CONSTRAINED LINEAR MODELS

by

Thomas M. Gerig and A. R. Gallant

Institute of Statistics
Mimeo Series No. 814

Raleigh 1972

ABSTRACT

A presentation of the theory of linear models subject to equality constraints on the parameters is set forth. No rank conditions on the matrices appearing in the model are required and reparameterization of the model is unnecessary in order to use the methods developed. The bias which may arise due to false restrictions or deficient rank of the input matrix is derived and convenient methods for detecting the presence of this bias in applications are given. Attention is given to accurate and efficient computing procedures and an example is provided to illustrate the application of these methods to data.

1. INTRODUCTION

It is very common in economic investigations to assume that a linear model gives an adequate representation of the data. Often, however, the investigator knows from the underlying theory that certain restrictions exist among the parameters. It is the aim of this paper to set out the important practical and theoretical aspects of constrained linear models theory using mathematical forms which are computationally convenient.

We have specified our model so as not to put any limitations on the dimensions or ranks of the matrices involved nor on the relationship of the constraints to the input matrix. Also, we have adopted an approach for our analysis which does not rely on a reparametrization of the model. We feel that these features are of practical importance because they allow the investigator to specify his model in exactly the form he wants and allows him to keep in touch with his original input variables throughout the analysis.

Attention shall be given to determining the effect of incorrectly specifying the restrictions and to matters of computational efficiency. Almost all of the theoretical results used but not proved in this paper can be found in Theil [3].

Section 2 contains notation and some matrix results which will be used in the paper. Section 3 spells out the basic properties of the estimators under the assumption that the model is correctly specified. Section 4 gives two decompositions of an arbitrary linear function of the parameters and gives a discussion of conditions under which bias may occur and how to eliminate it. Section 5 discusses the singular value decomposition of a general matrix, indicates how this can be used to obtain the Moore-Penrose inverse of a matrix. These results provide the tools necessary for implementing the methods of earlier sections. Section 6 gives an example which illustrates the use of the model and how the formulas should be applied to data.

The Model. Suppose that

$$y = X\beta + e \quad \text{and} \quad R\beta = r ,$$

where y is an $(n \times 1)$ vector of observations, X is an $(n \times p)$ matrix of fixed input variables, β is a $(p \times 1)$ vector of unknown parameters, e is an $(n \times 1)$ vector of uncorrelated random variables each with mean zero and variance σ^2, R is a $(q \times p)$ matrix of constants and r is a $(q \times 1)$ vector of constants. The equations $R\beta = r$ shall be referred to as the constraints and we shall assume that they are a consistent set of equations.

In Section 2 we shall demonstrate that a more general model can be handled by the methods we present. In fact, it will be shown that, if $\mathrm{Var}(e) = \Sigma\,\sigma^2$, where Σ is a known, positive semi-definite (possibly singular) matrix and σ^2 is unknown, the model can be reduced to the form we give

2. REDUCTION TO STANDARD FORM

In this section, we show how a model with $\mathrm{Var}(e) = \sigma^2\,\Sigma$ can be reduced to the form given in the preceeding section and set forth our notation and certain matrix relations to be used in the sequel.

The model appears in first form as

$$y^* = X^*\beta + e^* \quad \text{and} \quad R^*\beta = r^*$$

where y^*: $(n^* \times 1)$, X^*: $(n^* \times p)$, β: $(p \times 1)$, e^*: $(n^* \times 1)$, R^*: $(q^* \times p)$ and r^*: $(q^* \times 1)$. We make no restriction on the order or rank of X^* and R^* but we do require $R^*\beta = r^*$ to be a consistent set of equations. We shall take $\mathcal{E}(e^*) = 0$ and $\text{Var}(e^*) = \sigma^2 \Sigma$, where Σ is known and positive semi-definite. The case when Σ is full rank is handled in the usual way by finding a non-singular matrix T: $(n^* \times n^*)$ such that $T \Sigma T' = I_{n^*}$. Then the model may be transformed to

$$y = X\beta + e \quad \text{and} \quad R\beta = r ,$$

where $y = Ty^*$: $(n \times 1)$, $X = TX^*$: $(n \times p)$, and $e = Te^*$: $(n \times 1)$. Then $n = n^*$, $q = q^*$ and $\text{Var}(e) = T\,\text{Var}(e^*)T' = \sigma^2\, T \Sigma T' = \sigma^2 I$. Notice that the original parameters β are not altered by this transformation.

If Σ is singular then we can find a non-singular matrix T such that $T \Sigma T' = \begin{pmatrix} I_n & 0 \\ 0 & 0 \end{pmatrix}$, $n < n^*$. We partition $T = \begin{pmatrix} T_{(1)} \\ T_{(2)} \end{pmatrix}$, where $T_{(1)}$: (n, n^*) and $T_{(2)}$: $(n^* - n, n^*)$. Transforming the model as before we obtain the relations

$$T_{(1)}y^* = T_{(1)}X^*\beta + T_{(1)}e^*$$

$$T_{(2)}y^* = T_{(2)}X^*\beta + T_{(2)}e^*$$

$$R^*\beta = r^* .$$

Now $Var(T_{(2)}e^*) = T_{(2)}Var(e^*)\, T'_{(2)} = 0$ so that $T_{(2)}e^* = 0$ and $T_{(2)}X^*\beta = T_{(2)}y^*$ become known linear restrictions on the parameters. Appending these to the previous restrictions we obtain the restrictions $R\beta = r$ with

$$R = \begin{pmatrix} T_{(2)}X^* \\ R^* \end{pmatrix} : \ (q \times p) \quad \text{and} \quad r = \begin{pmatrix} T_{(2)}y^* \\ r^* \end{pmatrix} : \ (q \times 1) ,$$

where $q = q^* + n^* - n$. We will assume, additionally, that the new set of restrictions $R\beta = r$ are consistent. The remainder of the model is obtained by setting $y = T_{(1)}y^*$: $(n \times 1)$, $X = T_{(1)}X^*$: $(n \times p)$ and $e = T_{(1)}e^*$: $(n \times 1)$. Then $Var(e) = T_{(1)}Var(e^*)T_{(1)} = \sigma^2 I_n$ as required.

We have seen that whether or not Σ is of full rank there is a non-singular transformation matrix T which may be used to reduce the model to standard form. We will consider in Section 5 how T may be obtained in practice.

We will make extensive use of the Moore-Penrose inverse in what is to follow. We define it here and defer the discussion of computation to Section 5.

Definition. (Theil [3], pp. 269-274). Let A be an $(m \times n)$ matrix. Then there exists a matrix A^+ of order $(n \times m)$ which satisfies $A A^+ A = A$, $A^+A A^+ = A^+$, $(A A^+)' = A A^+$, and $(A^+A)' = A^+A$. The matrix A^+ is unique and is called the Moore-Penrose (pseudo) inverse of A .

The following matrix notation shall be used in what is to follow. If A is an arbitrary $(m \times n)$ matrix and a is an $(n \times 1)$ vector then let

$$A' = \text{the transpose of } A ,$$

$$\|a\|^2 = a'a ,$$

$$P_A = A^+A ,$$

$$Q_A = I - A^+A ,$$

$$P_A^* = A\,A^+ , \text{ and}$$

$$Q_A^* = I - A\,A^+ ,$$

whose dimensions are, respectively, $(n \times m)$, (1×1) , $(n \times n)$, $(n \times n)$, $(m \times m)$, and $(m \times m)$. The ranks of the last four matrices are, respectively, rank (A), n-rank (A), rank (A), and m-rank (A) .

Notation which is specific to the model (in standard form) is as follows:

$$W = X(I - R^+R) = XQ_R ,$$

$$V = \begin{pmatrix} X \\ R \end{pmatrix} ,$$

$$\tilde{\beta} = Q_R(XQ_R)^+\{y - XR^+r\} + R^+r ,$$

$$\hat{\beta} = X^+y ,$$

$$SSE(\beta) = (y - X\beta)'(y - X\beta) = \|y - X\beta\|^2 ,$$

$$\tilde{\sigma}^2 = SSE(\tilde{\beta})/(n - \text{rank}(W)) ,$$

$$\hat{\sigma}^2 = SSE(\hat{\beta})/(n - \text{rank}(X)) ,$$

whose dimensions are, respectively, $(n \times p)$, $(n + q \times p)$, $(p \times 1)$, $(p \times 1)$, (1×1) , (1×1) , and (1×1) .

The following matrix relations are easily verified using the four properties of the Moore-Penrose inverse. Much of the verification may be found in Theil ([3], pp. 269-274) .

P_A, Q_A, P_A^*, Q_A^* are symmetric and idempotent,

$$(A')^+ = (A^+)' ,$$

$$A^+(A^+)' = (A'A)^+ ,$$

$$(A'A)^+A' = A'(A\,A')^+ = A^+ ,$$

$R\,R^+r = r$ provided $R\beta = r$ are consistent,

$R(\beta - R^+r) = 0$ provided $R\beta = r$,

$Q_R(\beta - R^+r) = (\beta - R^+r)$ provided $R\beta = r$,

$$P_R R^+ r = R^+ r ,$$

$$Q_R R^+ = 0 ,$$

$$R\ Q_R = 0 ,$$

$$P_R R^+ = R^+ , \text{ and}$$

$$R\ P_R = R .$$

In the remaining sections we will assume that the above relations are known and will use them repeatedly without reference to this section.

3. STATISTICAL PROBERTIES OF $\tilde{\beta}$

The following theorems parallel the standard results in (unconstrained) linear models theory. In each case, our theorem is followed by the corresponding result for an unconstrained linear model stated as a corollary. The proof of each corollary is obtained by setting $q = 1$, $R = 0$ and $r = 0$ then applying the theorem which preceeds it.

The reader who is primarily interested in applications of these results is invited to read the statements of the theorems, skip the proofs, and go on to the next section where he will find what we feel is a more applications oriented interpretation of the properties of the estimator $\tilde{\beta}$.

THEOREM 1:

$$\tilde{\beta} = Q_R(X\ Q_R)^+(y - X\ R^+r) + R^+r$$

minimizes

$$SSE(\beta) = (y - X\beta)'(y - X\beta)$$

subject to the (consistent) constraints

$$R\beta = r \; .$$

PROOF: We will first verify that $R\tilde{\beta} = r$. Now there is a $\bar{\beta}$ such that $R\bar{\beta} = r$ since we assumed a consistent set of constraint equations. Then

$$R\tilde{\beta} = R\, Q_R (X\, Q_R)^+ (y - X\, R^+ r) + R\, R^+ r$$

$$= 0 + R\, R^+ R\, \bar{\beta} = R\, \bar{\beta} = r \; .$$

We now verify that $SSE(\tilde{\beta}) \le SSE(\bar{\beta})$ provided $\bar{\beta}$ satisfies $R\bar{\beta} = r$.

$$SSE(\bar{\beta}) = \| \, y - X\, P_R \bar{\beta} - X\, Q_R \bar{\beta} \, \|^2$$

$$= \| \, y - X\, R^+ r - X\, Q_R \tilde{\beta} + X\, Q_R (\tilde{\beta} - \bar{\beta}) \, \|^2$$

$$= \| \, y - X\, R^+ r - X\, Q_R \tilde{\beta} \, \|^2 + \| \, X\, Q_R (\tilde{\beta} - \bar{\beta}) \, \|^2$$

$$+ 2(y - X\, R^+ r - X\, Q_R \tilde{\beta})'(X\, Q_R)(\tilde{\beta} - \bar{\beta})$$

$$= \| \, y - X\, P_R \tilde{\beta} - X\, Q_R \tilde{\beta} \, \|^2 + \| \, X\, Q_R (\tilde{\beta} - \bar{\beta}) \, \|^2$$

$$+ 2(y - X R^+ r)'[I - (X Q_R)(X Q_R)^+] (X Q_R)(\tilde{\beta} - \bar{\beta})$$

$$= \| y - X \tilde{\beta} \|^2 + \| X Q_R(\tilde{\beta} - \bar{\beta}) \|^2 + 0$$

$$\geq SSE(\tilde{\beta}) . \quad \square$$

COROLLARY:

$$\hat{\beta} = X^+ y$$

is the unconstrained minimum of

$$SSE(\beta) = (y - X\beta)'(y - X\beta) .$$

THEOREM 2: There is a $\bar{\beta}$ of the form

$$\bar{\beta} = Ay + c$$

such that $\mathcal{E}(\lambda'\bar{\beta}) = \lambda'\beta$ for every β satisfying the consistent equations $R\beta = r$ if and only if there are vectors δ and ρ such that

$$\lambda' = \delta' X + \rho' R .$$

PROOF: (If) Let $\lambda' = \delta' X + \rho' R$. As we will see in the next section $\mathcal{E}(\lambda'\tilde{\beta}) = \lambda'\beta$ provided $R\beta = r$. Note that $\tilde{\beta}$ is of the required form.

(Only if) If β is of the form

$$\beta = R^+ r + Q_R \gamma$$

then $R\beta = r$ for all choices of γ . We will take β to be of this form and examine the consequences of various choices of γ under the assumption that there is a

$$\bar{\beta} = Ay + c$$

such that $\mathcal{E}(\lambda'\bar{\beta}) = \lambda'\beta$ for all γ . Under this assumption, for all γ

$$\lambda' A\, X\, R^{+} r + \lambda' A\, X\, Q_R \gamma + \lambda' c = \lambda' R^{+} r + \lambda' Q_R \gamma \; .$$

First set $\gamma = 0$, hence

$$\lambda' A\, X\, R^{+} r + \lambda' c = \lambda' R^{+} r \; ,$$

so that

$$\lambda' A\, X\, Q_R \gamma = \lambda' Q_R \gamma$$

for all choices of γ . By successive choice of the elementary vectors for γ we obtain

$$\lambda' A\, X\, Q_R = \lambda' Q_R$$

whence

$$\lambda' = \lambda' A\, X\, Q_R + \lambda' P_R$$

$$= \lambda' A\, X + \lambda' A\, X\, P_R + \lambda' P_R$$

$$= [\lambda' A] X + [\lambda' A\, X\, R^{+} + \lambda' R^{+}]\, R$$

$$= \delta' X + \rho' R \; . \quad \square$$

COROLLARY: There is a $\bar{\beta}$ of the form

$$\bar{\beta} = Ay + c$$

such that $\mathcal{E}(\lambda'\bar{\beta}) = \lambda'\beta$ for all β if and only if there is a vector δ such that

$$\lambda' = \delta' X .$$

THEOREM 3: Let $\bar{\beta}$ be any estimator of the form $\bar{\beta} = Ay + c$ and λ be of the form $\lambda' = \delta' X + \rho' R$. If $\mathcal{E}(\lambda'\bar{\beta}) = \lambda'\beta$ for all β satisfying the consistent equations $R\beta = r$ then

$$\mathrm{Var}(\lambda'\tilde{\beta}) \leq \mathrm{Var}(\lambda'\bar{\beta}) .$$

PROOF: From the proof of the previous theorem we have $\lambda' A\, X\, Q_R = \lambda' Q_R$. The variance of $\lambda'\tilde{\beta}$ is

$$\mathrm{Var}(\lambda'\tilde{\beta}) = \lambda' Q_R (X\, Q_R)^+ (X\, Q_R)^{+\prime} Q_R \lambda\, \sigma^2$$

$$= \lambda' Q_R (Q_R X' X\, Q_R)^+ Q_R \lambda\, \sigma^2 .$$

The variance of $\lambda'\bar{\beta}$ is

$$\mathrm{Var}(\lambda'\bar{\beta}) = \lambda' A\, A' \lambda\, \sigma^2$$

$$= \lambda' A [P_W^* + Q_W^*]\, A' \lambda\, \sigma^2$$

$$= \lambda' A (X\, Q_R)(X\, Q_R)^+ A' \lambda\, \sigma^2 + \lambda' A Q_W^* A' \lambda\, \sigma^2$$

$$= \lambda' A (X\, Q_R)(Q_R X' X\, Q_R)^+ (X\, Q_R)' A' \lambda\, \sigma^2 + \lambda' A\, Q_W^* A' \lambda\, \sigma^2$$

$$= \lambda' Q_R (Q_R X'X\, Q_R)^+ Q_R \lambda\, \sigma^2 + \lambda' A\, Q_W^* A' \lambda\, \sigma^2$$

$$\geq \mathrm{Var}(\lambda' \tilde{\beta}) \;. \quad \square$$

COROLLARY: Let $\bar{\beta}$ be any estimator of the form $\bar{\beta} = Ay + c$ and λ be of the form $\lambda' = \delta' X$. If $\mathcal{E}(\lambda'\bar{\beta}) = \lambda'\beta$ for all β then

$$\mathrm{Var}(\lambda'\hat{\beta}) \leq \mathrm{Var}(\lambda'\bar{\beta}) \;.$$

THEOREM 4:

$$\mathcal{E}(\mathrm{SSE}(\tilde{\beta})) = [n - \mathrm{rank}(W)]\, \sigma^2 \quad \text{provided} \quad R\beta = r \;.$$

PROOF.

$$\mathrm{SSE}(\tilde{\beta}) = \| \, y - X\, \tilde{\beta} \, \|^2$$

$$= \| \, y - X\, Q_R (X\, Q_R)^+ (y - X\, R^+ r) - X\, R^+ r \, \|^2$$

$$= \| \, Q_W^* e + Q_W^* X(\beta - R^+ r) \, \|^2 \;.$$

Now

$$Q_W^* X(\beta - R^+ r) = Q_W^* W(\beta - R^+ r) + Q_W^* X\, P_R (\beta - R^+ r) = 0$$

since $Q_W^* W = 0$ and $P_R(\beta - R^+ r) = 0$ provided $R\beta = r$. We now have that

$$\mathrm{SSE}(\tilde{\beta}) = e' Q_W^* e \;,$$

where Q_W^* is symmetric and idempotent with rank $Q_W^* = n - \mathrm{rank}\,(W)$. Thus $\mathcal{E}(e' Q_W^* e) = [n - \mathrm{rank}(W)]\, \sigma^2$. $\square$

COROLLARY:

$$\mathcal{E}(SSE(\hat{\beta})) = [n - \text{rank}(X)]\, \sigma^2 .$$

THEOREM 5: Let e be distributed as a multivariate normal $N_n\{0, \sigma^2 I\}$ and let α be chosen between zero and one.

a) If λ is of the form $\lambda' = \delta' X + \rho' R$ then

$$P[\lambda'\tilde{\beta} - \epsilon \le \lambda'\beta \le \lambda'\tilde{\beta} + \epsilon] \ge 1 - \alpha ,$$

where

$$\epsilon^2 = (\lambda' Q_R (W'W)^+ Q_R \lambda)\, \tilde{\sigma}^2\, F\{\alpha; 1, n - \text{rank}(W)\}$$

provided $R\beta = r$.

b) If Λ is a matrix of the form $\Lambda' = \Delta' X + \zeta' R$ then

$$P[S \ge F\{\alpha; \text{rank}(\Lambda' Q_R), n - \text{rank}(W)\} \mid \beta = \beta^\circ] \le \alpha ,$$

where

$$S = \frac{(\Lambda'\tilde{\beta} - \Lambda'\beta^\circ)' (\Lambda' Q_R (W'W)^+ Q_R \Lambda)^+ (\Lambda'\tilde{\beta} - \Lambda'\beta^\circ)[\text{rank}(\Lambda' Q_R)]^+}{\tilde{\sigma}^2}$$

provided $R\beta^\circ = r$.

$F\{\alpha; f_1, f_2\}$ denotes the α level percentage point of an F random variable with f_1 degrees freedom for the numerator and f_2 for the denominator.

PROOF: Part (a) follows from Part (b) when Λ' is a $(1 \times p)$ row vector. To prove Part (b) we write

$$\Lambda' = \Delta' X + \quad ' R = \Delta' X\, Q_R + (\Delta' X\, R^+ + \quad ')R = \Delta' W + \Gamma' R$$

so that S may be written

$$S = \frac{(\Delta' W\tilde{\beta} - \Delta' W\beta^\circ)'(\Delta' W(W'W)^+ W'\Delta)^+(\Delta' W\tilde{\beta} - \Delta' W\beta^\circ)/f_1}{SSE(\tilde{\beta})/f_2}$$

$$= \frac{N/f_1}{D/f_2} ,$$

where $f_1 = \text{rank}(\Lambda Q_R)$ and $f_2 = n - \text{rank}(W)$ (set $1/f_1 = 0$ if $\text{rank}(\Lambda Q_R) = 0$) .

If $\Delta' W = 0$ then $S = 0$ and Part (b) follows trivially. We will therefore consider the case when $\Delta' W \neq 0$. From the proof of Theorem 4, $D/\sigma^2 = e' Q_W^* e/\sigma^2$. Since Q_W^* is idempotent with rank f_2 we have by Theil ([3], p. 83) that D/σ^2 is distributed as a χ^2 random variable with f_2 degrees freedom.

If β° is the true value of β and $R\beta^\circ = r$ we may write

$$\Delta' W(\tilde{\beta} - \beta^\circ) = \Delta' W[W^+\{e + X(\beta^\circ - R^+ r)\} + R^+ r - \beta^\circ]$$

$$= \Delta' W\, W^+ e + \Delta' W[W^+ X(\beta^\circ - R^+ r) - (\beta^\circ - R^+ r)]$$

$$= \Delta' W\, W^+ e + \Delta' W[W^+ W(\beta^\circ - R^+ r) - (\beta^\circ - R^+ r)]$$

$$= \Delta' W\, W^+ e + \Delta'[W(\beta^\circ - R^+ r) - W(\beta^\circ - R^+ r)]$$

$$= \Delta' P_W^* e .$$

Then N becomes

$$N = e'P_W^*\Delta(\Delta'W(W'W)^+W'\Delta)^+\Delta'P_W^*e$$

$$= e'P_W^*\Delta(\Delta'P_W^*\Delta)^+\Delta'P_W^*e$$

$$= e'(P_W^*\Delta)(P_W^*\Delta)^+e \ .$$

In general $\text{rank}(AB) \le \text{rank}(A)$ so that in particular $\text{rank}(\Delta'W) = \text{rank}(\Delta'P_W^*W) \le \text{rank}(\Delta'P_W^*) \le \text{rank}(\Delta'W)$ whence $f_1 = \text{rank}(\Lambda'Q_R) = \text{rank}(\Delta'W) = \text{rank}(\Delta'P_W^*)$. Again citing Theil ([3], p.83), N/σ^2 is distributed as a χ^2 with f_1 degrees freedom. Since $Q_W^*(P_W^*\Delta)(P_W^*\Delta)^+ = 0$ we have by Theil ([3], p. 84) that N and D are independent.

It follows that if $\Delta'W \neq 0$ and $R\beta^\circ = r$ then $S = \dfrac{N/\sigma^2 f_1}{D/\sigma^2 f_2}$ is distributed as an F with f_1 numerator degrees of freedom and f_2 for the denominator. ▯

COROLLARY: Let e be distributed as a multivariate normal $N_n\{0, \sigma^2 I\}$ and let α be chosen between zero and one.

a) If λ is of the form $\lambda' = \delta'X$ then

$$P[\lambda'\hat{\beta} - \epsilon \le \lambda'\beta \le \lambda'\hat{\beta} + \epsilon] \ge 1 - \alpha \ ,$$

where

$$\epsilon^2 = (\lambda'(X'X)^+\lambda)\ \hat{\sigma}^2\ F\{\alpha;\ 1,\ n - \text{rank}(X)\} \ .$$

b) If Λ is a matrix of the form $\Lambda' = \Delta'X$ then

$$P[S \ge F\{\alpha;\ \text{rank}(\Lambda),\ n - \text{rank}(X)\} \mid \beta = \beta^\circ] \le \alpha \ ,$$

where

$$S = \frac{(\Lambda'\hat{\beta} - \Lambda'\beta^{\circ})'(\Lambda'(X'X)^{+}\Lambda)^{+}(\Lambda'\hat{\beta} - \Lambda'\beta^{\circ})/\text{rank}(\Lambda)}{\hat{\sigma}^2} .$$

4. SOURCES OF BIAS

In the preceeding section we saw that $\lambda'\tilde{\beta}$ is unbiased for $\lambda'\beta$ provided $\lambda' = \delta'X + \rho'R$ and $R\beta = r$. In this section, we will examine the bias which results when either $\lambda' \neq \delta'X + \rho'R$ or $R\beta \neq r$ or both. As a result of this examination, we will be able to characterize those $(\lambda'\beta)$'s which are estimated unbiasedly by $(\lambda'\tilde{\beta})$ even when $R\beta \neq r$ and determine what additional information is necessary to allow unbiased estimation of $\lambda'\beta$ when the condition $\lambda' = \delta'X + \rho'R$ is not satisfied. We have deferred proofs of the less obvious claims made in this section to the Appendix in order to focus attention on the main points of the discussion.

Recalling the notation and relations given in Section 2 we can write

$$\tilde{\beta} = Q_R(X\,Q_R)^{+}(y - X\,R^{+}r) + R^{+}r$$

$$= \beta - (I - Q_R(X\,Q_R)^{+}X)P_R(\beta - R^{+}r) - Q_V(\beta - R^{+}r) + Q_R(X\,Q_R)^{+}e$$

$$= \beta + B_1(\beta) + B_2(\beta) + Q_R(X\,Q_R)^{+}e .$$

Consider the estimation of an arbitrary linear function of the parameters, $\lambda'\beta$. It is clear from the decomposition of $\tilde{\beta}$ that $\mathcal{E}(\lambda'\tilde{\beta}) = \lambda'\beta + \lambda'B_1(\beta) + \lambda'B_2(\beta)$. We will consider the conditions on λ and β which eliminate the two sources of bias $\lambda' B_1(\beta)$ and $\lambda' B_2(\beta)$.

The first source of bias is due to specification error since $B_1(\beta) = 0$ for all β satisfying $R\beta = r$. The second source is due to the deficient rank of V since $Q_V = 0$ if $\text{rank}(V) = p$.

There do exist linear functions of the parameters, $\lambda'\beta$, for which $\lambda' B_1(\beta) = \lambda' B_2(\beta) = 0$ for arbitrary choice of β . These are the parametric functions which are estimated unbiasedly by $\lambda'\tilde{\beta}$ whether or not the restrictions $R\beta = r$ are correctly specified. Consider λ of the form $\lambda' = \delta' Q_R (X\ Q_R)^+ X$. It is not difficult to verify that for such λ

$$\mathcal{E}(\lambda'\tilde{\beta}) = \mathcal{E}(\lambda'\hat{\beta}) = \lambda'\beta ,$$

$$\lambda\ \tilde{\beta} = \lambda\ \hat{\beta} ,$$

and

$$\text{Var}(\lambda'\tilde{\beta}) = \text{Var}(\lambda'\hat{\beta}) = \lambda'(X'X)^+ \lambda\ \sigma^2 .$$

An easy test for λ of this form is to check whether

$$\lambda' Q_R (X\ Q_R)^+ X = \lambda' .$$

This test follows from the fact that $Q_R (X\ Q_R)^+ X$ is idempotent hence λ' is of the form $\lambda' = \delta' Q_R (X\ Q_R)^+ X$ if and only if $\lambda' Q_R (X\ Q_R)^+ X = \lambda'$.

In general, bias of the form $\lambda' B_1(\beta)$ is best eliminated by not using $\lambda'\tilde{\beta}$ to estimate $\lambda'\beta$ if $R\beta \neq r$. That is, do not use false restrictions. (Toro-Vizcarrondo and Wallace [4] consider the question of using possibly false restrictions to reduce mean square error under the condition that $\operatorname{rank}(X) = p$ and $\operatorname{rank}(R) = q$.)

The second source of bias $\lambda' B_2(\beta)$ is eliminated when $\lambda' = \delta' X + \rho' R$ since $X\, Q_V = R\, Q_V = 0$. (In fact, $\lambda' B_2(\beta) = 0$ for all β satisfying $R\beta = r$ if and only if $\lambda' = \delta' X + \rho' R$ by Theorem 2). Discussions of estimability (Theil [3], pp. 147, 152) revolve around this second component of bias and the conditions under which it vanishes.

Consider, now, an attempt to estimate an arbitrary function of the parameters $\lambda'\beta$ using $\lambda'\tilde{\beta}$ when $R\beta = r$. Since

$$I = P_R + P_W + Q_V$$

we can write

$$\lambda' = \lambda' P_R + \lambda' P_W + \lambda' Q_V$$

$$= \lambda_1' + \lambda_2' + \lambda_3' \,.$$

It can be verified that

$$\mathcal{E}(\lambda'\tilde{\beta}) = \mathcal{E}(\lambda_1'\tilde{\beta}) + \mathcal{E}(\lambda_2'\tilde{\beta}) + \mathcal{E}(\lambda_3'\tilde{\beta})$$

$$\operatorname{Var}(\lambda'\tilde{\beta}) = \operatorname{Var}(\lambda_1'\tilde{\beta}) + \operatorname{Var}(\lambda_2'\tilde{\beta}) + \operatorname{Var}(\lambda_3'\tilde{B}) \,.$$

If the true but unknown value of β satisfies $R\beta = r$ then these expectations and variances are

$$\mathcal{E}(\lambda_1'\tilde{\beta}) = \lambda_1'\beta \ , \qquad \operatorname{Var}(\lambda_1'\tilde{\beta}) = 0 \ ,$$

$$\mathcal{E}(\lambda_2'\tilde{\beta}) = \lambda_2'\beta \ , \qquad \operatorname{Var}(\lambda_2'\tilde{\beta}) = \lambda_2'W^+(W^+)'\lambda_2\sigma^2 \ ,$$

$$\mathcal{E}(\lambda_3'\tilde{\beta}) = 0 \ , \qquad \operatorname{Var}(\lambda_3'\tilde{\beta}) = \lambda_3'W^+(W^+)'\lambda_3\sigma^2 \ .$$

Inspection of these expectations and variances indicates that the component $\lambda_1'\tilde{\beta}$ of $\lambda'\tilde{\beta}$ is the constant $\lambda_1'R^+r$ (since $R\beta = r$ implies $P_R(\beta - R^+r) = 0$) regardless of the value taken on by the random variable y . Thus, any information about $\lambda_1'\beta$ contained in the sample y is completely overridden by the restriction $R\beta = r$.

The second component, $\lambda_2'\tilde{\beta}$, varies with y and is the portion of $\lambda'\beta$ estimated from the sample.

If the third component $\lambda_3' = \lambda' Q_V$ of λ is not zero, then $\lambda_3'\beta$ (and hence $\lambda'\beta$) cannot be estimated unbiasedly using $\lambda'\tilde{\beta}$. If the estimation of $\lambda'\beta$ is important to the econometric investigation the investigator must augment V by row vectors which will yield λ_3' as a linear combination and recompute $\tilde{\beta}$ using the additional information. V can be augmented by appending additional data

$$y_{(2)} = X_{(2)}\beta + e_{(2)}$$

and additional restrictions

$$R_{(2)}\beta = r_{(2)}$$

to the original model. If observations with the rows of $X_{(2)}$ as inputs can be obtained and restrictions $R_{(2)}\beta = r_{(2)}$ can be deduced such that

$$\lambda_3' = a'X_{(2)} + b'R_{(2)}$$

then $\lambda'\beta$ can be estimated unbiasedly by $\tilde{\beta}$ computed from the augmented model provided the true value of β satisfies

$$\begin{pmatrix} R \\ R_{(2)} \end{pmatrix} \beta = \begin{pmatrix} r \\ r_{(2)} \end{pmatrix} .$$

In summary, we recommend that the results of this section be used in applications to estimate a linear parametric function $\lambda'\beta$ as follows. First, check that $\lambda' Q_R (X\, Q_R)^+ X \neq \lambda'$ since if equality holds, the estimate based on $\lambda'\tilde{\beta}$ coincides with the unrestricted least squares estimate $\lambda'\hat{\beta}$. This may be either a comfort or a disappointment, depending on the application. The variance estimate $\tilde{\sigma}^2$ has more degrees freedom than the estimator $\hat{\sigma}^2$ but the extra degrees of freedom may not be worth the extra bother of computing $\tilde{\beta}$. Second, check that $\lambda' Q_V = 0$ to be sure $\lambda'\tilde{\beta}$ estimates $\lambda'\beta$ unbiasedly. Thirdly, one may wish to compute $\lambda_1' = \lambda' P_R$ and $\lambda_2' = \lambda' P_W$ to determine the information which is due to the restrictions $R\beta = r$ and that which is due to the sample y .

5. COMPUTATIONS

For a given matrix A of order $(m \times n)$ with $m \geq n$ we may decompose A as

$$A = U\,S\,V' ,$$

where U is $(m \times n)$, S is an $(n \times n)$ diagonal matrix, V' is $(n \times n)$ and

$$I_n = U'U = V'V = V\,V' .$$

This is called the singular value decomposition of A [2]. Let s_i denote the diagonal elements of S. Set $s_i^+ = 1/s_i$ if $s_i > 0$, set $s_i^+ = 0$ if $s_i = 0$ and form the diagonal S^+ matrix from the elements s_i^+. Then the Moore-Penrose (pseudo) inverse of A is given by

$$A^+ = V\,S^+U'$$

and the rank of A is the same as the rank of S^+. (If $m < n$ compute $B = (A')^+$ using this method and set $A^+ = B'$.)

A listing of a FORTRAN subroutine to obtain the singular value decomposition of A may be found in [1]. The subroutine as listed is for a COMPLEX matrix A, but we had no difficulty in converting it to REAL*8 from the COMPLEX version. We have had good results using an IBM 370/165 setting the parameters ETA = 1.D - 14 and TOL = 1.D - 60; we take S(I) = 0 if S(I) . LT . S(1) * 1.D - 13 .

If y and X are too large for storage in core but $y'y$, $X'y$, and $X'X$ can be stored then the computational formulas

$$\tilde{\beta} = Q_R(Q_R X'X\,Q_R)^+ Q_R(X'y - X'X\,R^+ r) + R^+ r$$

$$C(\tilde{\beta}\tilde{\beta}') = Q_R(Q_R X'X\,Q_R)^+ Q_R\,\sigma^2$$

$$\tilde{\sigma}^2 = (y'y - 2\tilde{\beta}'X'y + \tilde{\beta}'X'X\tilde{\beta})/(n - \text{rank}(Q_R X'X\, Q_R))$$

may be used. If the formulas

$$\tilde{\beta} = Q_R(X\, Q_R)^+(y - X\, R^+ r) + R^+ r$$

$$C(\tilde{\beta}\tilde{\beta}) = Q_R(X\, Q_R)^+(X\, Q_R)^{+\prime} Q_R\, \sigma^2$$

$$\tilde{\sigma}^2 = (y - X\,\tilde{\beta})'(y - X\,\tilde{\beta})/(n - \text{rank}(X\, Q_R))$$

are feasible, their use should improve the accuracy of the computations by avoiding unnecessary matrix multiplications.

For the computation of the transformation matrix T we will make the assumption that two matrices the size of Σ: $(n^* \times n^*)$ may be stored in core. The singular value decomposition subroutine can be used to obtain U, S, V (since $U = V$ in this case) and the diagonal matrix S stored as a vector with diagonal elements $s_1 \geq s_2 \geq \ldots \geq s_{n^*} \geq 0$. If Σ is non-singular, form the diagonal $(n^* \times n^*)$ matrix D with elements $d_i = (s_i)^{-\frac{1}{2}}$ and $T = D\,U'$. If Σ has rank $n < n^*$ then S will have elements $s_1 \geq s_2 \geq \ldots \geq s_n > s_{n+1} = \ldots = s_{n^*} = 0$. Form the diagonal $(n \times n)$ matrix $D_{(1)}$ with elements $d_i = (s_i)^{-\frac{1}{2}}$ and partition $U' = \begin{pmatrix} U'_{(1)} \\ U'_{(2)} \end{pmatrix}$ where $U'_{(1)}$ is $(n \times n^*)$ and $U'_{(2)}$ is $(n^* - n \times n)$. Then $T_{(1)} = D_{(1)} U'_{(1)}$ and $T_{(2)} = U'_{(2)}$.

In most applications where Σ is known it will be patterned in such a way that knowledge of T for small n^* can be used to deduce the form of T for the problem at hand. Thus the storage requirement is not as stringent as it would first appear.

6. EXAMPLE

Consider a series of quarterly measurements on a variate y with an unconstrained model given by

$$y_{ti} = a + bt + Q_i + e_{ti} ,$$

where the years are denoted by $t = 1, 2, \ldots, 32$ and the quarters by $i = 1, 2, 3, 4$. For the first thirty years the parameters were estimated subject to the constraints

$$\Sigma_{i=1}^{4} Q_i = 0$$

$$Q_1 - Q_4 = 0$$

$$Q_2 - Q_3 = 0$$

yielding the estimates

$$\tilde{\beta}_{30} = (1.1581, .53227, 1.0386, -1.0386, -1.0386, 1.0386)' .$$

We suspect that the last two restrictions are false and that the data follow a quarterly effects pattern rather than the winter/summer pattern used to estimate β from the first thirty years. Our problem will be to estimate β subject to the constraint

$$Q_1 + Q_2 + Q_3 + Q_4 = 0$$

and test the hypotheses

$$T = \begin{pmatrix} 5.3279 & .26223 & 0 \\ 0 & 0 & 10.954 \\ -9.5715 & -194.47 & 0 \end{pmatrix}$$

was computed using the method suggested in the preceeding section.

By combining $T\ \Lambda'\tilde{\beta}_{30}$ and $T\ \Lambda'$ with the data for the years 31 and 32 we obtain X and y as given in Table 1.

TABLE 1

y	X						
6.0307	5.3279	- 0.26223	1.3320	1.3320	1.3320	1.3320
11.377	0.	0.	2.7385	-2.7385	-2.7385	-2.7385
-114.60	-9.5715	-194.47	-2.3929	-2.3929	-2.3929	-2.3929
18.52	1.	31.	1.	0.	0.	0.
16.65	1.	31.	0.	1.	0.	0.
16.71	1.	31.	0.	0.	1.	0.
18.79	1.	31.	0.	0.	0.	1.
19.00	1.	32.	1.	0.	0.	0.
17.03	1.	32.	0.	1.	0.	0.
16.91	1.	32.	0.	0.	1.	0.
19.61	1.	32	0.	0.	0.	1.

The estimate of

$$\beta = (a,\ b,\ Q_1,\ Q_2,\ Q_3,\ Q_4)'$$

subject to

$$Q_1 + Q_2 + Q_3 + Q_4 = 0$$

is

$$H_1: \; Q_1 = Q_4$$

$$H_2: \; Q_2 = Q_3$$

using as much of the information from the previous study as possible.

The matrix $Q_R(X Q_R)^+X$ and variance-covariance matrix of $\tilde{\beta}_{30}$ can be obtained since we know the form taken by X and $R\beta = r$ for the first thirty years. Observe that $\tilde{\beta}_{30}$ obtained in the previous study must coincide with $\tilde{\beta}$ as defined in this paper since $\text{rank}(V) = p = 6$.

The linearly independent rows of $Q_R(X Q_R)^+X$ are

$$\Lambda' = \begin{pmatrix} 1 & 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & -1/4 & -1/4 & 1/4 \end{pmatrix}$$

so that $\Lambda'\beta$ is estimated unbiasedly by

$$\Lambda'\tilde{\beta}_{30} = (1.1581, .53227, 1.0386)'$$

with variance-covariance matrix

$$\Lambda'\text{Var}(\tilde{\beta})\,\Lambda = \begin{pmatrix} .035057 & -.0017241 & 0 \\ -.0017241 & .00011123 & 0 \\ 0 & 0 & .0083333 \end{pmatrix} \sigma^2 \quad .$$

The transformation matrix

$$\tilde{\beta} = (1.161,\ 0.532,\ 0.821,\ -1.026,\ -1.056,\ 1.261)' .$$

The (uncorrelated) estimates of $Q_2 - Q_3$ and $Q_1 - Q_4$ are .030 and -.440 each with the same 6 d.f. estimated variance of .0207. The respective F_6^1 values are .435 and 9.35. We fail to reject H_1 and reject H_2 at a significance level of .025.

The vectors

$$\lambda_1' = (0,\ 0,\ 1,\ 0,\ 0,\ -1)$$

$$\lambda_2' = (0,\ 0,\ 0,\ 1,\ -1,\ 0)$$

are each of the form

$$\lambda' = \delta' P_W \quad \text{and} \quad \lambda' = \delta' Q_R (X\ Q_R)^+ X .$$

Thus the estimates of $Q_1 - Q_4$ and $Q_2 - Q_3$ vary with the sample data and are estimated unbiasedly even if the restriction is false.

If we use the outcome of our tests to re-estimate β subject to

$$Q_1 + Q_2 + Q_3 + Q_4 = 0$$

$$Q_1 - Q_4 = 0$$

we obtain the results given in Table 2.

TABLE 2

$\tilde{\beta}'$					
1.161	.532	.821	-1.041	-1.041	1.261

$\hat{Var}(\tilde{\beta})$					
.0058	-.000027	0	0	0	0
-.000027	.0000016	0	0	0	0
0	0	.0046	-.00014	-.00014	-.0043
0	0	-.00014	.00014	.00014	-.00014
0	0	-.00014	.00014	.00014	-.00014
0	0	-.0043	-.00014	-.00014	.0046

P_R					
0	0	0	0	0	0
0	0	0	0	0	0
0	0	1/4	1/4	1/4	1/4
0	0	1/4	3/4	-1/4	1/4
0	0	1/4	-1/4	3/4	1/4
0	0	1/4	1/4	1/4	1/4

P_W					
1	0	0	0	0	0
0	1	0	0	0	0
0	0	3/4	-1/4	-1/4	-1/4
0	0	-1/4	1/4	1/4	-1/4
0	0	-1/4	1/4	1/4	-1/4
0	0	-1/4	-1/4	-1/4	3/4

$Q_R(X\ Q_R)^+X$					
1	0	1/4	1/4	1/4	1/4
0	1	0	0	0	0
0	0	3/4	-1/4	-1/4	-1/4
0	0	-1/4	1/4	1/4	-1/4
0	0	-1/4	1/4	1/4	-1/4
0	0	-1/4	-1/4	-1/4	3/4

APPENDIX

1. Verification of the properties of P_R, P_W, Q_V. Let A be a $(m \times n)$ matrix and let T be $(m \times m)$ and non-singular. It follows that $(TA)^+(TA) = A^+A$. To see this observe that

$$[(TA)^+(TA) - A^+A]'[(TA)^+(TA) - A^+A]$$

$$= (TA)^+(TA)[I - A^+A] + A^+T^{-1}TA[I - (TA)^+(TA)]$$

$$= (TA)^+T\,0 + A^+T^{-1}0 = 0 .$$

Since

$$\begin{pmatrix} X\,Q_R \\ R \end{pmatrix} = \begin{pmatrix} I & -X\,R^+ \\ 0 & I \end{pmatrix} \begin{pmatrix} X \\ R \end{pmatrix} = TV ,$$

where T is non-singular, we obtain

$$P_V = \begin{pmatrix} X\,Q_R \\ R \end{pmatrix}^+ \begin{pmatrix} X\,Q_R \\ R \end{pmatrix}$$

$$= \begin{pmatrix} X\,Q_R \\ R \end{pmatrix}' \begin{pmatrix} X\,Q_R X' & 0 \\ 0 & RR' \end{pmatrix}^+ \begin{pmatrix} X\,Q_R \\ R \end{pmatrix}$$

$$= \begin{pmatrix} X\,Q_R \\ R \end{pmatrix}' \begin{pmatrix} (X\,Q_R X')^+ & 0 \\ 0 & (RR')^+ \end{pmatrix} \begin{pmatrix} X\,Q_R \\ R \end{pmatrix}$$

$$= P_W + P_R ,$$

hence $I = P_R + P_W + Q_V$. Now

$$Q_V = (I - P_V) = (I - P_R - P_W)$$

$$= (Q_R - P_W) = (Q_R - (X\ Q_R)^+(X\ Q_R)\ Q_R)$$

$$= (I - P_W)Q_R = Q_W Q_R \ .$$

Since Q_V, Q_W, Q_R are symmetric and idempotent

$$Q_V = Q_W Q_R = Q_R' Q_W' = Q_R Q_W$$

$$= Q_R(Q_R Q_W) = Q_R Q_V = Q_R Q_W Q_R \ .$$

Lastly

$$P_W P_R = (X\ Q_R)^+(X\ Q_R)P_R = (X\ Q_R)^+ X\ 0 = 0$$

$$P_R Q_V = P_R(I - P_R - P_W) = P_R(Q_R - P_W) = 0 - 0 = 0$$

$$P_W Q_V = P_W(I - P_R - P_W) = P_W(Q_W - P_R) = 0 - 0 = 0 \ .$$

2. Verification that $Var(\lambda'\tilde{\beta}) = Var(\lambda_1'\tilde{\beta}) \pm Var(\lambda_2'\tilde{\beta}) + Var(\lambda_3'\tilde{\beta})$. It is required to show that $P_R\ Cov(\widetilde{\beta\beta'})P_W = P_R\ Cov(\widetilde{\beta\beta'})Q_V =$ $P_W\ Cov(\widetilde{\beta\beta'})Q_V = 0$.

Now $Cov(\widetilde{\beta\beta'}) = Q_R(X\ Q_R)^+(X\ Q_R)^{+'}Q_R\sigma^2$. Since $P_R Q_R = 0$ we have the first two equalities.

$$\sigma^{-2}P_W \mathrm{Cov}(\tilde{\beta}\tilde{\beta}')Q_V = P_W Q_R (X\ Q_R)^{+}(X\ Q_R)^{+'} Q_R Q_V$$

$$= (X\ Q_R)^{+}(X\ Q_R)(X\ Q_R)^{+}(X\ Q_R)^{+'} Q_V$$

$$= W^{+}\{W^{+}\}' Q_W Q_R$$

$$= W^{+}\{W'(W\ W')^{+}\}' Q_W Q_R$$

$$= W^{+}(W\ W')^{+'} W\ Q_W Q_R$$

$$= W^{+}(W\ W')^{+'} 0\ Q_R = 0\ .$$

REFERENCES

[1] Businger, P. A. and Golub, G. H.: "Singular Value Decomposition of a Complex Matrix," Communications of the A C M, 12, 1969, pp. 564-65.

[2] Golub, G. and Kahan, W.: "Calculating the Singular Values and Pseudo-Inverse of a Matrix," S I A M Journal on Numerical Analysis, 2, 1965, pp. 205-24.

[3] Theil, Henri: Principles of Econometrics. John Wiley and Sons, Inc., New York, 1971.

[4] Wallace, T. D. and Toro-Vizcarrondo, C. E.: "A Test of the Mean Square Error Criterion for Restrictions in Linear Regression," Journal of the American Statistical Association, 63, 1968, pp. 558-72.

APPENDIX

Information for TUCC Users

Three Fortran subroutines are described below which can be used to analyze constrained linear models data. They are stored at TUCC and may be called by users through Fortran programs. Briefly, REGR2 estimates β for the linear model $y = X\beta + e$ subject to the consistent constraints $R\beta = r$, REGR 3 estimates and gives the decomposition of a set of linear functions of the parameters β, and REGR4 tests the hypothesis $H_0: G\beta = G\beta_0$, where β_0 satisfies the equations $R\beta_0 = r$. To illustrate their use, a Fortran program with an input subroutine and some sample data are also given.

The following is the Job Control Language (JCL) required to access the subroutines.

JCL TO RUN THE FORTRAN (G) COMPILER.

```
//JOBNAME JOB ACCOUNT,NAME
// EXEC FTGCG
//C.SYSIN DD *
      (SOURCE PROGRAM)
//G.SYSLIB DD DSN=NCS.ES.B4139.GALLANT.GALLANT,DISP=SHR
//          DD DSN=SYS1.FORTLIB,DISP=SHR
//          DD DSN=SYS1.SUBLIB,DISP=SHR
//G.SYSIN DD *
      (DATA CARDS)
```

JCL TO RUN THE FORTRAN (H) COMPILER.

```
//JOBNAME JOB ACCOUNT,NAME
// EXEC FTHCG
//C.SYSIN DD *
      (SOURCE PROGRAM)
//G.SYSLIB DD DSN=NCS.ES.B4139.GALLANT.GALLANT,DISP=SHR
//          DD DSN=SYS1.FORTLIB,DISP=SHR
//          DD DSN=SYS1.SUBLIB,DISP=SHR
//G.SYSIN DD *
      (DATA CARDS)
```

DGMPNT 10/6/71

PURPOSE
PRINT A MATRIX

USAGE
CALL DGMPNT(A,N,M)

ARGUMENTS
A - INPUT N BY M MATRIX
    STORED COLUMNWISE (STORAGE MODE OF 0)
    ELEMENTS OF A ARE REAL*8
N - NUMBER OF ROWS IN A
M - NUMBER OF COLUMNS IN A

REGR2 2/22/72

PURPOSE
ESTIMATE B FOR THE LINEAR MODEL Y=X*B+E SUBJECT TO THE CONSTRAINTS RR*B=R.

USAGE
CALL REGR2(YPY,XPY,XPX,RR,R,N,IP,IQ,B,C,VAR,IDF,P1,P2,P3,P4)

SUBROUTINES CALLED
DGMPRD, DGMADD, DGMSUB, DAPLUS, DSVD

ARGUMENTS
YPY - INPUT SCALAR CONTAINING (Y-TRANSPOSE)*Y.
REAL*8
XPY - INPUT VECTOR OF LENGTH IP CONTAINING (X-TRANSPOSE)*Y.
ELEMENTS OF XPY ARE REAL*8
XPX - INPUT IP BY IP MATRIX CONTAINING (X-TRANSPOSE)*X. STORED COLUMNWISE (STORAGE MODE OF 0).
ELEMENTS OF XPX ARE REAL*8
RR - INPUT IQ BY IP MATRIX OF CONSTRAINTS. STORED COLUMNWISE (STORAGE MODE OF 0)
ELEMENTS OF RR ARE REAL*8
R - INPUT VECTOR OF LENGTH IQ CONTAINING THE RIGHT HAND SIDE OF THE CONSTRAINT EQUATIONS.
ELEMENTS OF R ARE REAL*8
N - NUMBER OF OBSERVATIONS.
INTEGER
IP - NUMBER OF PARAMETERS IN THE MODEL. IP MUST BE LESS THAN 101
INTEGER
IQ - NUMBER OF CONSTRAINTS. IQ MUST BE GREATER THAN 0 AND LESS THAN OR EQUAL TO IP.
INTEGER
B - ESTIMATE OF THE PARAMETERS SUBJECT TO THE CONSTRAINTS.
VECTOR OF LENGTH IP.
ELEMENTS OF B ARE REAL*8
C - ESTIMATED IP BY IP VARIANCE-COVARIANCE MATRIX OF B. STORED COLUMNWISE (STORAGE MODE OF 0).
ELEMENTS OF C ARE REAL*8
VAR - ESTIMATED VARIANCE.
REAL*8
IDF - DEGREES FREEDOM OF VAR.
INTEGER
P1 - ROW SPACE OF SPECIFIED PARAMETRIC FUNCTIONS. ESTIMATED UNBIASEDLY BY B PROVIDED THE TRUE VALUE SATISFIES RR*B=R.
P2 - ROW SPACE OF REMAINING PARAMETRIC FUNCTIONS ESTIMATED UNBIASEDLY BY B PROVIDED THE TRUE VALUE SATISFIES RR*B=R.
P3 - ROW SPACE OF PARAMETRIC FUNCTIONS ESTIMATED SUBJECT TO BIAS BY B.
P1,P2,P3 ARE SYMMETRIC IDEMPOTENT IP BY IP MATRICES STORED COLUMNWISE (STORAGE MODE OF 0). P1+P2+P3=I, P1*P2=P2*P3=P1*P3=0. ELEMENTS ARE REAL*8.

P4 - ROW SPACE OF PARAMETRIC FUNCTIONS ESTIMATED UNBIASEDLY BY B
WHETHER OR NOT THE TRUE VALUE SATISFIES RR*B=R.
IDEMPOTENT IP BY IP MATRIX STORED COLUMNWISE (STORAGE MODE
OF 0). ELEMENTS ARE REAL*8

REGR3 2/25/72

PURPOSE
DECOMPOSE AND ESTIMATE A SET OF IG LINEAR PARAMETRIC FUNCTIONS USING THE OUTPUT FROM SUBROUTINE REGR2.

USAGE
CALL REGR3(G,B,C,P1,P2,P3,P4,IG,IP,GB,GCG,I1,I2,I3,I4)

ARGUMENTS
G - INPUT IG BY IP MATRIX OF COEFFICIENTS. STORED COLUMNWISE (STORAGE MODE 0).
ELEMENTS OF G ARE REAL*8
B - INPUT VECTOR OF LENGTH IP RETURNED BY REGR2.
ELEMENTS OF B ARE REAL*8
C - INPUT IP BY IP MATRIX RETURNED BY REGR2. STORED COLUMNWISE (STORAGE MODE 0)
ELEMENTS OF C ARE REAL*8
P1 - INPUT IP BY IP MATRIX RETURNED BY REGR2. ON RETURN CONTAINS THE IG BY IP MATRIX G*P1 STORED COLUMNWISE (STORAGE MODE 0).
P2 - INPUT IP BY IP MATRIX RETURNED BY REGR2. ON RETURN CONTAINS THE IG BY IP MATRIX G*P2 STORED COLUMNWISE (STORAGE MODE 0).
P3 - INPUT IP BY IP MATRIX RETURNED BY REGR2. ON RETURN CONTAINS THE IG BY IP MATRIX G*P3 STORED COLUMNWISE (STORAGE MODE 0).
P4 - INPUT IP BY IP MATRIX RETURNED BY REGR2. ON RETURN CONTAINS THE IG BY IP MATRIX G*P4 STORED COLUMNWISE (STORAGE MODE 0).
ELEMENTS OF P1, P2, P3, P4 ARE REAL*8.
IG - NUMBER OF LINEAR PARAMETRIC FUNCTIONS TO BE ESTIMATED.
INTEGER
IP - NUMBER OF PARAMETERS (LENGTH OF B).
INTEGER
GB - VECTOR OF LENGTH IG CONTAINING THE ESTIMATES OF THE LINEAR PARAMETRIC FUNCTIONS, G*B.
ELEMENTS OF GB ARE REAL*8
GCG - ESTIMATED IG BY IG VARIANCE-COVARIANCE MATRIX OF CB. STORED COLUMNWISE (STORAGE MODE 0).
ELEMENTS OF GCG ARE REAL*8
I1 - VECTOR OF LENGTH IG.
I2 - VECTOR OF LENGTH IG.
I3 - VECTOR OF LENGTH IG.
I4 - VECTOR OF LENGTH IG.
I1(I)=0 IF ROW I OF G SATISFIES GI*P1=0.
I1(I)=1 IF ROW I OF G SATISFIES GI*P1=GI.
I1(I)=-1 IF NEITHER OF THE ABOVE ARE SATISFIED BY GI.
SIMILARLY FOR I2, I3, I4.
ELEMENTS OF I1, I2, I3, I4 ARE INTEGERS.

REMARK
BE SURE P1, P2, P3, P4 ARE DIMENSIONED LARGE ENOUGH TO CONTAIN MAX(IP*IP,IG*IP) ELEMENTS IN THE CALLING PROGRAM.

REGR4 3/15/72

PURPOSE
TEST H:GB=0 USING OUTPUT FROM REGR2 AND REGR3.

USAGE
CALL REGR4(GB,GCG,IG,IDF,F,IR,SF,P1,P2,P3,P4)

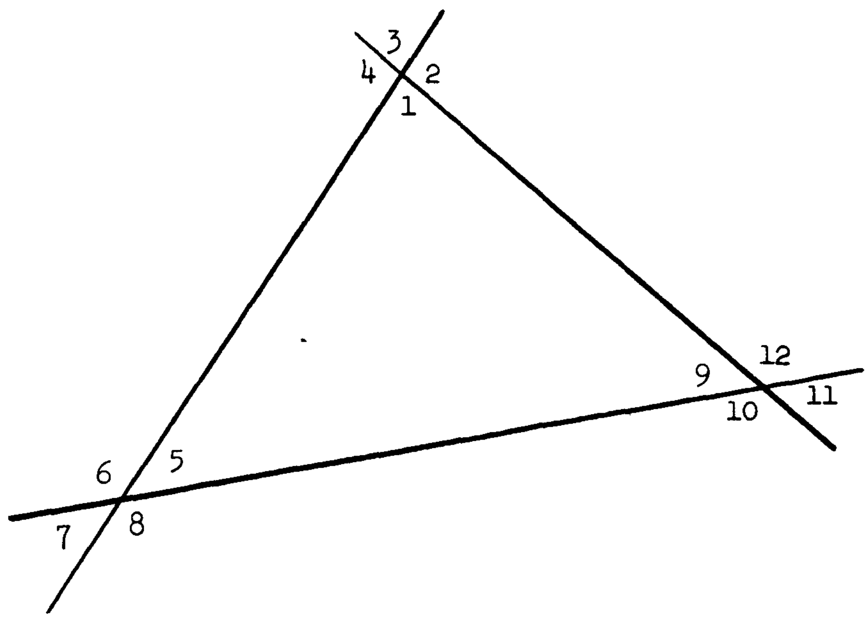
SUBROUTINES CALLED
DAPLUS, DGMPRD, BDTR, DSVD

ARGUMENTS
GB - INPUT VECTOR OF LENGTH IG RETURNED BY REGR3.
ELEMENTS OF GB ARE REAL*8.
GCG - INPUT IG BY IG MATRIX RETURNED BY REGR3. STORED COLUMNWISE (STORAGE MODE 0).
ELEMENTS OF GCG ARE REAL*8.
IG - LENGTH OF GB; NUMBER OF ROWS AND COLUMNS IN GCG.
IG MUST BE LESS THAN 100.
INTEGER
IDF - INPUT INTEGER RETURNED BY REGR2; DENOMINATOR D.F. FOR F.
INTEGER
F - COMPUTED F STATISTIC
REAL*8
IR - COMPUTED NUMERATOR D.F. FOR F, RANK OF GCG.
INTEGER
SF - SIGNIFICANCE LEVEL OF F. (I.E. 1-CDF(F)).
REAL*8
P1 - IG BY IG MATRIX USED AS WORKSPACE.
P2 - IG BY IG MATRIX USED AS WORKSPACE.
P3 - IG BY IG MATRIX USED AS WORKSPACE.
P4 - IG BY IG MATRIX USED AS WORKSPACE.
ELEMENTS OF P1,P2,P3,P4 ARE REAL*8.

REMARK
THE RESULTS RETURNED BY REGR4 ARE INVALID IF B=0 DOES NOT SATISFY RR*B=R. TO TEST H:GB=G*BO WHERE RR*BO=R INPUT G*(B-BO) INSTEAD OF GB.

Sample Problem

Measurements are taken on the 12 angles of the following figure.



We assume that the following linear model is appropriate to describe the data.

$$y = X\beta + e \quad \text{subject to} \quad R\beta = r \; ,$$

where y: (12×1) and X: (12×6) are given in Table 3, β_1 corresponds to angles 1, 3, β_2 to 2, 4, β_3 to 5, 7 β_4 to 6, 8, β_5 to 9, 11, β_6 to 10, 12,

$$R = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

and

$$r = (180,\ 180,\ 180,\ 180)' \ .$$

We wish to test the hypothesis that the triangle is equilateral. That is H_o: $G\beta = 0$, where

$$G = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} .$$

It is interesting to note that the test of H_o can also be interpreted as a test of no regression effect.

The quantities $\tilde{\beta}$, $\hat{Var}(\tilde{\beta})$, P_R , P_W , and $Q_R(XQ_R)^+X$ are given in Table 4. The results of the test of H_o are

$$G\hat{\beta} = \begin{pmatrix} -2.025 \\ -0.500 \end{pmatrix} ,$$

$$\hat{Var}(G\hat{\beta}) = \begin{pmatrix} 0.2749 & 0.1375 \\ 0.1375 & 0.2749 \end{pmatrix} ,$$

and

$$F^2_{10} = 8.095 \quad (p = 0.0081) .$$

It is also found that the rows of G are neither in the row space of R nor in that of XQ_R , but are orthogonal to the row space of Q_V . This indicates that information from both the restrictions and the data went into the estimation of the $G\beta$ and that, if the restrictions are valid, the estimates are unbiased. Finally, it is found that the rows of G are not in $Q_R(XQ_R)^+X$ indicating that if the restrictions are false then the estimates are biased.

TABLE 3

Angle	Measurements: y	Design Matrix: X					
1	59.1	1	0	0	0	0	0
2	120.5	0	1	0	0	0	0
3	58.6	1	0	0	0	0	0
4	122.1	0	1	0	0	0	0
5	60.4	0	0	1	0	0	0
6	119.8	0	0	0	1	0	0
7	61.3	0	0	1	0	0	0
8	118.7	0	0	0	1	0	0
9	60.1	0	0	0	0	1	0
10	120.7	0	0	0	0	0	1
11	59.2	0	0	0	0	1	0
12	121.5	0	0	0	0	0	1

TABLE 4

		$\tilde{\beta}$			
59.158	120.842	61.183	188.817	59.658	120.342
		$\hat{Var}(\tilde{\beta})$			
.09164	-.09164	-.04582	.04582	-.4582	.04582
-.09164	.09164	.04582	-.04582	.04582	-.04582
-.04582	.04582	.09164	-.09164	-.04582	.04582
.04582	-.04582	-.09164	.09164	.04582	-.04582
-.04582	.04582	-.04582	.04582	.09164	-.09164
.04582	-.04582	.04582	-.04582	-.09164	.90164
		P_R			
4/6	2/6	1/6	-1/6	1/6	-1/6
2/6	4/6	-1/6	1/6	-1/6	1/6
1/6	-1/6	4/6	2/6	1/6	-1/6
-1/6	1/6	2/6	4/6	-1/6	1/6
1/6	-1/6	1/6	-1/6	4/6	2/6
-1/6	1/6	-1/6	1/6	2/6	4/6
		P_W			
2/6	-2/6	-1/6	1/6	-1/6	1/6
-2/6	2/6	1/6	-1/6	1/6	-1/6
-1/6	1/6	2/6	-2/6	-1/6	1/6
1/6	-1/6	-2/6	2/6	1/6	-1/6
-1/6	1/6	-1/6	1/6	2/6	-2/6
1/6	-1/6	1/6	-1/6	-2/6	2/6
		$Q_R(XQ_R)^+X$			
2/6	-2/6	-1/6	1/6	-1/6	1/6
-2/6	2/6	1/6	-1/6	1/6	-1/6
-1/6	1/6	2/6	-2/6	-1/6	1/6
1/6	-1/6	-2/6	2/6	1/6	-1/6
-1/6	1/6	-1/6	1/6	2/6	-2/6
1/6	-1/6	1/6	-1/6	-2/6	2/6

```
// EXEC FTGCG
//C.SYSIN DD *
C
C  MAIN: CONSTRAINED LINEAR MODELS (C L M)
C
C  PURPOSE:
C  TO ESTIMATE B FOR THE LINEAR MODEL Y=X*B+E SUBJECT TO THE CONSTRAINTS
C  RR*B=R, TO DECOMPOSE AND ESTIMATE A SET OF IG LINEAR PARAMETRIC
C  FUNCTIONS (G*B), AND TO TEST THE HYPOTHESIS H:G*B=G*B0.
C
C  INPUT:
C  THE USER MUST SUPPLY AN INPUT SUBROUTINE OF THE FOLLOWING FORM --
C  SUBROUTINE INPUT(N,IP,IQ,IG,YPY,XPY,XPX,RR,R,G,GB0)
C  WHERE
C  N   - NUMBER OF OBSERVATIONS.
C        INTEGER
C  IP  - NUMBER OF PARAMETERS IN THE MODEL
C        INTEGER
C  IQ  - NUMBER OF CONSTRAINTS
C        INTEGER
C  IG  - NUMBER OF LINEAR PARAMETRIC FUNCTIONS TO BE ESTIMATED
C        INTEGER
C  YPY - SCALAR CONTAINING Y'Y
C        REAL*8
C  XPY - VECTOR OF LENGTH IP CONTAINING X'Y
C        ELEMENTS ARE REAL*8
C  XPX - VECTOR OF LENGTH IP*IP CONTAINING X'X STORED COLUMNWISE
C        ELEMENTS ARE REAL*8
C  RR  - VECTOR OF LENGTH IQ*IP CONTAINING THE MATRIX OF CONSTRAINTS
C        STOPED COLUMNWISE
C        ELEMENTS ARE REAL*8
C  R   - VECTOR OF LENGTH IQ CONTAINING THE RIGHT HAND SIDE OF THE
C        CONSTRAINT EQUATIONS
C        ELEMENTS ARE REAL*8
C  G   - VECTOR OF LENGTH IG*IP CONTAINING THE MATRIX OF COEFFICIENTS
C        FOR THE SET OF LINEAR PARAMETRIC FUNCTIONS STORED COLUMNWISE
C        ELEMENTS ARE REAL*8
C  GB0 - VECTOR OF LENGTH IG CONTAINING THE RIGHT HAND SIDE OF THE
C        HYPOTHESIS EQUATIONS
C        ELEMENTS ARE REAL*8
C  THE ABOVE ARRAYS MAY BE DIMENSIONED IN THE SUBROUTINE AS FOLLOWS
C  REAL*8 YPY,XPY(1),XPX(1),RR(1),R(1),G(1),GB0(1)
C
C  SUBROUTINES USED:
C  DGMPNT,REGR2,REGR3,REGR4,INPUT
C
      IMPLICIT REAL*8 (A-H,O-Z)
      REAL*8 XPY(10),XPX(100),RR(100),R(10),B(10),C(100),P1(100),P2(100)
      REAL*8 P3(100),P4(100),G(100),GB(10),GCG(100),GB0(10)
      INTEGER I1(10),I2(10),I3(10),I4(10)
      CALL INPUT(N,IP,IQ,IG,YPY,XPY,XPX,RR,R,G,GB0)
      WRITE(3,1000)
      WRITE(3,1006) N
      WRITE(3,1007) IP
      WRITE(3,1008) IQ
      WRITE(3,1010) IG
      WRITE(3,1001) YPY
      WRITE(3,1002)
      CALL DGMPNT(XPY,IP, 1)
      WRITE(3,1003)
      CALL DGMPNT(XPX,IP,IP)
      IF(IQ.EQ.0) GO TO 100
```

```
      WRITE(3,1004)
      CALL DGMPNT( RR,IQ,IP)
      WRITE(3,1005)
      CALL DGMPNT(  R,IQ, 1)
      GO TO 400
100   IQ=1
      R(1)=0.0
      DO 300 I=1,IP
300   RR(I)=0.0
400   CALL REGR2(YPY,XPY,XPX,RR,R,N,IP,IQ,B,C,VAR,IDF,P1,P2,P3,P4)
      IF(IG.NE.0) WRITE(3,1009)
      IF(IG.NE.0) CALL DGMPNT(G,IG,IP)
      IF(IG.NE.0) WRITE(3,1031)
      IF(IG.NE.0) CALL DGMPNT(GB0,IG,1)
      WRITE(3,1011)
      CALL DGMPNT(  B,IP,1)
      WRITE(3,1012)
      CALL DGMPNT(  C,IP,IP)
      WRITE(3,1013) VAR
      WRITE(3,1014) IDF
      WRITE(3,1015)
      CALL DGMPNT( P1,IP,IP)
      WRITE(3,1016)
      CALL DGMPNT( P2,IP,IP)
      WRITE(3,1017)
      CALL DGMPNT( P3,IP,IP)
      WRITE(3,1018)
      CALL DGMPNT( P4,IP,IP)
      IF(IG.EQ.0) STOP
      CALL REGR3(G,B,C,P1,P2,P3,P4,IG,IP,GB,GCG,I1,I2,I3,I4)
      WRITE(3,1019)
      CALL DGMPNT(GB,IG,1)
      WRITE(3,1020)
      CALL DGMPNT(GCG,IG,IG)
      WRITE(3,1021)
      CALL DGMPNT(P1,IG,IP)
      WRITE(3,1022)
      CALL DGMPNT(P2,IG,IP)
      WRITE(3,1023)
      CALL DGMPNT(P3,IG,IP)
      WRITE(3,1024)
      CALL DGMPNT(P4,IG,IP)
      WRITE(3,1025)
      WRITE(3,1029) (I1(I),I=1,IG)
      WRITE(3,1026)
      WRITE(3,1029) (I2(I),I=1,IG)
      WRITE(3,1027)
      WRITE(3,1029) (I3(I),I=1,IG)
      WRITE(3,1028)
      WRITE(3,1029) (I4(I),I=1,IG)
      DO 200 I=1,IG
200   GB(I)=GB(I)-GB0(I)
      CALL REGR4(GB,GCG,IG,IDF,F,IR,SF,P1,P2,P3,P4)
      WRITE(3,1030) F,IR,IDF,SF
      STOP
1000  FORMAT('1'///' ANALYSIS OF THE MODEL Y=X*B+E SUBJECT TO RR*B=R'//)
1001  FORMAT(///'0YPY - (Y-TRANSPOSE)*Y'//' ',D15.8)
1002  FORMAT(///'0XPY - (X-TRANSPOSE)*Y')
1003  FORMAT(///'0XPX - (X-TRANSPOSE)*X')
1004  FORMAT(///'0RR - COEFFICIENT MATRIX OF THE RESTRICTIONS RR*B=R')
1005  FORMAT(///'0R - RIGHT HAND SIDE OF THE RESTRICTIONS RR*B=R')
1006  FORMAT(///'0N - NUMBER OF OBSERVATIONS'//' ',I5)
```

```
1007  FORMAT(///'0IP - NUMBER OF PARAMETERS'//' ',I5)
1008  FORMAT(///'0IQ - NUMBER OF RESTRICTIONS'//' ',I5)
1009  FORMAT(///'0G - COEFFICIENT MATRIX OF G*B')
1010  FORMAT(///'0IG - NUMBER OF ROWS IN G'//' ',I5)
1011  FORMAT(///'0B - ESTIMATE OF Y=X*B+E SUBJECT TO RR*B=R')
1012  FORMAT(///'0C - ESTIMATED VARIANCE-COVARIANCE MATRIX OF ESTIMATE')
1013  FORMAT(///'0VAR - ESTIMATE OF VAR(E(I))'//' ',D15.8)
1014  FORMAT(///'0IDF - NUMBER OF D.F. FOR VAR ESTIMATE'//' ',I5)
1015  FORMAT(///'0P1 - ROW SPACE SPECIFIED BY RR*B=R')
1016  FORMAT(///'0P2 - P1+P2 IS ROW SPACE EST. UNBIASEDLY IF RR*B=R')
1017  FORMAT(///'0P3 - ROW SPACE ESTIMATED WITH BIAS')
1018  FORMAT(///'0P4 - ROW SPACE EST. UNBIASEDLY EVEN IF RR*B.NE.R')
1019  FORMAT(///'0GB - ESTIMATE OF G*B')
1020  FORMAT(///'0GCG - ESTIMATED VARIANCE-COVARIANCE MATRIX')
1021  FORMAT(///'0G*P1')
1022  FORMAT(///'0G*P2')
1023  FORMAT(///'0G*P3')
1024  FORMAT(///'0G*P4')
1025  FORMAT(///'0I1 - ROWS = (0,1) IF ROWS G (ORTHOG,IN) ROW SPACE P1')
1026  FORMAT(///'0I2 - ROWS = (0,1) IF ROWS G (ORTHOG,IN) ROW SPACE P2')
1027  FORMAT(///'0I3 - ROWS = (0,1) IF ROWS G (ORTHOG,IN) ROW SPACE P3')
1028  FORMAT(///'0I4 - ROWS = 1 IF ROWS OF G ARE IN THE ROW SPACE P4')
1029  FORMAT(' '/(' ',I8))
1030  FORMAT(///'0F,DF1,DF2,P - TEST OF H:GB=GB0'/'0',F15.5,2I5,F15.5)
1031  FORMAT(///'0GB0 - HYPOTHESIZED VALUE OF G*B')
      END
      SUBROUTINE INPUT(N,IP,IQ,IG,YPY,XPY,XPX,RR,R,G,GB0)
      REAL*8 YPY,XPY(1),XPX(1),RR(1),G(1),R(1),X(10),Y,GB0(1)
      READ(1,15)N,IP,IQ,IG
      YPY=0.0
      DO 10 I=1,IP
      XPY(I)=0.
      DO 10 J=1,IP
10    XPX((J-1)*IP+I)=0.
      DO 20 I=1,N
      READ(1,11) Y,(X(J),J=1,IP)
      YPY=YPY+Y*Y
      DO 30 J=1,IP
      XPY(J)=XPY(J)+Y*X(J)
      DO 30 K=1,IP
      IJ=(K-1)*IP+J
30    XPX(IJ)=XPX(IJ)+X(J)*X(K)
20    CONTINUE
      DO 13 I=1,IQ
13    READ(1,11)R(I),(RR((J-1)*IQ+I),J=1,IP)
      IF(IG.EQ.0) RETURN
      DO 40 I=1,IG
40    READ(1,11) (G((J-1)*IG+I),J=1,IP)
      READ(1,11) (GB0(I),I=1,IG)
      RETURN
11    FORMAT(7F5.1)
15    FORMAT(4I5)
      END
//G.SYSLIB DD DSN=NCS.ES.B4139.GALLANT.GALLANT,DISP=SHR
//          DD DSN=SYS1.FORTLIB,DISP=SHR
//          DD DSN=SYS1.SUBLIB,DISP=SHR
//G.SYSIN DD *
00012000060000400002
59.1 1.   0.   0.   0.   0.   0.
120.50.   1.   0.   0.   0.   0.
58.6 1.   0.   0.   0.   0.   0.
122.10.   1.   0.   0.   0.   0.
```

```
                                                  CLM 1870
60.4 0.    0.    1.    0.    0.    0.             CLM 1880
119.80.    0.    0.    1.    0.    0.             CLM 1890
61.3 0.    0.    1.    0.    0.    0.             CLM 1900
118.70.    0.    0.    1.    0.    0.             CLM 1910
60.1 0.    0.    0.    0.    1.    0.             CLM 1920
120.70.    0.    0.    0.    0.    1.             CLM 1930
59.2 0.    0.    0.    0.    1.    0.             CLM 1940
121.50.    0.    0.    0.    0.    1.             CLM 1950
180. 1.    0.    1.    0.    1.    0.             CLM 1960
180. 1.    1.    0.    0.    0.    0.             CLM 1970
180. 0.    0.    1.    1.    0.    0.             CLM 1980
180. 0.    0.    0.    0.    1.    1.             CLM 1990
1.0  0.0   -1.0  0.0   0.0                        CLM 2000
1.0  0.0   0.0   0.0   -1.0                       CLM 2010
0.0  0.0
```