_{max}}{d_i}, \quad i = 1, \ldots, p, \tag{27}
\]
If \( Q \) has mutually orthogonal columns, the ideal data, then the condition indexes would be all equal to one, the smallest possible value. Therefore it has been suggested that the columns of \( Q \) be scaled to have all length equal to one before the SVD is performed. This scaling is accomplished by REG with the NOINT option, and the condition indexes can be obtained with the option COLLIN when the procedure is applied to (3).

5.3 Variance Decomposition Proportions

Another important result taken from the SVD is the production of the Variance Decomposition Proportions. Note that we can express the covariance matrix of \( \hat{\beta} \) as
\[
\text{cov}(\hat{\beta}) = \sigma_1^2 (Q' Q)^{-1} = \sigma_1^2 T D^{-2} T', \tag{28}
\]
and for the \( i^{th} \) component of \( \beta \)
\[
\text{var}(\hat{\beta}_{(i)}) = \sigma_1^2 T_{i, i} D_{i, i}^{-2}, \tag{29}
\]
therefore (29) decomposes \( \hat{\beta}_{(i)} \) into components, each one associated with each singular value. REG with NOINT applied to (3) produces (29) with the COLLIN option when \( Q \) has been scaled.
6 Influence Diagnostics

In this section we study some of the diagnostic statistics produced by REG with the INFLUENCE option as applied to the model (1).

The influence diagnostics produced with the option INFLUENCE under model (1) or (3) do not have too much significance. First of all, since REG does not accept a general covariance matrix for the observations the way that we have defined it after (1), we cannot use the INFLUENCE option applied to (1). Secondly, this option applied to (3) will produce diagnostics that apply to the transformed data, and consequently they are not of much interest.

The main difficulty lies in the fact that the diagnostics are produced by deletion of observations. Case deletion has no significance if it is done with REG applied to (2) because deletion of rows of X and Q does not achieve the corresponding deletion of rows of y and X and corresponding rows and columns of H in the original model equation (1). Rather than using INFLUENCE in REG applied to (3), we want to produce case deletion diagnostics applied to (1) by selecting the rows and columns to delete from this model. This can be done with some programming statements in a DATA step, as we show in the example.

In the discussion that follows we study some of the regression diagnostics that can be applied to the model that we are considering. Recall from Hurtado and Gerig (1989) that the estimated value of \( \beta \) say \( \hat{\beta}_1 \), computed after the \( p \)th observation has been deleted, and assuming that the rank of the 'new' X has not changed and the new covariance matrix of the observations is still pd, is given by

\[
\hat{\beta}_1 = \hat{\beta} - (X'H^{-1}X)^{-1}X'H^{-1}(y - X\hat{\beta})
\]

where \( h_{11} \) is the \( p \)th diagonal element of \( H^{-1} \), \( m_1 = \hat{\xi}'M\hat{\xi} \) is the \( p \)th column of the identity matrix, and \( M = H^{-1}X(X'H^{-1}X)^{-1}X'H^{-1} \). The estimated value of \( \sigma^2 \) computed after the removal of the \( p \)th observation, say \( \sigma_i^2 \), is given by

\[
(n - p - 1)\sigma_i^2 = (n - p)\sigma_i^2 - \left(\frac{\sigma_i^2}{h_{11}^2} - \sigma_i^2\right).
\]

6.1 The Cook’s D Statistic

Assuming \( H = I \) and under normality of y, Cook (1977) proposed a statistic based on the confidence ellipsoid generated by all \( \beta^* \), such that

\[
\left(\beta^* - \hat{\beta}\right)'X'X(\beta^* - \hat{\beta}) \leq F(1 - \alpha, p, n - p).
\]

If we replace \( \beta^* \) with \( \hat{\beta}_1 \), we obtain

\[
D_r(X'X) = \frac{(\hat{\beta}_1 - \hat{\beta})'X'X(\hat{\beta}_1 - \hat{\beta})}{\sigma_i^2},
\]

a statistic that Cook suggested to use and compare for significance with \( F(1 - \alpha, p, n - p) \). \( D_r(X'X) \) however is not distributed as \( F \). The concept of this ellipsoid centered at \( \hat{\beta} \) with contours determined by the eigenvalues and eigenvectors of \( X'X \) can be extended to cover the case that we are studying. The new formulation is given by

\[
D_r(X'H^{-1}X) = \frac{(\hat{\beta}_1 - \hat{\beta})'X'H^{-1}X(\hat{\beta}_1 - \hat{\beta})}{\sigma_i^2},
\]

which is motivated by the fact that

\[
\frac{(\beta^* - \hat{\beta})'X'H^{-1}X(\beta^* - \hat{\beta})}{\sigma_i^2} \sim F(p, n - p).
\]

6.2 The ‘Hat’ Matrix

When \( H \) is the identity matrix, Hoaglin and Welsch (1978) define the Hat Matrix as the matrix that projects onto the column space of \( X \) that is \( X(X'X)^{-1}X' \). As we can see from (30) and (31), this matrix is of fundamental importance in the determination of influential observations because for \( H = I \) we have that \( h_{11} - m_1 = 1 - \hat{\xi}'X(X'X)^{-1}X'\hat{\xi} \), so that observations with values of \( m_1 \) close to 1 can be considered influential in the estimation of \( \beta \) and \( \sigma^2 \). But when \( H \) is any pd matrix, from (30) and (31) we see that we need to consider the diagonal elements of \( PH = H^{-1} - M \). Observations with \( h_{12} - m_1 \approx 0 \) are highly influential and should be studied. In particular since \( PH \) is pd (positive semidefinite) we have that \( h_{12} - m_1 \geq 0 \). Also, it is possible to show that \( h_{12} - m_1 \leq h_{11} \). Note that when \( H = I \), \( PH \) represents the orthogonal projection matrix onto the error space in (1), but when \( H \) is any pd matrix this property is lost, making the analysis of the magnitude of the diagonal elements of \( PH \) more complicated. For example, the diagonal elements of \( M \) are no longer bounded above by 1, but by the diagonal elements of \( H^{-1} \). A test to 0 trivially for every observation when \( n = p \) (and rank(\( X \)) = \( p \)) and also when a column of \( X \) is identical to some other column except at the \( p \)th observation or a column of \( X \) is a column of the identity matrix so that one regression parameter is spent in only one data point.

The distributional assumptions for \( X \) that lead Beale, Kuh, and Welch (1980) to determine \( 2p \sigma_i \) as a cutoff point for the Hat matrix are of no use here.

6.3 The COVRATIO Statistic

The COVRATIO statistic measures the change in the determinant of the covariance matrix of the regression estimates by dropping the \( p \)th observation. Assuming without loss of generality that the \( p \)th observation being deleted is the last one in the data set, we have

\[
COVRATIO = \frac{\det[\sigma_i^2(X'H^{-1}X)_{-p}]^{1/2}}{\det[\sigma_i^2(X'H^{-1}X)^{-1}]^{1/2}}
\]

\[
= \frac{\sigma_i^2}{\sigma_i^2}\frac{\det[[X'H^{-1}X]_{-p}]}{\det[[X'H^{-1}X]^{-1}]^{1/2}}
\]

\[
= \frac{1}{\frac{1}{\sigma_i^2} + \frac{c_p}{h_{11}^2}}
\]

where \( X = [X_1, X_2, X_3] \) being a \( n - 1 \times p \) submatrix of \( X \), \( H_{11} \) is the upper left submatrix of \( H \) and

\[
c_p = \left(\frac{\xi'H^{-1}X_{-p}(y - X\hat{\beta})}{h_{12}^2 - m_1}\right).
\]

A Value of COVRATIO close to 1 indicates that the observation being deleted is not influential in the estimation of the covariance matrix of the regression parameters.

7 Example

In the following example, we have a data set with 10 observations, 3 independent variables, \( X_1, X_2 \) and \( X_3 \) and a dependent variable \( Y \).

\%let nobs=10; \%let p=3; \%let nobs2 =\%eval(nobs+2); \%let nobs3 =\%eval(nobs+3); \%let nobs4 =\%eval(nobs+4);
data a;
  input xy1-xybobslppl;
  title1 'Data set to analyse';
  run;

data cov(type=cov);
  title1 'Matrix B of the Cov';
  title2 'matrix of the obs';
  run;

*--obtain the Eigenvalues and the Eigenvectors of H--;*
  proc princomp coy outstat=outstat;
  title1 'Spectral decomposition of H';
  var obs1-obs&Uobs;
  run;

*proc print data=outstat;
  title1 'outstat= from PRINCOMP';
  run;*---prepare the data
  *--outstat2 contains
  Lambda^{(-1/2)} Gamma'--;*
  proc print data=outstat2;
  title1 'Data to enter to REG';
  run;

*proc print data=a;
  title1 'Data set to enter to REG';
  run;*----------run PROC REG----------;*
  proc reg data=ab outest=outest;
covout;
title1 'PROC REG output';
**estimation of parameters and
  collinearity diagnostics--;
unrestr:model abkpp1~ab1-abip/noint
covb corrb xpx itol vif collin;
**--estimation of parameters

***estimation under constraints--;
***-----------(example)------------;
restr: model abkpp1~ab1-abip/noint
covb;
restrict ab1=ab2;
run;
proc print data=outest;
title1 'OUTEST= ds from REG';
run;
proc print data=out:
title1 'OUT= ds from REG';
run;
**usual the influence
diagnostics--;
data infl;set
cutest (in=one)
ab (in=two)
cut (in=three)
keeps (in=four)
cov (in=five)
outest2(in=three);
keep resid y resid y stdy resid y stud y std resid y stud y std resid y
array cov{tp,tp} _temporary_;
array alp{&p} _temporary_; array ab{*} _temporary_; array axb{knobs ,Ilppi} _temporary_; _temporary_; _temporary_; _temporary_; _temporary_; _temporary_; array x{*} _temporary_; array llX{lmobs .tp} _temporary_; array vy{bobs} _temporary_; array pred{bobs} _temporary_; array resy{tnobs} _temporary_; array g{bobs ,tnobs} _temporary_; array diagg{bobs} _temporary_; array stdp{bobs} _temporary_; array stdr{bobs} _temporary_; array stdi{btobs} _temporary_; array student{tnobs} _temporary_; array a{knobs .tnobs} _temporary_; array b{1lp ,bobs} _temporary_; array c{tnobs .Ilp} _temporary_; array m{tnobs .mobs} _temporary_; array diag{knobs} _temporary_; array diag{knots} _temporary_; array diag{knobs} _temporary_; array resid{knobs} _temporary_; array hires{knobs} _temporary_; array siz{knobs} _temporary_; array rstdres{knobs} _temporary_; array covratic{knobs} _temporary_; array a.al{knobs .tp} _temporary_; array xix{tp,tp} _temporary_; array prec{knobs .tp} _temporary_; array cookd{knobs} _temporary_; retain s2;
if one then do;
  if _MODEL_='UNRESTR' & _TYPE_='PARMS'
  then do s=1 to &p;
alp{s}=ab{s};
  end;
  if _MODEL_='UNRESTR' & _TYPE_='COV'
  then i+i;
  if i=1 then s2=RMSE*RMSE.
do j=1 to &p;--load cov(alp)/s2
          in a matrix--;
cov{i,j}=ab{j}/s2;
  end;
  end;
if two then do;
l=1;
do j=1 to &ppi;
  ab1(j)=ab{j};--load the data transformed in a matrix--;
  if j=ppi then vy{i}=y;--load
          original data into a matrix--;
  end;
if four then do;
k=1;
  resid{k}=res:*--load the z
          residuals into a vector--;
if five then do;
r=1;
  diag=r=obs:*--load the z
          residuals into a vector--;
if three then do;
  *--compute predicted &
          residuals--;
do i=1 to &nobs;
pred(i)=0;
do j=1 to &p;
  pred(i)=pred(i)+mx{i,j}*alp{j};
  end;
  resid(i)=vy(i)-pred(i);
  *alp{j};
  end;
end;
*--form diag(G)=
  diag(inv(\'inv(H)I)\'X');--
*--form inv(\'inv(H)X\')X--;
do i=1 to &p;
do j=1 to &nobs;
a(i,j)=0;
do l=1 to &p;
a(i,j)=a(i,j)+cov{i,l}*mx{j,l};
  end;
end;
*--form diag(G)=
do i=1 to &nobs;
diag(i)=0;
**--form STDP, STDR, STD1 & STUDENT--**

do i=1 to knobs;
  stdp{i}=sqrt(s2*diagg{i});
  stdr{i}=sqrt(s2*(diagh{i}-diagg{i}));
  stdi{i}=sqrt(s2*(diagh{i}+diagg{i}));
  student{i}=resy{i}/stdr{i};
end;

**--form M = inv(X'inv(H)X)X'inv(H)--;**

do i=1 to knobs;
  do j=1 to knobs;
    a{i,j}=0;
    do l=1 to kp;
      a{i,j}=a{i,j}+eov{i,l}.*axb{j,l};
    end;
  end;
end;

**--form C = inv(H)X--;**

do i=1 to &:nobs;
  do j=1 to &:p;
    c{i,j}=0;
    do l=1 to &:nobs;
      c{i,j}=c{i,j}+g{l,i}.*axb{l,j};
    end;
  end;
end;

**--form diag(inv(H)); & diag(Ph)--**

do i=1 to &:nobs;
  diag(i)=0;
  do l=1 to &:nobs;
    diag(i)=diag(i)+g{l,i}.*axb{l,i};
end;

end;
if cookd(&nobs) = . then do i=1 to &nobs;
   resid=resy{i};
   rstd=rstudent{i};
   dgph=diagph{i};
   dgbi=diaghi{i};
   dg =diag{i};
   covrat=covratio{i};
   predicty=pred{i};
   stdpred=stdp{i};
   stdres=stdr{i};
   studnt=student{i};
   cook=cookd{i};
   stdrpred=stdi{i};
   output;
end;
run;

proc print data=infl (rename=(residy=res rstd=rstudent
dgph=diagph dghi=diaghi
dgm=diagm covrat=eovratio
predicty=pred stdpred=stdp
stdres=stdr student=student
cook=cookd stdrpred=stdi);
title1 'Influence statistics';
var res rstudent diagph diaghi
dgm covrat pred stdp stdr
student cookd stdi;
run;

8 Acknowledgments
I wish to thank Professor Thomas Gerig (North Carolina State
University), Warren Sarle and Scott Sweetland (both at SAS Insti-
tute Inc.) for their many helpful comments during earlier versions
of this article.

9 List of References
Belady, D. A., Kuh, E. and Welsch, R. E. (1980), Regression Di-
gnostics: Identifying Influential Data and Sources of Collinearity,
New York: John Wiley and Sons, Inc.

Cook, R. D. (1977), "Detection of Influential Observation in Linear
Regression", Technometrics, 19, 15–18.

Hoaglin D. C. and Welsch, R. E. (1978), "The Hat Matrix in Re-
gression and ANOVA", The American Statistician, 32, pp 17–22
and Corrigenda 32, p. 146.

Linear Models with Applications to Variance Components", SAS
Users Group International Conference Proceedings: SUGI 14 ,

Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., Lee,
York: John Wiley and Sons, Inc.

Searle, S. R. (1971), Linear Models, New York: John Wiley and
Sons, Inc.

Theil, H. (1971) Principles of Econometrics, New York: John Wi-
ley and Sons, Inc.