SCHUR COMPLEMENTS AND STATISTICS

by

DIANE VALÉRIE OUELLETTE

A thesis submitted to the Faculty of Graduate Studies and Research, in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics
McGill University
Montreal

July 1978

SCHUR COMPLEMENTS AND STATISTICS

by

DIANE VALÉRIE OUELLETTE

ABSTRACT

In this thesis we discuss various properties of matrices of the type

$$S = H - GE^{-1}F ,$$

which we call the Schur complement of E in

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} .$$

The matrix E is assumed to be nonsingular. When E is singular or rectangular we consider the generalized Schur complement $S = H - GE^{-}F$, where E^{-} is a generalized inverse of E. A comprehensive account of results pertaining to the determinant, the rank, the inverse and generalized inverses of partitioned matrices, and the inertia of a matrix is given both for Schur complements and generalized Schur complements. We survey the known results in a historical perspective and obtain several extensions. Numerous applications in numerical analysis and statistics are included. The thesis ends with an exhaustive bibliography of books and articles related to Schur complements.

Department of Mathematics
McGill University
Montreal

M.Sc.
July 1978

SCHUR COMPLEMENTS AND STATISTICS

par

DIANE VALÉRIE OUELLETTE

RÉSUMÉ

Dans cette thèse, on étudie les propriétés des matrices du type

$$S = H - GE^{-1}F ,$$

que nous appelons le complément de Schur de E dans

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix},$$

si la matrice E est nonsingulière. Quand E est singulière ou rectangulaire, nous considérons le complément de Schur généralisé $S = H - GE^{-}F$, où E^{-} est un inverse généralisé de E. On présente des résultats concernant l'inertie d'une matrice, le déterminant, le rang, l'inverse et les inverses généralisés des matrices fractionnées pour les compléments de Schur et les compléments de Schur généralisés. Nous examinons ces résultats dans une perspective historique et obtenons plusieurs généralisations. Nous incluons de nombreuses applications en analyse numérique et en statistique. La thèse prend fin avec une bibliographie complète des livres et des articles se rapportant aux compléments de Schur.

Département de Mathématiques
Université McGill
Montréal

M.Sc.
Juillet 1978

ACKNOWLEDGEMENTS

I am especially grateful to Professor George P.H. Styan, who suggested the topic for this thesis and continued to offer support and expert advice to the very end. He also greatly helped in the compilation of the references and bibliography.

Special thanks go to Raffaella Bruno for the excellent job in typing this thesis. I would also like to thank Patricia Babecki and Maura Crilly for completing the typescript.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION AND NOTATION

§1.1 *Introduction*.

"In recent years, the designation "Schur complement" has been applied to any matrix of the form $D-CA^{-1}B$. These objects have undoubtedly been encountered from the time matrices were first used. But today under this new name and with new emphasis on their properties, there is greater awareness of the widespread appearance and utility of Schur complements."

- Cottle (1974).

Our purpose in this thesis is to present a unified treatment covering both the Schur complement

$$S = H - GE^{-1}F \tag{1.1}$$

and the generalized Schur complement

$$S = H - GE^{-}F \,, \tag{1.2}$$

where E^{-} is a generalized inverse of E satisfying $EE^{-}E = E$. We discuss various properties of matrices of the type (1.1) and (1.2) and present both early and recent results. We also show how Schur complements may be used to obtain concise proofs of some well-known and some not so well-known formulas.

Issai Schur (1917) appears to be the first author to explicitly consider a matrix of the form (1.1). He used (1.1) to prove that

$$\begin{vmatrix} E & F \\ G & H \end{vmatrix} = |E|.|H-GE^{-1}F| \,, \tag{1.3}$$

where $|\cdot|$ denotes determinant. The matrix E is assumed to be nonsingular. We present (1.3) in Theorem 2.1.

Emilie V. Haynsworth (1968, p. 74) appears to be the first author to give the name Schur complement to a matrix of the form (1.1). Following her, we refer to

$$S = H - GE^{-1}F \tag{1.4}$$

as the *Schur complement of* E *in* A , where the partitioned matrix

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} . \tag{1.5}$$

The notation

$$S = (A/E) = H - GE^{-1}F \tag{1.6}$$

is convenient.

We may consider the Schur complement of any nonsingular submatrix in A . However, for notational convenience, it is preferable to shift the nonsingular submatrix either to the upper left-hand corner or to the lower right-hand corner of A . This is equivalent to pre-multiplication and/or post-multiplication of A by a permutation matrix.

In the book by Bodewig (1956, 1959), the formula (1.3) is said (1956, p. 189; 1959, p. 218) to date from Frobenius (1849-1917), who obtained, cf. Frobenius (1908),

$$\begin{vmatrix} E & \underset{\sim}{f} \\ \underset{\sim}{g}' & h \end{vmatrix} = h|E| - \underset{\sim}{g}'(adj\ E)\ \underset{\sim}{f} , \tag{1.7}$$

where *adj* denotes adjugate matrix. In (1.7) $\underset{\sim}{f}$ and $\underset{\sim}{g}$ are column vectors while h is a scalar. Boerner (1975) reports that Schur (1875-1941) was a student of Frobenius. We present (1.7) in Theorem 2.3.

Banachiewicz (1937) appears to be the first author to express the inverse of a partitioned matrix in terms of the Schur complement. When the partitioned matrix A in (1.5) and the submatrix E are both nonsingular then the Schur complement of E in A ,

$$S = (A/E) = H - GE^{-1}F \tag{1.8}$$

is also nonsingular, cf. (1.14) below, and

(1.9) $$A^{-1} = \begin{pmatrix} E^{-1} + E^{-1}FS^{-1}GE^{-1} & -E^{-1}FS^{-1} \\ -S^{-1}GE^{-1} & S^{-1} \end{pmatrix},$$

cf. Theorem 2.7.

Banachiewicz (1937) obtained (1.9) in Cracovian notation, where matrices are multiplied column by column (see Appendix A for further details).

The formula (1.9) is often attributed to Schur (1917), cf. e.g., Marsaglia & Styan (1974b, p. 437), but apparently was not discovered until 1937 by Banachiewicz. We will refer to (1.9) as the Schur-Banachiewicz inverse formula.

When the partitioned matrix A in (1.5) and the submatrix H are both nonsingular, then it follows similarly that the Schur complement of H in A

(1.10) $$T = (A/H) = E - FH^{-1}G$$

is also nonsingular and

(1.11) $$A^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}FH^{-1} \\ -H^{-1}GT^{-1} & H^{-1} + H^{-1}GT^{-1}FH^{-1} \end{pmatrix}.$$

When A , E and H are all three nonsingular then

(1.12) $$(E - FH^{-1}G)^{-1} = E^{-1} + E^{-1}F(H-GE^{-1}F)^{-1}GE^{-1},$$

which was observed by Duncan (1944) and reestablished by Woodbury (1950). Equation (1.12) lead to formulas like, cf. Sherman & Morrison (1949, 1950), Bartlett (1951),

(1.13) $$(E + \underset{\sim}{f}\underset{\sim}{g}')^{-1} = E^{-1} - E^{-1}\underset{\sim}{f}\underset{\sim}{g}'E^{-1}/(1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{f}),$$

cf. Corollary 2.6.

Bodewig (1947) has shown that by establishing a count of the number of operations required, the usual method of calculating the determinant of the partitioned matrix (1.5) is preferable to Schur's formula (1.3). Bodewig (1947) claims, however, that the opposite is true when the inverse is calculated, cf. (1.9).

Louis Guttman (1946) established that, if the matrix E in (1.5) is nonsingular, then

$$(1.14) \qquad r(A) = r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(H - GE^{-1}F) = r(E) + r(A/E)$$

where $r(\cdot)$ denotes rank. We present this as Theorem 2.5 . In other words, rank is additive on the Schur complement, cf. Marsaglia & Styan (1974a, p. 291). Wedderburn (1934) and Householder (1964) gave related results on rank, which turn out to be special cases of (1.14). See Theorems 2.6a and 2.6b .

We conclude Chapter II by showing how Schur complements may be used to prove theorems of Cauchy (1812) and Jacobi (1834).

In Chapter III, we discuss various properties of the Schur complement of a nonsingular matrix which have appeared more recently. It seems; cf. the survey paper on Schur complements by Cottle (1974), that from 1952 through 1967 no research papers with results on Schur complements were published.

In a study of the inertia of a partitioned matrix, Haynsworth (1968) showed that when the partitioned matrix A in (1.5) is Hermitian and E is nonsingular, then

$$(1.15) \qquad In\,A = In\,E + In(A/E) ,$$

that is, inertia of a Hermitian matrix is additive on the Schur complement. In Theorem 3.1, we show how rank additivity and inertia additivity are related.

Crabtree & Haynsworth (1969) and Ostrowski (1971) prove that if we partition E as well, i.e.,

$$A = \left(\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right) = \left(\begin{array}{cc|c} K & L & F_1 \\ M & N & F_2 \\ \hline G_1 & G_2 & H \end{array}\right) , \tag{1.16}$$

with E and K both nonsingular, then

$$(A/E) = ((A/K)/(E/K)) , \tag{1.17}$$

cf. Theorem 3.3. This result, called the *quotient property*, has lead to several determinant inequalities, cf. Haynsworth (1970b) and Hartfiel (1973).

We conclude Chapter III by describing an interpretation for the Schur complement as the coefficient matrix of a quadratic form restricted to the null space of a matrix, as developed by Cottle (1974).

In Chapter IV, we extend the results in Chapters II and III to generalized Schur complements, cf. (1.2). Let the partitioned matrix A in (1.5) and the submatrix E both be square. If either

$$r(E,F) = r(E) \tag{1.18}$$

or

$$r\begin{pmatrix} E \\ G \end{pmatrix} = r(E) \tag{1.19}$$

then

$$|A| = |E| \cdot |H-GE^{-}F| \tag{1.20}$$

for every g-inverse E^- , cf. Theorem 4.1 .

Following Meyer (1973), Marsaglia & Styan (1974a), Carlson, Haynsworth & Markham (1974) and Carlson (1975), we establish several results on rank. Among these, we show, cf. Corollary 4.3, that rank is additive on the Schur complement

(1.21) $$r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(H - GE^{-}F) ,$$

when (1.18) and (1.19) hold.

Following Rohde (1965), Pringle & Rayner (1970), Bhimasankaram (1971), Marsaglia & Styan (1974b) and Burns, Carlson, Haynsworth & Markham (1974), we investigate conditions under which the Schur-Banachiewicz inversion formula (1.9) works with generalized inverses replacing regular inverses, cf. Theorem 4.6 .

Following Carlson, Haynsworth & Markham (1974), we find that inertia continues to be additive on the (generalized) Schur complement, that is,

(1.22) $$In\begin{pmatrix} E & F \\ F' & H \end{pmatrix} = In E + In(H-F'E^{-}F) = In E + In(H-F'E^{+}F) ,$$

where the partitioned matrix is real and symmetric, if

(1.23) $$r(E) = r(E,F) ,$$

cf. (1.18), where E^{-} is any g-inverse of E , cf. page 63.

The quotient property may be extended using generalized Schur complements so that, if in (1.16)

(1.24) $$r(E) = r(E,F) = r\begin{pmatrix} E \\ G \end{pmatrix}$$

and

(1.25) $$r(K) = r(K,L) = r\begin{pmatrix} K \\ M \end{pmatrix}$$

hold then (1.17) is still true. We conclude Chapter IV by showing how readily results like

(1.26a) $$|I - FG| = |I - GF|$$

and

(1.26b) $$\psi(I - FG) = \psi(I - GF)$$

may be established using Schur complements. In (1.26b) $\psi(\cdot)$ denotes nullity.

Chapter V contains a number of algorithms for matrix inversion and for generalized inversion which make use of Schur complements. "The bordering method" published in the book by Frazer, Duncan & Collar (1938), the variant given by Jossa (1940), and the "second-order enlargement" method due to Louis Guttman (1946) are described and are accompanied by numerical examples. A similar method called "geometric enlargement", due to Louis Guttman (1946), is also given.

One of the most useful algorithms, perhaps, is that of partitioned Schur complements outlined by Louis Guttman (1946) and later developed by Zlobec & Chan (1974). Wilf (1959) elaborated a method of rank annihilation, while Edelblute (1966) considered a special case of the above algorithm which simplifies the calculations performed.

Following Newman (1962) and Westlake (1968, p. 31) we show in Section 5.5 how Schur complements may be used to obtain the inverse of a complex matrix using real operations only.

We also present an algorithm due to Zlobec (1975), which computes a g-inverse of a partitioned matrix using partitioned Schur complements. Generalized inversion has also been studied by Ahsanullah & Rahman (1973) who have extended the method of rank annihilation.

Further details of some of these algorithms are given in the books by Faddeeva (1959, pp. 105-111) and Faddeev & Faddeeva (1963, pp. 161-167, 173-178).

In Chapter VI, we describe the areas of mathematical statistics in which Schur complements arise. An excellent example of this is the covariance matrix in a conditional multivariate normal distribution.

In Section 6.2 we consider partial covariances and partial correlation coefficients, and prove the well-known recursion formula for partial correlation coefficients using the quotient property (1.17).

In Section 6.3 we study several special covariance and correlation structures; we easily evaluate, using Schur complements, the determinant, rank, characteristic roots, and inverse of each structure.

In Section 6.4 we show how a quadratic form which follows a chi-squared distribution may be expressed as a Schur complement. We extend this result to show that the Schur complement in a Wishart matrix is also Wishart, and that the Schur complement in the matrix-variate beta distribution is also beta, cf. Mitra (1970).

We conclude Chapter VI and this thesis by showing how the Cramér-Rao inequality for a minimum variance unbiased estimator of a vector-valued parameter may be proved using the inertia additivity of Schur complements, cf. (1.15).

The concept of Schur complement has recently been extended by Ando (1978) as the matrix

$$(1.27) \qquad \begin{pmatrix} 0 & 0 \\ 0 & H-GE^{-1}F \end{pmatrix};$$

he refers to

$$(1.28) \qquad \begin{pmatrix} E & F \\ G & GE^{-1}F \end{pmatrix}$$

as a Schur compression. It follows at once that (1.27) and (1.28) are rank additive, cf. (1.14). Ando uses these new definitions to extend the quotient property (1.17). We hope to consider other extensions at a later time.

§1.2 *Notation.*

Matrices are denoted by capital letters, column vectors by underscored lower case letters and scalars by lower case letters. An n×n matrix A may also be denoted by $\{a_{ij}\}_{i,j=1,\ldots,n}$ and a diagonal matrix whose entries are $a_{11}, a_{22}, \ldots, a_{nn}$ on the diagonal by *diag* $(a_{11}, a_{22}, \ldots, a_{nn})$. In particular, $I = \{\delta_{ij}\}$ represents the identity matrix with δ_{ij} the Kronecker delta, $\underset{\sim}{e}$ or $\underset{\sim}{e}^{(n)}$ the n×1 column vector of ones, $\underset{\sim}{e}_i$ or $\underset{\sim}{e}_i^{(n)}$ the n×1 column vector with all elements zero except for unity in the *ith* position. The transpose of a matrix A is denoted A' , with $\underset{\sim}{a}'$ the row vector corresponding to the column vector $\underset{\sim}{a}$. The determinant is denoted by $|\cdot|$, the adjugate (or adjoint) matrix by *adj* and the trace *tr* . Rank is denoted by $r(\cdot)$ and nullity by $\psi(\cdot)$. We call A^- a *generalized inverse* (or g-inverse) of A if $AA^-A = A$, cf. Rao (1962) and Rao & Mitra (1971). If, in addition, $A^-AA^- = A^-$, or $r(A) = r(A^-)$, then $A^- = A_r^-$, a reflexive g-inverse. If, in addition, the projectors AA_r^- and A_r^-A are both symmetric, then $A_r^- = A^+$, the unique *Moore-Penrose* g-inverse of A .

We denote the characteristic roots of A by $ch(A)$, with $ch_j(A)$ being the *jth* largest when the roots are real. The inertia *In*A of a real symmetric matrix A is the ordered triple (π, ν, δ) , where π is the number of positive characteristic roots of A , ν the number of negative and δ the number of zero roots of A . Thus for a symmetric matrix A we have that $\pi + \nu = r(A)$, the rank of A , and $\delta = \psi(A)$, the nullity of A . In this thesis, positive definite (pd), positive semidefinite (psd) and nonnegative definite (nnd) matrices are always real and symmetric. A matrix is pd if $\nu = \delta = 0$, psd if $\nu = 0$ and $\delta \geq 1$, nnd if $\nu = 0$. Some authors (e.g., Haynsworth, 1968) use positive semidefinite where we use nonnegative definite.

The symbol ~ following a random variable means distributed as. Other symbols used in statistics are: $\mathcal{E}$ for expected value, $\mathcal{V}$ and *cov* for variance and covariance. We denote the normal distribution by $\mathcal{N}$, the Wishart distribution by $\mathcal{W}$ and the matrix variate beta distribution by $\mathcal{B}$.

Finally we point out that (qed) is used to indicate the end of a proof.

CHAPTER II

EARLY RESULTS ON SCHUR COMPLEMENTS

We are concerned with matrices of the form

$$S = H - GE^{-1}F \tag{2.1}$$

Emilie V. Haynsworth (1968, p. 74) appears to be the first author to give the name *Schur complement* to (2.1). Following her, we refer to (2.1) as the *Schur complement of* E *in* A , where the square matrix

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}. \tag{2.2}$$

The notation

$$S = (A/E) = H - GE^{-1}F \tag{2.3}$$

is convenient.

§2.1 *Determinants.*

The first explicit mention of a matrix of the form (2.1) appears to be by Issai Schur (1875-1941), who used (2.1) to prove (1917, Hilfssatz, pp. 216-7):

THEOREM 2.1 (Schur, 1917). *Let the matrix* E *in* (2.2) *be nonsingular. Then*

$$\begin{vmatrix} E & F \\ G & H \end{vmatrix} = |E|.|H - GE^{-1}F| , \tag{2.4}$$

where $|\cdot|$ *denotes determinant.*

Proof (Banachiewicz, 1937, p. 51). We may write

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} E & 0 \\ G & I \end{pmatrix} \begin{pmatrix} I & E^{-1}F \\ 0 & H - GE^{-1}F \end{pmatrix} ; \tag{2.5}$$

taking determinants, we obtain (2.4). (qed)

Similarly, it may be shown that if the matrix H in (2.2) is non-singular then

$$\begin{vmatrix} E & F \\ G & H \end{vmatrix} = |H|.|E - FH^{-1}G| . \tag{2.6}$$

In the notation (2.3) we thus see that

$$|(A/E)| = |A|/|E| \quad \text{and} \quad |(A/H)| = |A|/|H|. \tag{2.7}$$

An immediate consequence of (2.4) and (2.6) is the following

COROLLARY 2.1. *Let* F *be* m×n *and* G n×m. *Then*

$$|I_m - FG| = |I_n - GF| . \tag{2.8}$$

Proof. Put $E = I_m$ and $H = I_n$ in (2.2). Then (2.8) follows at once using (2.4) and (2.6). (qed)

An alternate proof of (2.8), due to George Tiao, is given in the Appendix to the paper by Irwin Guttman (1971).

THEOREM 2.2 (Schur, 1917). *Consider the matrix* (2.2), *where* E, F, G, *and* H *are all* n×n , *and*

$$EG = GE . \tag{2.9}$$

Then

$$|A| = \begin{vmatrix} E & F \\ G & H \end{vmatrix} = |EH - GF| . \tag{2.10}$$

Proof. Suppose first that $|E| \neq 0$. Then (2.4) holds. Hence $|A| = |EH - EGE^{-1}F| = |EH - GEE^{-1}F|$, using (2.9), and so (2.10) follows. Now suppose that $|E| = 0$. Then $|E + xI| \neq 0$ for all $x \neq -ch(E)$, where $ch(\cdot)$ denotes characteristic root. Let

$$(2.11) \qquad B = \begin{pmatrix} E + xI & F \\ G & H \end{pmatrix} ;$$

then (2.9) $\Leftrightarrow$ $(E + xI)G = G(E + xI)$. Thus

$$(2.12) \qquad |B| = |EH + xH - GF| ;$$

as $x \to 0$, the matrix $B \to A$ and (2.12) becomes (2.10). (qed)

It is easily seen that (2.10) need not imply (2.9), since when $F = 0$ and E (or H) is nonsingular then (2.10) holds whether or not G is chosen to commute with E.

An immediate consequence of (2.10) is that A is nonsingular if and only if $EH - GF$ is nonsingular. In a paper by Herstein and Small (1975) it is shown that, for a fairly wide class $\mathcal{R}$ of rings, if the matrix (2.2) is over $\mathcal{R}$, where E, F, G and H are all $n \times n$ over $\mathcal{R}$ and (2.9) holds, then A is invertible if and only if $EH - GF$ is invertible. The authors state, as an example, that the result is true when $\mathcal{R}$ is a (right) artinian ring.

In the book by E. Bodewig (1956, 1959), the formula (2.4) is said (1956, p. 189; 1959, p. 218) to date from Frobenius (1849-1917), who obtained the following theorem (1908, p. 405):

THEOREM 2.3 (Frobenius, 1908). *Consider the matrix*

$$(2.13) \qquad A = \begin{pmatrix} E & \underset{\sim}{f} \\ \underset{\sim}{g}' & h \end{pmatrix} ,$$

where h *is a scalar,* $\underset{\sim}{f}$ *and* $\underset{\sim}{g}$ *are column vectors and* E *is a square matrix. Then*

$$(2.14) \qquad |A| = h|E| - \underset{\sim}{g}'(\mathit{adj}\,E)\underset{\sim}{f} ,$$

where adj *denotes adjugate matrix.*

Proof. Suppose first that $|E| \neq 0$. Using (2.4),

$$|A| = (h - \underset{\sim}{g}'E^{-1}\underset{\sim}{f})|E| \tag{2.15}$$

follows. Since

$$E^{-1} = (adj E)/|E|, \tag{2.16}$$

(2.14) follows. Now suppose that $|E| = 0$. Then $|E+xI| \neq 0$ for all $x \neq -ch(E)$. Let B be defined similarly to (2.11),

$$B = \begin{pmatrix} E+xI & \underset{\sim}{f} \\ \underset{\sim}{g}' & h \end{pmatrix}; \tag{2.17}$$

then

$$|B| = h|E+xI| - \underset{\sim}{g}'[adj(E+xI)]\underset{\sim}{f}. \tag{2.18}$$

As $x \to 0$, the matrix $B \to A$ and (2.18) becomes (2.14). (qed)

We notice that if $h \neq 0$ in (2.13) then using (2.6),

$$|A| = h|E - \underset{\sim}{f}\underset{\sim}{g}'/h| \tag{2.19a}$$

$$= |hE - \underset{\sim}{f}\underset{\sim}{g}'|/h^{n-1}, \tag{2.19b}$$

when E is $n \times n$. When $h = 1$ this simplifies further:

$$|A| = |E - \underset{\sim}{f}\underset{\sim}{g}'| \tag{2.20a}$$

$$= |E| - \underset{\sim}{g}'(adj E)\underset{\sim}{f} \tag{2.20b}$$

$$= (1 - \underset{\sim}{g}'E^{-1}\underset{\sim}{f})|E|, \tag{2.20c}$$

using (2.14) and (2.15). This leads at once to the following related result:

THEOREM 2.4 (Bodewig, 1956, p. 36; 1959, p. 42). *Let the matrix* E *be nonsingular and let the matrix* B *have rank* 1 . *Then*

$$(2.21)\qquad |E+B| = (1+\mathit{tr}E^{-1}B)|E| ,$$

where *tr* *denotes trace.*

Proof. Since B is of rank 1 , we may write

$$(2.22)\qquad B = \underset{\sim}{f}\underset{\sim}{g}'$$

as a full rank decomposition. Then applying (2.20c) gives

$$(2.23)\qquad |E+B| = (1+\mathit{tr}E^{-1}\underset{\sim}{f}\underset{\sim}{g}')|E| ,$$

and using (2.22), (2.21) follows. (qed)

When E is singular Theorem 2.4 reduces to

COROLLARY 2.2 (Bodewig, 1956, p. 36; 1959, p. 42). *Let the matrix* E *be singular and let the unit rank matrix* B *be defined by* (2.22). *Then*

$$(2.24)\qquad |E+B| = \mathit{tr}[(\mathit{adj}E)B] = \underset{\sim}{g}'(\mathit{adj}E)\underset{\sim}{f} .$$

We will see later, in §2.4, how Schur complements are related to Jacobi's theorem on the determinant of a minor (Jacobi, 1834; cf. Mirsky, 1955, p. 25).

§2.2 *Rank.*

Schur's determinant formula shows that the partitioned matrix

$$(2.24a)\qquad A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$$

is singular whenever the Schur complement $S = (A/E) = H - GE^{-1}F$ is singular

(E is assumed to be nonsingular). This result may be strengthened to show that *rank is additive* on the Schur complement, viz.

$$r(A) = r(E) + r(A/E) \tag{2.24b}$$

cf. Marsaglia and Styan (1974a, p. 291).

THEOREM 2.5 (Louis Guttman, 1946). *Let the matrix* E *in* (2.2) *be nonsingular. Then*

$$r(A) = r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(H - GE^{-1}F) = r(E) + r(A/E) , \tag{2.25}$$

where $r(\cdot)$ *denotes rank.*

Proof. Since E is nonsingular we may write, cf. (2.5),

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} I & 0 \\ GE^{-1} & I \end{pmatrix}\begin{pmatrix} E & 0 \\ 0 & H - GE^{-1}F \end{pmatrix}\begin{pmatrix} I & E^{-1}F \\ 0 & I \end{pmatrix}, \tag{2.26}$$

which yields (2.25). (qed)

Using the notation (2.3), we may write, cf. (2.7),

$$r(A/E) = r(A) - r(E) , \tag{2.27a}$$

and when H is nonsingular,

$$r(A/H) = r(A) - r(H). \tag{2.27b}$$

Theorem 2.5 readily yields

COROLLARY 2.3 (Louis Guttman, 1946). *If* A *and* E *in* (2.2) *are both nonsingular then the Schur complement* $(A/E) = H - GE^{-1}F$ *is also nonsingular.*

In the book by Wedderburn (1934, p. 69) a rank reduction procedure is presented which turns out to be a special case of (2.25). Let the matrix H be nonnull. Then there clearly exist vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ so that $\underset{\sim}{a}'H\underset{\sim}{b} \neq 0$. Consider the matrix

$$(2.28) \qquad A = \begin{pmatrix} \underset{\sim}{a}'H\underset{\sim}{b} & \underset{\sim}{a}'H \\ H\underset{\sim}{b} & H \end{pmatrix} = \begin{pmatrix} \underset{\sim}{a}' \\ I \end{pmatrix} H(\underset{\sim}{b}, I) .$$

Applying (2.25) yields

$$(2.29) \qquad r(A) = r(\underset{\sim}{a}'H\underset{\sim}{b}) + r(H - H\underset{\sim}{b}\underset{\sim}{a}'H/\underset{\sim}{a}'H\underset{\sim}{b}) = r(H)$$

using (2.28), and so we have proved:

THEOREM 2.6a (Wedderburn, 1934, p. 69). *If the matrix* $H \neq 0$, *then there exist vectors* $\underset{\sim}{a}$, $\underset{\sim}{b}$ *such that* $\underset{\sim}{a}'H\underset{\sim}{b} \neq 0$ *and*

$$(2.30) \qquad r(H - H\underset{\sim}{b}\underset{\sim}{a}'H/\underset{\sim}{a}'H\underset{\sim}{b}) = r(H) - 1.$$

Theorem 2.6a was extended in the book by Householder (1964) as an exercise.

THEOREM 2.6b (Householder, 1964, p. 33). *Let* $\underset{\sim}{u}$ *and* $\underset{\sim}{v}$ *be column vectors. Then for* $\lambda \neq 0$

$$(2.31) \qquad r(H - \underset{\sim}{u}\underset{\sim}{v}'/\lambda) < r(H)$$

if and only if there exist vectors $\underset{\sim}{a}$ *and* $\underset{\sim}{b}$ *such that* $\underset{\sim}{u} = H\underset{\sim}{b}$, $\underset{\sim}{v} = H'\underset{\sim}{a}$ *and* $\lambda = \underset{\sim}{a}'H\underset{\sim}{b} \neq 0$.

Proof. It suffices to prove the "only if" part. Consider the matrix

$$(2.32) \qquad A = \begin{pmatrix} \lambda & \underset{\sim}{v}' \\ \underset{\sim}{u} & H \end{pmatrix} .$$

Using (2.25) we obtain

$$(2.33) \qquad r(H) \le r(A) = 1 + r(H - \underset{\sim}{u}\underset{\sim}{v}'/\lambda) < 1 + r(H)$$

when (2.31) holds. Hence $r(H) = r(A)$ and so there exist vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ so that

$$(2.34a) \qquad (\lambda, \underset{\sim}{v}') = \underset{\sim}{a}'(\underset{\sim}{u}, H) ,$$

$$(2.34b) \qquad \begin{pmatrix} \lambda \\ \underset{\sim}{u} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{v}' \\ H \end{pmatrix} \underset{\sim}{b} ,$$

and thus $\underset{\sim}{u} = H\underset{\sim}{b}$, $\underset{\sim}{v}' = \underset{\sim}{a}'H$ and $\lambda = \underset{\sim}{a}'\underset{\sim}{u} = \underset{\sim}{v}'\underset{\sim}{b} = \underset{\sim}{a}'H\underset{\sim}{b}$. (qed)

Wedderburn (1934, p. 68) derived (2.30) using "the Lagrange method of reducing quadratic forms to a normal form", while Rao (1973, p. 69) refers to Theorem 2.6a as "Lagrange's Theorem"; for an extension see §4.6, Theorem 4.11.

§2.3 *Matrix inversion.*

Banachiewicz (1937, p. 54) appears to be the first author to study the inverse of a partitioned matrix. The formula, (2.37) below, is often attributed to Schur, who, it seems, did not proceed further than the determinant formulas (2.4) and (2.10). Banachiewicz (1937) obtained (2.37) in Cracovian notation, where matrices are multiplied column by column (see Appendix A for further details); he also rediscovered Theorem 2.1 and proved it using (2.5).

THEOREM 2.7 (Banachiewicz, 1937; Frazer, Duncan & Collar, 1938, p. 113). *Suppose that*

$$(2.35) \qquad A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$$

and E *are both nonsingular. Then the Schur complement*

(2.36) $$S = H - GE^{-1}F$$

is also nonsingular and

(2.37) $$A^{-1} = \begin{pmatrix} E^{-1}+E^{-1}FS^{-1}GE^{-1} & -E^{-1}FS^{-1} \\ -S^{-1}GE^{-1} & S^{-1} \end{pmatrix} = \begin{pmatrix} E^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} E^{-1}F \\ -I \end{pmatrix} S^{-1}(GE^{-1}, -I).$$

Proof. The first part is Corollary 2.3. To prove (2.37) we invert (2.5) giving

(2.38) $$A^{-1} = \begin{pmatrix} I & -E^{-1}FS^{-1} \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} E^{-1} & 0 \\ -GE^{-1} & I \end{pmatrix},$$

which yields (2.37). (qed)

COROLLARY 2.4 (Duncan, 1944). *Suppose that both* A , *given by* (2.35), *and* H *are nonsingular. Then the Schur complement*

(2.39) $$T = (A/H) = E - FH^{-1}G$$

is nonsingular, and

(2.40) $$A^{-1} = \begin{pmatrix} T^{-1} & -T^{-1}FH^{-1} \\ -H^{-1}GT^{-1} & H^{-1}+H^{-1}GT^{-1}FH^{-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & H^{-1} \end{pmatrix} + \begin{pmatrix} -I \\ H^{-1}G \end{pmatrix} T^{-1}(-I, FH^{-1}).$$

Hotelling (1943), moreover, noted that if A, E and H are all nonsingular then

(2.41) $$A^{-1} = \begin{pmatrix} T^{-1} & -E^{-1}FS^{-1} \\ -H^{-1}GT^{-1} & S^{-1} \end{pmatrix},$$

which involves four inverses, while (2.37) and (2.40) each require only two (cf. Waugh, 1945). Duncan (1944) observed that (2.37) = (2.40), so that

(2.42) $$(E - FH^{-1}G)^{-1} = E^{-1}+E^{-1}F(H - GE^{-1}F)^{-1}GE^{-1},$$

which Woodbury (1950) reestablished.

THEOREM 2.8 (Woodbury, 1950). *Let*

$$A = \begin{pmatrix} E & -FH \\ HG & H \end{pmatrix}, \tag{2.43}$$

and let E *be nonsingular. If either* A *or* (A/E) *is nonsingular then* A, (A/E), H *and* (A/H) *are all nonsingular. Moreover,*

$$(E+FHG)^{-1} = E^{-1} - E^{-1}FH(H+HGE^{-1}FH)^{-1}HGE^{-1} \tag{2.44a}$$

$$= E^{-1} - E^{-1}F(H^{-1}+GE^{-1}F)^{-1}GE^{-1} . \tag{2.44b}$$

Proof. Since E is nonsingular we may write

$$r(A) = r(E) + r(A/E) = r(E) + r(H+HGE^{-1}FH) \tag{2.45}$$

using (2.25). Assume H is of rank h . We may write

$$H = KL' \tag{2.46}$$

as a full rank decomposition, where K and L have full column rank h . The matrix A may now be written as

$$A = \begin{pmatrix} E & -FKL' \\ KL'G & KL' \end{pmatrix} \tag{2.47}$$

$$= \begin{pmatrix} I & 0 \\ 0 & K \end{pmatrix}\begin{pmatrix} E & -FK \\ L'G & I_h \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & L' \end{pmatrix} . \tag{2.48}$$

Then

$$r(A) = r\begin{pmatrix} E & -FK \\ L'G & I_h \end{pmatrix} \tag{2.49}$$

$$= r(I_h) + r(E+FKL'G) \tag{2.50}$$

$$= r(H) + r(E+FHG) . \tag{2.51}$$

It is easily seen that (2.45) and (2.51) imply that A, (A/E), H and $E + FHG = (A/H)$ are all nonsingular when A, or (A/E), is nonsingular. Hence, using (2.42), (2.44) follows. (qed)

Woodbury (1950) implied that (2.44a) might hold if H is singular. However this cannot be for if E and its Schur complement $H + HGE^{-1}FH = (A/E)$ are both nonsingular then by (2.45) A must be nonsingular. Thus, using (2.51), the fact that both A and $E + FHG$ are nonsingular implies H nonsingular.

From Theorem 2.8 readily follows:

COROLLARY 2.5 (Woodbury, 1950). *Suppose that*

$$A = \begin{pmatrix} E & -h\underset{\sim}{f} \\ h\underset{\sim}{g}' & h \end{pmatrix} \tag{2.52}$$

and E *are both nonsingular and* $h \neq 0$. *Then the Schur complement*

$$(A/E) = h(1 + h\underset{\sim}{g}'E^{-1}\underset{\sim}{f}) \neq 0, \tag{2.53}$$

and the Schur complement

$$(A/h) = E + h\underset{\sim}{f}\underset{\sim}{g}' \tag{2.54}$$

is nonsingular; and

$$(E + h\underset{\sim}{f}\underset{\sim}{g}')^{-1} = E^{-1} - hE^{-1}\underset{\sim}{f}\underset{\sim}{g}'E^{-1}/(1 + h\underset{\sim}{g}'E^{-1}\underset{\sim}{f}) . \tag{2.55}$$

Woodbury (1950) observed that John W. Tukey independently found that

$$(I + h\underset{\sim}{f}\underset{\sim}{g}')^{-1} = I - h\underset{\sim}{f}\underset{\sim}{g}'/(1 + h\underset{\sim}{g}'\underset{\sim}{f}), \tag{2.56}$$

which follows immediately by substituting $E = I$ in (2.55).

COROLLARY 2.6 (Bartlett, 1951). *Suppose that both* A *, given by* (2.52), *and* E *are nonsingular. Let* h = 1. *Then the Schur complement*

$$(A/E) = 1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{f} \neq 0 \;, \tag{2.57}$$

and the Schur complement

$$(A/h) = (A/1) = E + \underset{\sim}{f}\underset{\sim}{g}' \tag{2.58}$$

is nonsingular; and

$$(E + \underset{\sim}{f}\underset{\sim}{g}')^{-1} = E^{-1} - E^{-1}\underset{\sim}{f}\underset{\sim}{g}'E^{-1}/(1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{f}). \tag{2.59}$$

Sherman and Morrison (1949, 1950) obtained the following results, which are all special cases of Corollary 2.5:

$$(E + h\underset{\sim}{e}_i\underset{\sim}{g}')^{-1} = E^{-1} - hE^{-1}\underset{\sim}{e}_i\underset{\sim}{g}'E^{-1}/(1 + h\underset{\sim}{g}'E^{-1}\underset{\sim}{e}_i), \quad h\underset{\sim}{g}'E^{-1}\underset{\sim}{e}_i \neq -1, \tag{2.60}$$

$$(E + h\underset{\sim}{f}\underset{\sim}{e}_j')^{-1} = E^{-1} - hE^{-1}\underset{\sim}{f}\underset{\sim}{e}_j'E^{-1}/(1 + h\underset{\sim}{e}_j'E^{-1}\underset{\sim}{f}), \quad h\underset{\sim}{e}_j'E^{-1}\underset{\sim}{f} \neq -1, \tag{2.61}$$

$$(E + h\underset{\sim}{e}_i\underset{\sim}{e}_j')^{-1} = E^{-1} - hE^{-1}\underset{\sim}{e}_i\underset{\sim}{e}_j'E^{-1}/(1 + h\underset{\sim}{e}_j'E^{-1}\underset{\sim}{e}_i), \quad h\underset{\sim}{e}_j'E^{-1}\underset{\sim}{e}_i \neq -1, \tag{2.62}$$

where $\underset{\sim}{e}_k$ denotes a column vector with all elements zero except for unity in the k*th* position.

The formula (2.62) shows what happens to the inverse E^{-1} when the scalar h is added to the (i,j)*th* element of E ; the modified matrix remains nonsingular $\Leftrightarrow h\underset{\sim}{e}_j'E^{-1}\underset{\sim}{e}_i \neq -1$. If the row vector $\underset{\sim}{g}'$ is added to the i*th* row of E then the modified matrix remains nonsingular $\Leftrightarrow \underset{\sim}{g}'E^{-1}\underset{\sim}{e}_i \neq -1$, and then, cf. (2.60),

$$(E + \underset{\sim}{e}_i\underset{\sim}{g}')^{-1} = E^{-1} - E^{-1}\underset{\sim}{e}_i\underset{\sim}{g}'E^{-1}/(1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{e}_i). \tag{2.63}$$

Similarly if the column vector $\underset{\sim}{f}$ is added to the jth column of E then the modified matrix remains nonsingular $\Leftrightarrow \underset{\sim}{e}_j' E^{-1}\underset{\sim}{f} \neq -1$, and then, cf. (2.61),

$$(2.64) \qquad (E+\underset{\sim}{f}\underset{\sim}{e}_j')^{-1} = E^{-1} - E^{-1}\underset{\sim}{f}\underset{\sim}{e}_j' E^{-1}/(1+\underset{\sim}{e}_j'E^{-1}\underset{\sim}{f}).$$

§2.4 *Theorems of Cauchy (1812) and Jacobi (1834).*

It is well known (cf. e.g., Aitken, 1939, p. 53) that for any square matrix A

$$(2.65) \qquad A(adjA) = (adjA)A = |A|I,$$

and so if A is n×n , taking determinants of (2.65) yields

$$(2.66) \qquad |adjA| = |A|^{n-1},$$

which is due to Cauchy (1812). This result was extended by Jacobi (1834) as follows (see also Aitken, 1939, p. 103):

THEOREM 2.9 (Jacobi, 1834). *Consider the* n×n *matrix*

$$(2.67) \qquad A = \begin{pmatrix} E & F \\ G & H \end{pmatrix},$$

where E *is* m×m. *Let*

$$(2.68) \qquad A^* = adjA = \begin{pmatrix} E^* & F^* \\ G^* & H^* \end{pmatrix},$$

where E^* *is* m×m. *Then*

$$(2.69) \qquad |H^*| = |A|^{n-m-1}\,|E|; \qquad m = 0,1,\ldots,n-1.$$

Proof. When $m = 0$ the matrix E disappears and (2.69) reduces to (2.66). When $m = n-1$ (2.69) is trivially true. So assume $1 \le m \le n-2$. If $|A| = 0$ then $r(H^*) \le r(adjA) \le 1$ in view of (2.65), and with $n-m \ge 2$ it follows that $|H^*| = 0$ and so (2.69) holds. Now assume $|A| \ne 0$. Then (2.65) implies that

$$adjA = |A|A^{-1} . \tag{2.70}$$

Suppose first that $|E| \ne 0$. We may write

$$H^* = |A|(H - GE^{-1}F)^{-1} , \tag{2.71}$$

using (2.37). Taking determinants, we obtain

$$|H^*| = |A|^{n-m}/|H - GE^{-1}F| = |A|^{n-m}/|(A/E)| \tag{2.72a}$$

$$= |A|^{n-m}/(|A|/|E|), \tag{2.72b}$$

using (2.7), and so (2.69) follows. It remains only to consider the case when $|E| = 0$. Suppose then that $|H^*| \ne 0$. Using (2.27b) shows that since $|A| \ne 0$,

$$n = r(adjA) = r(H^*) + r(adjA/H^*) = n-m + r(adjA/H^*), \tag{2.73}$$

and so $(adjA/H^*)$ is nonsingular. The inverse of the Schur complement of H^* in $adjA$ is $E/|A|$, cf. (2.40). Hence $|E| \ne 0$, a contradiction. Thus $|E| = 0$ implies $|H^*| = 0$ and so (2.69) holds. (qed)

Similarly, it may be shown that

$$|E^*| = |A|^m|H|; \qquad m = 0,1,\dots,n-1. \tag{2.74}$$

CHAPTER III

RECENT RESULTS ON SCHUR COMPLEMENTS

In the previous chapter we studied many early results pertaining to Schur complements of a nonsingular matrix.

We now proceed to discuss various properties of the Schur complement of a nonsingular matrix which have appeared more recently. It seems (cf. Cottle, 1974) that from 1952 through 1967 no research papers with results on Schur complements were published. In fact the term "Schur complement" appears to originate with Haynsworth (1968) in a study of the *inertia* of a partitioned real symmetric matrix.

§3.1 *Inertia.*

The inertia of a symmetric matrix A is the ordered triple

$$\text{(3.1)} \qquad In\,A = (\pi, \nu, \delta),$$

where π is the number of positive characteristic roots of A, ν the number of negative and δ the number of zero roots of A. Thus $\pi + \nu = r(A)$, the rank of A, and δ is the nullity of A. Sylvester (1852) proved that

$$\text{(3.2)} \qquad In\,A = In\,CAC'$$

for every nonsingular matrix C, cf. Marcus and Minc (1964, p. 83). The equation (3.2) is called *Sylvester's law of inertia*.

THEOREM 3.1 (Haynsworth, 1968). *Consider the* $(m+n) \times (m+n)$ *symmetric matrix*

$$\text{(3.3)} \qquad A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix},$$

where E *is* $m \times m$ *nonsingular. Then*

$$\text{(3.4)} \qquad In\,A = In\,E + In(A/E) .$$

Proof. We may write, cf. (2.26),

$$B = \begin{pmatrix} I_m & 0 \\ -F'E^{-1} & I_n \end{pmatrix} \begin{pmatrix} E & F \\ F' & H \end{pmatrix} \begin{pmatrix} I_m & -E^{-1}F \\ 0 & I_n \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & (A/E) \end{pmatrix} \tag{3.5}$$

Using (3.2),

$$\mathit{In}B = \mathit{In}A \ , \tag{3.6}$$

and since the characteristic roots of B are those of E and of (A/E), (3.4) follows. (qed)

A matrix of the form X'X is said to be *Gramian* or *nonnegative definite* (nnd). If X'X is singular, then it will be called *positive semidefinite* (psd). If X'X is nonsingular, then it is called *positive definite* (pd). Some authors (e.g., Haynsworth, 1968) use positive semidefinite where we use nonnegative definite. In this thesis, positive definite, positive semidefinite and nonnegative definite matrices are always symmetric. We note that the symmetric matrix A is nonnegative definite $\Leftrightarrow \nu = 0$, positive definite $\Leftrightarrow \nu = \delta = 0$, positive semidefinite $\Leftrightarrow \{\nu = 0$ and $\delta \geq 1\}$.

COROLLARY 3.1 (Haynsworth, 1968). *Consider the symmetric matrix*

$$A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix} , \tag{3.7}$$

and let E *be positive definite. Then*

$$A \text{ is nnd} \Leftrightarrow (A/E) \text{ is nnd} \tag{3.8a}$$

$$A \text{ is psd} \Leftrightarrow (A/E) \text{ is psd} \tag{3.8b}$$

$$A \text{ is pd} \Leftrightarrow (A/E) \text{ is pd.} \tag{3.8c}$$

When E and H are both $n \times n$ nonsingular then the difference between them has the same rank as the difference between their inverses, for

$$E - H = -H(E^{-1} - H^{-1})E. \tag{3.9}$$

We now show that when E and H are both $n \times n$ positive definite and $E - H$ is positive (semi)definite then $H^{-1} - E^{-1}$ is positive (semi-) definite also. See also Theorem 4.13.

THEOREM 3.2. *Let* E *and* H *both be* $n \times n$ *positive definite, and suppose that* $E - H$ *is positive (semi)definite. Then* $H^{-1} - E^{-1}$ *is positive (semi)definite, and*

$$r(E - H) = r(H^{-1} - E^{-1}). \tag{3.10}$$

Proof. Consider the matrix

$$A = \begin{pmatrix} E & I_n \\ I_n & H^{-1} \end{pmatrix}, \tag{3.11}$$

where E, H^{-1} are pd. Since $E - H = (A/H^{-1})$ is nnd and H^{-1} is pd, A is nnd by (3.8a). Also, since E is pd, $(A/E) = H^{-1} - E^{-1}$ is nnd. (qed)

Haynsworth (1968) extended Theorem 3.1 by considering the inertia of partitioned Schur complements. We begin by partitioning the $(m+n) \times (m+n)$ symmetric matrix A as in (3.3), where E is $m \times m$ nonsingular. We then compute the Schur complement $S = (A/E)$ and obtain (3.4). We partition the Schur complement

$$(A/E) = S = \begin{pmatrix} E_1 & F_1 \\ F_1' & H_1 \end{pmatrix}, \tag{3.12}$$

where E_1 is $m_1 \times m_1$, nonsingular, and compute the Schur complement $S_1 = (S/E_1)$.

We obtain

(3.13) $$InA = InE + InE_1 + InS_1 \ .$$

We partition

(3.14) $$S_1 = \begin{pmatrix} E_2 & F_2 \\ F_2' & H_2 \end{pmatrix},$$

where E_2 is $m_2 \times m_2$, nonsingular. We compute $S_2 = (S_1/E_2)$ and repeat the procedure performed with S_1. We obtain

(3.14a) $$InA = InE + InE_1 + InE_2 + InS_2 \ .$$

We may continue this process by defining E_{i+1} as the top left-hand $m_{i+1} \times m_{i+1}$ nonsingular submatrix of the $(n - \sum_{j=1}^{i} m_j) \times (n - \sum_{j=1}^{i} m_j)$ Schur complement $S_i = (S_{i-1}/E_i)$. The process stops as soon as a Schur complement, S_k, say, is a scalar or has no top left nonsingulat submatrix. Then

(3.15) $$InA = InE + \sum_{j=1}^{k} InE_j + InS_k \ .$$

§3.2 *The quotient property and related determinant inequalities.*

Consider the matrix

(3.16) $$A = \left(\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right) = \left(\begin{array}{cc|c} K & L & F_1 \\ M & N & F_2 \\ \hline G_1 & G_2 & H \end{array}\right),$$

where E and K are nonsingular. Then the Schur complement (E/K) is a nonsingular leading principal submatrix of the Schur complement (A/K); Crabtree & Haynsworth (1969) and Ostrowski (1971) proved that

(3.17) $$(A/E) = ((A/K)/(E/K)),$$

which they called the *quotient property*.

We note that the parallel relationship

$$AE^{-1} = (AK^{-1})(EK^{-1})^{-1} \tag{3.18}$$

also holds.

THEOREM 3.3 (Crabtree & Haynsworth, 1969). *Consider the matrix* (3.16), *where both* E *and* K *are nonsingular. Then the Schur complement* (E/K) *is a nonsingular leading principal submatrix of the Schur complement* (A/K). *Moreover,* (3.17) *holds.*

Proof. Since

$$(A/K) = \begin{pmatrix} N & F_2 \\ G_2 & H \end{pmatrix} - \begin{pmatrix} M \\ G_1 \end{pmatrix} K^{-1}(L, F_1) \tag{3.19}$$

$$= \begin{pmatrix} N - MK^{-1}L & F_2 - MK^{-1}F_1 \\ G_2 - G_1K^{-1}L & H - G_1K^{-1}F_1 \end{pmatrix}, \tag{3.20}$$

$N - MK^{-1}L = (E/K)$ is a leading principal submatrix of (A/K). Since $|E| \neq 0$ and $|K| \neq 0$, it follows, using (2.7), that

$$|(E/K)| = |E|/|K| \neq 0 , \tag{3.21}$$

and so (E/K) is nonsingular. Also,

$$((A/K)/(E/K)) = H - G_1K^{-1}F_1 - (G_2 - G_1K^{-1}L)(E/K)^{-1}(F_2 - MK^{-1}F_1) \tag{3.22}$$

$$= H - (G_1, G_2)\begin{pmatrix} K^{-1} + K^{-1}L(E/K)^{-1}MK^{-1} & -K^{-1}L(E/K)^{-1} \\ -(E/K)^{-1}MK^{-1} & (E/K)^{-1} \end{pmatrix}\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \tag{3.23}$$

$$= H - GE^{-1}F = (A/E), \tag{3.24}$$

using (2.37). (qed)

Haynsworth (1970b) extended Theorem 3.2 by showing that if

$$A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix}, \tag{3.25}$$

cf. (3.7), and

$$B = \begin{pmatrix} K & L \\ L' & N \end{pmatrix} \tag{3.26}$$

are both $(m+n) \times (m+n)$ nonnegative definite matrices, where E and K are both $m \times m$ positive definite, then

$$((A+B)/(E+K)) - (A/E) - (B/K) \tag{3.27}$$

is nonnegative definite. To prove this we use the following:

LEMMA 3.1 (Haynsworth, 1970b). *Let* E *and* K *both be* $m \times m$ *positive definite. Then if* F *and* L *are arbitrary* $m \times n$ *matrices*,

$$F'E^{-1}F + L'K^{-1}L - (F+L)'(E+K)^{-1}(F+L) \tag{3.28}$$

is nonnegative definite with the same rank as

$$F - EK^{-1}L \; . \tag{3.29}$$

Proof. We may rewrite (3.28) as follows:

$$(F', L') \begin{pmatrix} E^{-1} - (E+K)^{-1} & -(E+K)^{-1} \\ -(E+K)^{-1} & K^{-1} - (E+K)^{-1} \end{pmatrix} \begin{pmatrix} F \\ L \end{pmatrix} . \tag{3.30}$$

Since E and K are pd so is $E+EK^{-1}E$. Applying (2.42) with $F = G = E$ and $H = -K$ then yields

$$(E+EK^{-1}E)^{-1} = E^{-1} - (E+K)^{-1} \tag{3.31}$$

positive definite, and so (3.30) may be written

$$(F', L') \begin{pmatrix} I \\ -K^{-1}E \end{pmatrix} (E+EK^{-1}E)^{-1}(I, -EK^{-1}) \begin{pmatrix} F \\ L \end{pmatrix}, \tag{3.32}$$

which is nnd with rank equal to the rank of $F - EK^{-1}L$. (qed)

THEOREM 3.4 (Haynsworth, 1970b). *Let*

$$A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix} \tag{3.33}$$

and

$$B = \begin{pmatrix} K & L \\ L' & N \end{pmatrix} \tag{3.34}$$

both be symmetric $(m+n) \times (m+n)$ *matrices, where* E *and* K *are both* $m \times m$. *If* A *and* B *are nonnegative definite and* E *and* K *are positive definite then*

$$((A+B)/(E+K)) - (A/E) - (B/K) \tag{3.35}$$

is nonnegative definite.

Proof. Since the sum of any two positive (nonnegative) definite matrices is positive (nonnegative) definite, $E+K$ is pd and $A+B$ is nnd. We have

$$A+B = \begin{pmatrix} E+K & F+L \\ F'+L' & H+N \end{pmatrix}, \tag{3.36}$$

so that the Schur complement

$$(3.37) \qquad ((A+B)/(E+K)) = H+N - (F+L)'(E+K)^{-1}(F+L) ,$$

is nnd. Hence

$$(3.38) \qquad ((A+B)/(E+K)) - (A/E) - (B/K) = F'E^{-1}F + L'K^{-1}L - (F+L)'(E+K)^{-1}(F+L),$$

which is nnd from Lemma 3.1. (qed)

Consider the $(m+n) \times (m+n)$ matrices A and B defined by (3.33) and (3.34), where E and K are both $m \times m$. Haynsworth (1970b) proved that if A and B are nonnegative definite and E and K are positive definite then

$$(3.39) \qquad |(A+B)/(E+K)| = \frac{|A+B|}{|E+K|} \geq \frac{|A|}{|E|} + \frac{|B|}{|K|} .$$

To prove (3.39), we use the following:

LEMMA 3.2. *Let* A *and* B *both be* $n \times n$ *nonnegative definite matrices. Then*

$$(3.40) \qquad |A+B| \geq |A| + |B| .$$

Proof. Suppose first that both $|A| = 0$ and $|B| = 0$. Then (3.40) clearly holds. Now suppose that $|A| \neq 0$. Then

$$(3.41) \qquad |A+B| = |A|\,|I + A^{-1}B|$$

$$(3.42) \qquad = |A| . \prod_{i=1}^{n} (1 + ch_i A^{-1}B)$$

$$(3.43) \qquad \geq |A| . [1 + \prod_{i=1}^{n} ch_i A^{-1}B]$$

since the characteristic roots $ch_i A^{-1}B \geq 0$. Using the fact that

$$(3.44) \qquad |A| . [1 + \prod_{i=1}^{n} ch_i A^{-1}B] = |A| . [1 + |A^{-1}B|] = |A| + |B| ,$$

(3.40) follows at once. (qed)

LEMMA 3.3. *Let* A *and* B *both be* $n\times n$ *positive definite matrices and let* A-B *be nonnegative definite. Suppose further that for* $i=1,\dots,n$, E_i *and* K_i *are the* $i\times i$ *leading principal submatrices of* A *and* B, *respectively. Then* $E_i - K_i$ *is nonnegative definite and*

$$(3.45)\qquad |E_i| \geq |K_i| \;,\; i=1,\dots,n.$$

COROLLARY 3.2 (Haynsworth, 1970b). *Let* A *and* B *both be* $(m+n)\times(m+n)$ *matrices defined by* (3.33) *and* (3.34), *where* E *and* K *are both* $m\times m$. *If* A *and* B *are nonnegative definite and* E *and* K *are positive definite then* (3.39) *holds.*

Proof. From Theorems 2.1 and 3.4 it follows that

$$(3.46)\qquad \frac{|A+B|}{|E+K|} = |((A+B)/(E+K))| \geq |(A/E)+(B/K)|$$

$$\geq |(A/E)| + |(B/K)| = \frac{|A|}{|E|} + \frac{|B|}{|K|} ,$$

using (3.35), (3.45) and (3.40). Hence (3.39) follows. (qed)

We may extend (3.39) using Lemma 3.3.

THEOREM 3.5 (Haynsworth, 1970b). *Let* A *and* B *both be* $n\times n$ *nonnegative definite matrices. Suppose further that* E_i *and* K_i, $i=1,\dots,n-1$, *are the* $i\times i$ *principal submatrices in the upper left corners of the matrices* A *and* B *respectively. If* $E_1,\dots,E_{n-1},K_1,\dots,K_{n-1}$ *are all positive definite then*

$$(3.47)\qquad |A+B| \geq |A|.\left(1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right)+|B|.\left(1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right) .$$

Proof. We will use induction on n. For $n=2$,

$$|A+B| = |E_1+K_1| . |(A+B)/(E_1+K_1)| \tag{3.48}$$

using (2.4). But

$$|E_1+K_1| \geq |E_1| + |K_1| \tag{3.49}$$

by (3.40), and

$$|(A+B)/(E_1+K_1)| \geq \frac{|A|}{|E_1|} + \frac{|B|}{|K_1|} \tag{3.50}$$

by (3.39). Hence, (3.47) holds for $n=2$. Now assume that (3.47) holds for A and B $n \times n$. If A_1 and B_1 are $(n+1) \times (n+1)$ nonnegative definite matrices, and $A = E_n$ and $B = K_n$ are $n \times n$ positive definite submatrices of A_1 and B_1, respectively, then

$$|A_1+B_1| = |E_n+K_n| . |(A_1+B_1)/(E_n+K_n)| \tag{3.51}$$

using (2.4). But, by the inductive assumption,

$$|E_n+K_n| \geq |E_n| . \left(1 + \sum_{i=1}^{n-1} \frac{|K_i|}{|E_i|}\right) + |K_n| . \left(1 + \sum_{i=1}^{n-1} \frac{|E_i|}{|K_i|}\right) \tag{3.52}$$

and by (3.39),

$$|(A_1+B_1)/(E_n+K_n)| \geq \frac{|A_1|}{|E_n|} + \frac{|B_1|}{|K_n|} . \tag{3.53}$$

Hence

$$(3.54)\quad |A_1+B_1| \geq \left\{ |E_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + |K_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right]\right\}\cdot\left\{\frac{|A_1|}{|E_n|}+\frac{|B_1|}{|K_n|}\right\}$$

$$= |A_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + \frac{|E_n|}{|K_n|}\cdot|B_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] +$$

$$\frac{|K_n|}{|E_n|}\cdot|A_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right] + |B_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right]$$

$$\geq |A_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + \frac{|E_n||B_1|}{|K_n|} + \frac{|K_n||A_1|}{|E_n|} + |B_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right]$$

$$= |A_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|K_i|}{|E_i|}\right] + |B_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|E_i|}{|K_i|}\right] .$$

Thus (3.47) holds for $(n+1)\times(n+1)$ matrices A and B and the induction proof is complete. (qed)

In the paper by Haynsworth (1970b), formula (3.47) was established with both A and B positive definite.

COROLLARY 3.3 (Haynsworth, 1970b). *If* A, B *and* $A-B$ *are* $n\times n$ *positive definite matrices, then*

$$(3.55)\qquad |A+B| > |A| + n|B| .$$

Proof. By Theorem 3.5,

$$(3.56)\quad |A+B| \geq |A|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + |B|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right].$$

But, since A−B is pd, $|E_i| > |K_i|$, cf. (3.45). Hence

$$|A+B| > |A|.\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + n|B| > |A| + n|B|. \tag{3.57}$$

(qed)

Hartfiel (1973) has improved (3.47) using the following result: If $f(x) = ax + bx^{-1}$, where $a,b > 0$, then $\min_{0<x<\infty} f(x)$ is achieved at $x = (b/a)^{\frac{1}{2}}$ and so

$$\min_{0<x<\infty} f(x) = f[(b/a)^{\frac{1}{2}}] = 2(ab)^{\frac{1}{2}} . \tag{3.58}$$

THEOREM 3.6 (Hartfiel, 1973). *Let* A *and* B *both be* $n\times n$ *nonnegative definite matrices. Suppose further that* E_i *and* K_i, $i=1,\dots,n-1$, *are the* $i\times i$ *principal submatrices in the upper left corners of the matrices* A *and* B, *respectively. If* $E_1,\dots,E_{n-1},K_1,\dots,K_{n-1}$ *are all positive definite then*

$$|A+B| \geq |A|.\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + |B|.\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right] + (2^n - 2n)(|A|.|B|)^{\frac{1}{2}} . \tag{3.59}$$

Proof. We will use induction on n. For $n=2$, (3.59) reduces to (3.47), and so (3.59) holds for $n=2$. Now assume that (3.59) holds for A and B $n\times n$. If A_1 and B_1 are $(n+1)\times(n+1)$ nonnegative definite, and $A = E_n$ and $B = K_n$ are $n\times n$ positive definite submatrices of A and B, respectively, then, cf. (3.51),

$$|A_1 + B_1| = |(E_n + K_n)|.|(A_1 + B_1)/(E_n + K_n)| . \tag{3.60}$$

But, by the inductive assumption,

$$|E_n + K_n| \geq |E_n|.\left[1+\sum_{i=1}^{n-1}\frac{|K_i|}{|E_i|}\right] + |K_n|.\left[1+\sum_{i=1}^{n-1}\frac{|E_i|}{|K_i|}\right] + (2^n - 2n)(|E_n|.|K_n|)^{\frac{1}{2}} , \tag{3.61}$$

and using (3.53) it follows that

$$(3.62)\qquad |A_1+B_1| \geq \left\{ |E_n| \cdot \left[1+\sum_{i=1}^{n-1} \frac{|K_i|}{|E_i|} \right] + |K_n| \cdot \left[1+\sum_{i=1}^{n-1} \frac{|E_i|}{|K_i|} \right] + (2^n-2n)(|E_n| \cdot |K_n|)^{\frac{1}{2}} \right\} \left[\frac{|A_1|}{|E_n|} + \frac{|B_1|}{|K_n|} \right]$$

$$(3.63)\qquad \geq |A_1| \cdot \left[1+\sum_{i=1}^{n} \frac{|K_i|}{|E_i|} \right] + |B_1| \cdot \left[1+\sum_{i=1}^{n} \frac{|E_i|}{|K_i|} \right] + \sum_{i=1}^{n-1} \left[\frac{|K_i|}{|E_i|} \cdot \frac{|E_n|}{|K_n|} \cdot |B_1| + \frac{|E_i|}{|K_i|} \cdot \frac{|K_n|}{|E_n|} \cdot |A_1| \right] + (2^n-2n)\,(|E_n| \cdot |K_n|)^{\frac{1}{2}} \left[\frac{|A_1|}{|E_n|} + \frac{|B_1|}{|K_n|} \right],$$

using (3.54). From (3.58) we see that

$$(3.64)\qquad \sum_{i=1}^{n-1} \left[\frac{|K_i| \cdot |E_n|}{|E_i| \cdot |K_n|} \cdot |B_1| + \frac{|E_i| \cdot |K_n|}{|K_i| \cdot |E_n|} \cdot |A_1| \right] \geq 2(n-1)(|A_1| \cdot |B_1|)^{\frac{1}{2}},$$

while using the arithmetic mean/geometric mean inequality, we have that

$$(3.65)\qquad \frac{|A_1|}{|E_n|} + \frac{|B_1|}{|K_n|} \geq 2 \left[\frac{|A_1| \cdot |B_1|}{|E_n| \cdot |K_n|} \right]^{\frac{1}{2}}.$$

Substituting (3.64) and (3.65) into (3.63) yields the lower bound:

$$|A_1| \cdot \left[1+\sum_{i=1}^{n} \frac{|K_i|}{|E_i|} \right] + |B_1| \cdot \left[1+\sum_{i=1}^{n} \frac{|E_i|}{|K_i|} \right] + (2^{n+1}-2(n+1))(|A_1| \cdot |B_1|)^{\frac{1}{2}},$$

as desired, since $2(n-1) + 2(2^n - 2n) = 2^{n+1} - 2(n+1)$. Thus (3.59) holds for $(n+1) \times (n+1)$ matrices A and B and the induction proof is complete. (qed)

In the paper by Hartfiel (1973), formula (3.59) was established when both A and B are positive definite.

COROLLARY 3.4 (Hartfiel, 1973). *If* A *and* B *are* $n\times n$ *positive definite matrices then*

$$|A+B| \geq |A| + |B| + (2^n-2)(|A|.|B|)^{\frac{1}{2}} . \tag{3.66}$$

Proof. By Theorem 3.6,

$$|A+B| \geq |A| + |B| + \sum_{i=1}^{n-1}\left[|A| . \frac{|K_i|}{|E_i|} + |B| . \frac{|E_i|}{|K_i|} \right] + (2^n - 2n)(|A|.|B|)^{\frac{1}{2}} \tag{3.67}$$

$$\geq |A| + |B| + 2(n-1)(|A|.|B|)^{\frac{1}{2}} + (2^n-2n)(|A|.|B|)^{\frac{1}{2}} ,$$

using (3.58), which

$$= |A| + |B| + (2^n-2)(|A|.|B|)^{\frac{1}{2}} . \quad \text{(qed)}$$

Corollary 3.4 allows us also to extend Corollary 3.3.

COROLLARY 3.5. *If* A, B *and* A-B *are all* $n\times n$ *positive definite then*

$$|A+B| > |A| + (2^n-1)|B| . \tag{3.68}$$

Proof. Since A-B is pd, $|A|^{\frac{1}{2}} > |B|^{\frac{1}{2}}$. Hence (3.66) implies (3.68). (qed)

§3.3 *Characteristic roots.*

If the $n \times m$ matrix X of linearly independent characteristic vectors corresponding to m roots of an $n \times n$ matrix A is available, then the remaining $n-m$ roots of A are the roots of a Schur complement in the matrix formed from A by replacing m of its columns with X. Thus let

$$AX = XD , \tag{3.69a}$$

where D is diagonal $m \times m$ and X has full column rank m. More generally, consider, cf. Haynsworth (1970a),

$$AX = XB , \tag{3.69b}$$

where B is an arbitrary $m \times m$ matrix. Goddard & Schneider (1955) call such an X a *commutator* and showed that m characteristic roots of A are characteristic roots of B.

THEOREM 3.7 (Haynsworth, 1970a). *Let the* $n \times n$ *matrix*

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} . \tag{3.70}$$

Suppose further that B *is an* $m \times m$ *matrix and that* X *is an* $n \times m$ *matrix of rank* m, *such that*

$$AX = XB \tag{3.71}$$

and

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \tag{3.72}$$

where X_1 *is an* $m \times m$ *nonsingular matrix. Then* m *characteristic roots of*

A *are characteristic roots of* B *and the remaining* $n-m$ *characteristic roots are characteristic roots of* (C/X_1) *where*

$$(3.73) \qquad C = \begin{pmatrix} X_1 & F \\ X_2 & H \end{pmatrix}.$$

Proof. Since X_1 is nonsingular, the $n \times n$ matrix

$$(3.74) \qquad J = \begin{pmatrix} X_1 & 0 \\ X_2 & I_{n-m} \end{pmatrix}$$

is nonsingular. Hence, using (3.71),

$$(3.75) \qquad J^{-1}AJ = \begin{pmatrix} B & X_1^{-1}F \\ 0 & H - X_2X_1^{-1}F \end{pmatrix}$$

has the same characteristic roots as A, and the proof is complete. (qed)

COROLLARY 3.6 (Haynsworth, 1970a). *Let the matrix* A *have* m *linearly independent (column) characteristic vectors corresponding to the characteristic roots* $\lambda_1,\ldots,\lambda_m$ *(not necessarily distinct). Suppose further that the columns of the* $n \times m$ *matrix* X *are the characteristic vectors and that* X *may be partitioned as in* (3.72). *Then the remaining* $n-m$ *characteristic roots of* A *are characteristic roots of* (C/X_1) *where* C *is defined by* (3.73).

Proof. Since $AX = XD$, where $D = \mathrm{diag}(\lambda_1,\ldots,\lambda_m)$; cf. (3.69a), Theorem 3.7 directly implies the result. (qed)

§3.4 *Quadratic forms.*

An alternate interpretation for the Schur complement is as the coefficient matrix of a quadratic form restricted to the null space of a matrix.

THEOREM 3.8 (Cottle, 1974). *Consider the quadratic form*

$$Q = \underset{\sim}{z}'A\underset{\sim}{z} = (\underset{\sim}{x}' \quad \underset{\sim}{y}')\begin{pmatrix} E & F \\ F' & H \end{pmatrix}\begin{pmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \end{pmatrix} , \tag{3.76}$$

where A *is symmetric and* E *is nonsingular. Let* Q_r *denote* Q *constrained by the system of equations*

$$E\underset{\sim}{x} + F\underset{\sim}{y} = \underset{\sim}{0} . \tag{3.77}$$

Then

$$Q_r = \underset{\sim}{y}'(A/E)\underset{\sim}{y} . \tag{3.78}$$

Proof. We may write

$$Q = \underset{\sim}{x}'E\underset{\sim}{x} + 2\underset{\sim}{x}'F\underset{\sim}{y} + \underset{\sim}{y}'H\underset{\sim}{y} . \tag{3.79}$$

Using (3.77) and the fact that E is nonsingular, we obtain

$$\underset{\sim}{x} = -E^{-1}F\underset{\sim}{y} . \tag{3.80}$$

Substituting (3.80) in (3.79), yields

$$Q_r = \underset{\sim}{y}'(H - F'E^{-1}F)\underset{\sim}{y} = \underset{\sim}{y}'(A/E)\underset{\sim}{y} . \tag{3.81}$$

(qed)

In Theorem 3.8 we restricted Q to the null space of a submatrix of A. More generally now let us restrict Q to the null space of the matrix $M = (K, L)$. Thus $M\underset{\sim}{z} = \underset{\sim}{0}$. We obtain

THEOREM 3.9 (Cottle, 1974). *Let* Q_s *denote the quadratic form* (3.76) *constrained by the system of equations*

$$K\underset{\sim}{x} + L\underset{\sim}{y} = \underset{\sim}{0} , \tag{3.82}$$

where K *is nonsingular. Let*

$$B = \left(\begin{array}{cc|c} 0 & K & L \\ K' & E & F' \\ \hline L' & F' & H \end{array}\right) \tag{3.83}$$

and

$$C = \begin{pmatrix} 0 & K \\ K' & E \end{pmatrix} . \tag{3.84}$$

Then

$$Q_s = \underset{\sim}{y}'(B/C)\underset{\sim}{y} . \tag{3.85}$$

Proof. Using (3.82) with K nonsingular, we obtain

$$\underset{\sim}{x} = -K^{-1}L\underset{\sim}{y} . \tag{3.86}$$

Substituting (3.86) in (3.79), we have

$$Q_s = \underset{\sim}{y}'[H - 2L'(K^{-1})'F + L'(K^{-1})'EK^{-1}L]\underset{\sim}{y} . \tag{3.87}$$

Since K is nonsingular, so is C and the inverse

$$(3.88)\qquad C^{-1} = \begin{pmatrix} -(K')^{-1}EK^{-1} & (K')^{-1} \\ K^{-1} & 0 \end{pmatrix}$$

is obtained using (2.40). Hence

$$(3.89)\qquad (B/C) = H - (L' \quad F')\begin{pmatrix} -(K')^{-1}EK^{-1} & (K')^{-1} \\ K^{-1} & 0 \end{pmatrix}\begin{pmatrix} L \\ F \end{pmatrix}$$

$$= H + L'(K')^{-1}EK^{-1}L - F'K^{-1}L - L'(K')^{-1}F ,$$

and so (3.87) = (3.85). (qed)

We may combine Theorems 3.8 and 3.9 to yield

THEOREM 3.10 (Cottle, 1974). *Let* Q_t *denote the quadratic form*

$$(3.90)\qquad \underset{\sim}{v}'B\underset{\sim}{v} = (\underset{\sim}{w}', \underset{\sim}{x}', \underset{\sim}{y}')\begin{pmatrix} 0 & K & L \\ K' & E & F \\ L' & F' & H \end{pmatrix}\begin{pmatrix} \underset{\sim}{w} \\ \underset{\sim}{x} \\ \underset{\sim}{y} \end{pmatrix}$$

constrained by (3.82), *with* E *and* K *nonsingular. Let* C *be defined as in* (3.84). *Then*

$$(3.91)\qquad Q_t = \underset{\sim}{y}'(B/C)\underset{\sim}{y} .$$

CHAPTER IV

RESULTS ON GENERALIZED SCHUR COMPLEMENTS

When the submatrix E in the partitioned matrix

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} \tag{4.1}$$

is rectangular, or square but singular, then the definition (2.1) of Schur complement cannot be used. Using generalized inverses, however, we may define (cf. Marsaglia and Styan, 1974a,b)

$$S = H - GE^{-}F \tag{4.2}$$

as a *generalized Schur complement of* E *in* A, where A^{-} is any solution to

$$AA^{-}A = A \; . \tag{4.3}$$

Following Rao (1962) and Rao and Mitra (1971), we will call A^{-} a *generalized inverse* (or g-inverse) of A. If $A^{-}AA^{-} = A^{-}$ also holds, then we will call A^{-} a *reflexive* g-inverse. Hence, a g-inverse A^{-} is *reflexive* if and only if it has the same rank as A (cf. proof in §4.6). A reflexive g-inverse A^{-} such that AA^{-} and $A^{-}A$ are both symmetric is unique and is denoted by A^{+}, the *Moore-Penrose* g-inverse of A. We note, however, that while a g-inverse and a reflexive g-inverse can always be found for matrices with elements over an arbitrary field, the Moore-Penrose g-inverse will only exist for those fields which have a "transpose" operator, so that $A'A$ and AA' are defined and have the same rank as both A and A', and so that $(A'B)' = B'A$. Cf. Marsaglia & Styan (1974b, p. 438).

Carlson, Haynsworth & Markham (1974) considered matrices over the complex field and used the Moore-Penrose g-inverse E^{+} in their definition of generalized Schur complement. Other writers, such as Rohde (1965), Khatri (1969), Meyer (1970), and Pringle & Rayner (1970), used (4.2) without giving it a name. See also Hartwig (1976a,b).

§4.1 *Determinants.*

When A is partitioned as in (4.1) and E is singular, then the analogue of Schur's determinant formula (2.4)

$$(4.4)\qquad \begin{vmatrix} E & F \\ G & H \end{vmatrix} = |E|.|H - GE^{-}F|$$

need not hold, e.g.,

$$(4.5)\qquad \left|\begin{array}{c|c} 0 & 1 \\ \hline 1 & 1 \end{array}\right| = -1 \neq 0 = |0|.|1 - 1.0^{-}.1| \ .$$

Sufficient conditions for (4.4) to hold, however, were obtained by Carlson, Haynsworth and Markham (1974) using Moore-Penrose g-inverses. We extend their results to arbitrary g-inverses using the following:

LEMMA 4.1 (Marsaglia & Styan, 1974a, p. 274, Th.5). *For matrices over an arbitrary field:*

$$(4.6)\qquad r(E,F) = r(E) + r([I - EE^{-}]F) = r([I - FF^{-}]E) + r(F)$$

for every E^{-} , F^{-} , *and*

$$(4.7)\qquad r\begin{pmatrix} E \\ G \end{pmatrix} = r(E) + r(G[I - E^{-}E]) = r(E[I - G^{-}G]) + r(G)$$

for every E^{-} , G^{-} .

Proof. We may write

$$r(E,F) = r\left[(E,F)\begin{pmatrix} I & -E^{-}F \\ 0 & I \end{pmatrix}\right] \tag{4.8}$$

$$= r(E,[I - EE^{-}]F) \tag{4.9}$$

$$= r(E) + r([I - EE^{-}]F) , \tag{4.10}$$

since the column spaces of E and $(I - EE^{-})F$ are virtually disjoint: if $\underset{\sim}{a} = E\underset{\sim}{b} = (I - EE^{-})F\underset{\sim}{c}$, then $(I - EE^{-})\underset{\sim}{a} = \underset{\sim}{0} = (I - EE^{-})F\underset{\sim}{c} = \underset{\sim}{a}$, as $I - EE^{-}$ is idempotent. This proves the first equation in (4.6). The second equation in (4.6) and both equations in (4.7) may be proved similarly. (qed)

THEOREM 4.1. *Let the matrix*

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} \tag{4.11}$$

have elements over an arbitrary field and suppose that both A *and* E *are square. If either*

$$r(E,F) = r(E) \tag{4.12}$$

or

$$r\begin{pmatrix} E \\ G \end{pmatrix} = r(E) , \tag{4.13}$$

then

$$|A| = |E| . |H - GE^{-}F| \tag{4.14}$$

for every g-inverse E^{-} .

Proof. It follows from Lemma 4.1 that (4.12) implies

$$EE^{-}F = F \tag{4.15}$$

for every g-inverse E^{-}. In this event, writing

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} E & 0 \\ G & I \end{pmatrix} \cdot \begin{pmatrix} I & E^{-}F \\ 0 & H - GE^{-}F \end{pmatrix} \tag{4.16}$$

and taking determinants yields (4.14). A similar proof works when (4.13) holds. (qed)

We note that neither $H - GE^{-}F$ nor its determinant is necessarily invariant under choice of E^{-}, when either (4.12) or (4.13), but not both, holds. However (4.14) shows that either $|E| = 0$ or E is nonsingular and $H - GE^{-}F = (A/E) = H - GE^{-1}F$.

When, however, both (4.12) and (4.13) hold (which is so when A has the structure (4.60) below, e.g., A nonnegative definite), then $H - GE^{-}F$ is invariant under choice of E^{-}, since (4.12) $\Rightarrow F = EL$ and (4.13) $\Rightarrow G = ME$, for some L and M. Hence $GE^{-}F = MEE^{-}EL = MEL = MEE^{\sim}EL$ for every g-inverse $E^{\sim}$.

COROLLARY 4.1. *If* A *and* H *in* (4.11) *are both square and if either*

$$r(G,H) = r(H) \tag{4.17}$$

or

$$r\begin{pmatrix} F \\ H \end{pmatrix} = r(H), \tag{4.18}$$

then

$$|A| = |H| \cdot |E - FH^{-}G| \tag{4.19}$$

for every g-inverse H^{-}.

We note that neither (4.12) nor (4.13) is *necessary* for (4.14) to hold, for if E is singular then (4.14) just says that A is singular, and if

$$(4.20)\qquad A = \left(\begin{array}{c|cc} 0 & 1 & 1 \\ \hline 1 & 1 & 1 \\ 1 & 1 & 1 \end{array}\right),$$

then both E and A are clearly singular. It would be interesting to find necessary and sufficient conditions for (4.14) to hold, viz., when does $|E| = 0$ imply $|A| = 0$?

Carlson, Haynsworth and Markham (1974) refer to the result in Theorem 4.2 below as Sylvester's determinant formula. We notice that this result parallels that of Jacobi (1834), our Theorem 2.9.

THEOREM 4.2. *Consider the* $n\times n$ *matrix*

$$(4.21)\qquad A = \begin{pmatrix} E & F \\ G & H \end{pmatrix},$$

where E *is* $m\times m$, *possibly singular. Let* $D = \{d_{ij}\}$, *where*

$$(4.22)\qquad d_{ij} = \begin{vmatrix} E & \underset{\sim}{f}_j \\ \underset{\sim}{g}_i' & h_{ij} \end{vmatrix};\ i,j = 1,2,\ldots,n-m,$$

and $\underset{\sim}{f}_j$, $\underset{\sim}{g}_i'$ *denote, respectively, the* jth *column of* F *and the* ith *row of* G, *and* $H = \{h_{ij}\}$. *If either*

$$(4.23)\qquad r(E,F) = r(E)$$

or

$$(4.24)\qquad r\begin{pmatrix} E \\ G \end{pmatrix} = r(E)$$

then

$$(4.25)\qquad D = |E|.(A/E)$$

for every generalized Schur complement $(A/E) = H - GE^{-}F$ *and*

$$(4.26)\qquad |D| = |E|^{n-m-1}|A| .$$

Proof. Theorem 4.1 yields $d_{ij} = |E|.(h_{ij} - \underset{\sim}{g}_i'E^{-}\underset{\sim}{f}_j)$, which gives (4.25) immediately, and hence (4.26), since D is $(n-m) \times (n-m)$. (qed)

§4.2 *Rank.*

When A is partitioned as in (4.1) and E is singular, then rank need not be additive on the generalized Schur complement, for

$$(4.27)\qquad r\left(\begin{array}{c|c} 0 & 1 \\ \hline 1 & 1 \end{array}\right) = 2 \neq r(0) + r(1 - 1.0^{-}.1) ,$$

which equals 0 or 1 according as 0^{-} is chosen as 1 or not 1.

We may, however, following Meyer (1973) and Marsaglia & Styan (1974a), establish the following:

THEOREM 4.3. *For matrices over an arbitrary field,*

$$(4.28)\qquad r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r\begin{pmatrix} 0 & (I - EE^{-})F \\ G(I - E^{-}E) & H - GE^{-}F \end{pmatrix} .$$

Three different choices of E^{-} *may be made.*

Proof. We note that

$$(4.28a)\qquad \begin{pmatrix} I & 0 \\ -GE^{-} & I \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} \begin{pmatrix} I & -E^{\sim}F \\ 0 & I \end{pmatrix} = \begin{pmatrix} E & X \\ Y & K \end{pmatrix} ,$$

where $E^{\sim}$ is a g-inverse of E, possibly different to E^{-},

$$X = (I - EE^{\sim})F \ ; \ Y = G(I - E^{-}E) \tag{4.29}$$

and

$$K = H - GE^{-}F - YE^{\sim}F \ . \tag{4.30}$$

Then

$$r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r\begin{pmatrix} E & X \\ Y & K \end{pmatrix} = r\begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} + r\begin{pmatrix} 0 & X \\ Y & K \end{pmatrix} , \tag{4.31}$$

since the columns (rows) of E are linearly independent of the columns of X (rows of Y). Since

$$\begin{pmatrix} 0 & X \\ Y & K \end{pmatrix} = \begin{pmatrix} 0 & X \\ Y & H - GE^{-}F \end{pmatrix} \begin{pmatrix} I & -E^{\sim}F \\ 0 & I \end{pmatrix} , \tag{4.32}$$

(4.28) follows, except that the choice of E^{-} in Y is the same as that in $H - GE^{-}F$. To relax this condition we note that with $E^{\#}$ as a g-inverse of E (possibly different to E^{-}), we have that

$$\begin{pmatrix} 0 & X \\ G(I - E^{-}E) & S \end{pmatrix} = \begin{pmatrix} 0 & X \\ G(I - E^{\#}E) & S \end{pmatrix} \begin{pmatrix} I - E^{-}E & 0 \\ 0 & I \end{pmatrix} \tag{4.33}$$

$$\begin{pmatrix} 0 & X \\ G(I - E^{\#}E) & S \end{pmatrix} = \begin{pmatrix} 0 & X \\ G(I - E^{-}E) & S \end{pmatrix} \begin{pmatrix} I - E^{\#}E & 0 \\ 0 & I \end{pmatrix} \tag{4.34}$$

where $S = H - GE^{-}F$, and hence

$$r\begin{pmatrix} 0 & X \\ G(I - E^{-}E) & S \end{pmatrix} = r\begin{pmatrix} 0 & X \\ G(I - E^{\#}E) & S \end{pmatrix} \tag{4.35}$$

is invariant under choice of E^{-} . This completes the proof. (qed)

Marsaglia and Styan (1974a, (8.5)) obtained Theorem 4.3 but required that the E^- in the lower right corner of

$$\begin{pmatrix} 0 & (I-EE^-)F \\ G(I-E^-E) & H-GE^-F \end{pmatrix} \tag{4.36}$$

must be either the E^- in the lower left or in the upper right corner.

COROLLARY 4.2. *For matrices over an arbitrary field,*

$$r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(H) + r\begin{pmatrix} E-FH^-G & F(I-H^-H) \\ (I-HH^-)G & 0 \end{pmatrix}. \tag{4.37}$$

Three different choices of H^- *may be made.*

We may expand the rank of (4.36) using Corollary 4.2 to obtain

$$r\begin{pmatrix} 0 & (I-EE^-)F \\ G(I-E^-E) & S \end{pmatrix} = r(S) + r\begin{pmatrix} U & V \\ W & 0 \end{pmatrix}, \tag{4.38}$$

where

$$U = -(I-EE^-)FS^-G(I-E^-E) \tag{4.39a}$$

$$V = (I-EE^-)F(I-S^-S) \tag{4.39b}$$

$$W = (I-SS^-)G(I-E^-E) \, . \tag{4.39c}$$

We now use

LEMMA 4.2 (Marsaglia & Styan, 1974a, (8.3)). *For matrices over an arbitrary field,*

$$r\begin{pmatrix} 0 & X \\ Y & S \end{pmatrix} = r(X) + r(Y) + r[(I-YY^-)S(I-X^-X)] \, . \tag{4.40}$$

Any choices of X^- *and* Y^- *may be made.*

Proof. Using Lemma 4.1 yields

$$(4.41)\qquad r\begin{pmatrix} O & X \\ \hline Y & S \end{pmatrix} = r(X) + r(Y,\ S)\begin{pmatrix} I & 0 \\ 0 & I - X^{-}X \end{pmatrix} .$$

$$(4.42)\qquad = r(X) + r(Y,\ S(I - X^{-}X)).$$

Applying (4.6) gives (4.40). (qed)

We now expand the rank of (4.36) using (4.40) to obtain

THEOREM 4.4. *For matrices over an arbitrary field,*

$$(4.43)\qquad r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(S) + r(V) + r(W) + r(Z) ,$$

where

$$(4.44)\qquad Z = (I - VV^{-})U(I - W^{-}W) ,$$

while U , V *and* W *are as in* (4.39). *The g-inverses may be any choices.*

Meyer (1973, Cor. 4.1) proved that

$$(4.45)\qquad r\begin{pmatrix} E & F \\ G & H \end{pmatrix} \le r(E) + r(S) + r(F) + r(G) .$$

To see this we notice that using (4.28) and (4.38) yields

$$(4.46)\qquad r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(S) + r\begin{pmatrix} U & V \\ W & 0 \end{pmatrix} ,$$

and

$$(4.47)\qquad r\begin{pmatrix} U & V \\ W & 0 \end{pmatrix} \le r(U,V) + r(W)$$

$$(4.48)\qquad \le r(U,V) + r(G)$$

$$(4.49)\qquad = r[(I-EE^-)F(-S^-G(I-E^-E),\ I-S^-S)] + r(G)$$

$$(4.50)\qquad \le r(F) + r(G)\ ,$$

which proves the inequality (4.45).

Meyer (1973, Th. 4.1) also proved:

THEOREM 4.5. *For matrices over an arbitrary field*

$$(4.51)\qquad r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(X) + r(Y) + r[(I-YY^-)(H-GE^-F)(I-X^-X)],$$

where X *and* Y *are as defined in* (4.29). *Any choices of g-inverses may be made.*

Proof. Immediate by applying Lemma 4.2 to (4.28). (qed)

Marsaglia and Styan (1974a, (8.6)) obtained (4.51) but required the E^- in (4.51) to be the same as that chosen in X or Y . In view of our proof of Theorem 4.3 this requirement is not needed.

We will refer to

$$(4.52)\qquad S = H - GE^{\sim}F$$

as *the generalized Schur complement of* E *in* A , *relative to the choice* $E^{\sim}$, where

$$(4.53)\qquad A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}.$$

COROLLARY 4.3 (Marsaglia & Styan, 1974a, p.291, Cor. 19.1). *For matrices over an arbitrary field, rank is additive on the Schur complement:*

$$(4.54)\qquad r\begin{pmatrix} E & F \\ G & H \end{pmatrix} = r(E) + r(H - GE^{\sim}F),$$

where $E^{\sim}$ *is a particular g-inverse of* E *, if and only if*

$$(4.55a)\qquad (I - EE^{-})F(I - S^{-}S) = 0$$

$$(4.55b)\qquad (I - SS^{-})G(I - E^{-}E) = 0$$

$$(4.55c)\qquad (I - EE^{-})FS^{-}G(I - E^{-}E) = 0 ,$$

where $S = H - GE^{\sim}F$ *, while* E^{-} *and* S^{-} *are any choices of g-inverses.*

Proof. Immediate from Theorem 4.4. (qed)

Corollary 4.3 was proved by Carlson, Haynsworth and Markham (1974) with $E^{\sim} = E^{+}$, the Moore-Penrose g-inverse. They assert that their proof can be used to cover the case where $E^{\sim}$ is a reflexive g-inverse. See also Carlson (1975, Th. A).

We note that if the conditions in Corollary 4.3 hold, then $|E| = 0$ implies $|A| = 0$; cf. the discussion before Theorem 4.2.

§4.3 *Generalized inverses.*

Our objective in this section is to investigate conditions under which the Schur-Banachiewicz inversion formula works with generalized inverses replacing regular inverses.

Consider

$$(4.56)\qquad A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$$

and

$$(4.57) \qquad B = \begin{pmatrix} E^- + E^-FS^-GE^- & -E^-FS^- \\ -S^-GE^- & S^- \end{pmatrix},$$

where

$$(4.58) \qquad S = H - GE^-F .$$

Rohde (1965) showed that if A is real and nonnegative definite then indeed B is a g-inverse of A. This result was extended by Pringle & Rayner (1970), who assumed that A has the structure

$$(4.59) \qquad A = \begin{pmatrix} K'K & K'L \\ L'K & 0 \end{pmatrix},$$

and later by Marsaglia and Styan (1974b, Cor. 1) for

$$(4.60) \qquad A = \begin{pmatrix} K'K & K'L \\ M'K & N \end{pmatrix},$$

cf. Corollary 4.6 below. More generally, Bhimasankaram (1971) and Burns, Carlson, Haynsworth and Markham (1974) showed that $B = A^-$ if and only if the conditions (4.55) hold. Applying Corollary 4.3 we then get

THEOREM 4.6 (Marsaglia & Styan, 1974b, p. 439). *Suppose that the matrix* A*, defined by* (4.56), *has elements over an arbitrary field, and that* $E^\sim$ *is a particular g-inverse of* E. *Let the Schur complement* $S = H - GE^\sim F$ *and let*

$$(4.61) \qquad B = \begin{pmatrix} E^\sim + E^\sim FS^-GE^\sim & -E^\sim FS^- \\ -S^-GE^\sim & S^- \end{pmatrix}.$$

Then:

(i) B *is a g-inverse of* A *for a particular g-inverse* S^- *if and only if rank is additive on the Schur complement (i.e.,* (4.54) *holds), and then* B *is a g-inverse of* A *for every g-inverse* S^-.

(ii) *The g-inverse* B *is reflexive if and only if* $E^\sim$ *and* S^- *are both reflexive g-inverses.*

(iii) *For complex* A, $B = A^+$, *the Moore-Penrose g-inverse of* A, *if and only if* $E^\sim = E^+$, $S^- = S^+$,

$$(4.62) \qquad r\begin{pmatrix} E \\ G \end{pmatrix} = r(E,F) = r(E)$$

and

$$(4.63) \qquad r\begin{pmatrix} F \\ H \end{pmatrix} = r(G,H) = r(S) = r(H - GE^{\sim}F) \ .$$

Proof. (i) Straightforward multiplication shows that $ABA = A \Leftrightarrow (4.55)$, and so B is a g-inverse of $A \Leftrightarrow (4.54)$ holds.

(ii) & (iii). These proofs are straightforward but more lengthy: we refer the reader to Marsaglia & Styan (1974b, pp. 438-9) for details.

(qed)

Bhimasankaram (1971) and Burns, Carlson, Haynsworth & Markham (1974) proved that the matrix B defined by (4.61) is a g-inverse of A if and only if (4.55) holds.

Similarly it may be shown that if $H^{\sim}$ is a particular g-inverse of H and $T = E - FH^{\sim}G$ is the generalized Schur complement relative to the choice $H^{\sim}$ then

$$(4.64) \qquad C = \begin{pmatrix} T^{-} & -T^{-}FH^{\sim} \\ -H^{\sim}GT^{-} & H^{\sim} + H^{\sim}GT^{-}FH^{\sim} \end{pmatrix}$$

is a g-inverse of A for a particular g-inverse T^{-} if and only if

$$(4.65) \qquad r(A) = r(H) + r(E - FH^{\sim}G) \ ,$$

and then C is a g-inverse of A for every g-inverse T^{-}. The g-inverse C is reflexive if and only if $H^{\sim}$ and T^{-} are both reflexive g-inverses. For complex A, $C = A^{+}$, the Moore-Penrose g-inverse of A, if and only if $H^{\sim} = H^{+}$, $T^{-} = T^{+}$, and

$$(4.66) \qquad r\begin{pmatrix} F \\ H \end{pmatrix} = r(G,H) = r(H)$$

and

$$(4.67) \qquad r\begin{pmatrix} E \\ G \end{pmatrix} = r(E,F) = r(T) = r(E - FH^{\sim}G) \ .$$

Since the Moore-Penrose g-inverse is unique, we obtain

COROLLARY 4.4. *Let the complex matrix*

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} , \tag{4.68}$$

and let

$$S = H - GE^{+}F \ , \ T = E - FH^{+}G \ . \tag{4.69}$$

Then

$$\begin{pmatrix} T^{+} & -E^{+}FS^{+} \\ -H^{+}GT^{+} & S^{+} \end{pmatrix} = A^{+} , \tag{4.70}$$

the Moore-Penrose g-inverse of A, *if*

$$r\begin{pmatrix} E \\ G \end{pmatrix} = r(E,F) = r(E) = r(T) \tag{4.71}$$

and

$$r\begin{pmatrix} F \\ H \end{pmatrix} = r(G,H) = r(H) = r(S) \ . \tag{4.72}$$

Burns, Carlson, Haynsworth & Markham (1974) noted that the Moore-Penrose g-inverse of A is given by (4.70) if (4.61) and (4.64) equal A^{+}, since the Moore-Penrose g-inverse is unique. Moreover, Theorem 4.6 yields:

COROLLARY 4.5. *Let* $E^{\sim} = E^{+}$ *and* $S^{-} = S^{+}$, *where*

$$S = H - GE^{+}F = (A/E) , \tag{4.73}$$

A *being defined by* (4.56). *If* (4.62) *and* (4.63) *hold, then* $B = A^{+}$ *and*

$$(A^{+}/S^{+})^{+} = (A^{+}/(A/E)^{+})^{+} = E \ . \tag{4.74}$$

Similarly, let $H^{\sim} = H^{+}$ and $T^{-} = T^{+}$ where

$$T = E - FH^{+}G = (A/H) , \tag{4.75}$$

A being defined by (4.56). If (4.66) and (4.67) hold, then $C = A^{+}$ and

$$(A^{+}/T^{+})^{+} = (A^{+}/(A/H)^{+})^{+} = H . \tag{4.76}$$

Burns, Carlson, Haynsworth & Markham (1974) proved that, if

$$G(I - E^{+}E) = 0 ; \quad (I - EE^{+})F = 0 ; \tag{4.77a}$$

$$(I - SS^{+})G = 0 ; \quad F(I - S^{+}S) = 0 , \tag{4.77b}$$

where $S = H - GE^{+}F$, then (4.74) holds. Using Lemma 4.1 it is easy to see that (4.77a) ⇔ (4.62) and (4.77b) ⇔ (4.63).

COROLLARY 4.6 (Marsaglia & Styan, 1974b). *Suppose that the real matrix* A *is defined by*

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} K'K & K'L \\ M'K & N \end{pmatrix} . \tag{4.78}$$

Then rank is additive on the Schur complement:

$$r(A) = r(K) + r(S) , \tag{4.79}$$

where

$$S = H - GE^{-}F = N - M'KK^{+}L \tag{4.80}$$

is independent of the choice of g-inverse E^{-} . *The matrix*

$$B = \begin{pmatrix} E^{-} + E^{-}FS^{-}GE^{-} & -E^{-}FS^{-} \\ -S^{-}GE^{-} & S^{-} \end{pmatrix} \tag{4.81}$$

is a g-inverse of A *for any choice of g-inverses* E^- , S^- . *Furthermore,* $B = A^+$, *the Moore-Penrose g-inverse of* A , *if and only if* $E^- = E^+, S^- = S^+$, *and*

$$(4.82) \qquad r\begin{pmatrix} F \\ H \end{pmatrix} = r(G,H) = r(H - GE^+F) = r(S) .$$

Proof. Since $K'K(K'K)^-K' = K'$, it follows that $(I - EE^-)F = 0$. Similarly, it may be shown that $G(I - E^-E) = 0$. Hence (4.55) holds and so rank is additive on the Schur complement (i.e., (4.79) holds). The Schur complement is unique since $GE^-F = M'K(K'K)^-K'L = M'KK^+L$. The conditions for B to equal A^+ follow since (4.62) and (4.63) reduce to (4.82). (qed)

Rohde (1965) obtained (4.81) for a g-inverse of A , where A is defined by (4.78) with $M = L$ and $N = L'L$. Pringle and Rayner (1970) also established that B , given by (4.81), is a g-inverse of A , where A is defined by (4.78), with $M = L$ and $N = 0$. The following corollary gives a different approach to Rohde's result. See also Ruohonen (1973).

COROLLARY 4.7 (Marsaglia & Styan, 1974b). *Suppose that the real nonnegative definite matrix* A *is defined by*

$$(4.83) \qquad A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix} = \begin{pmatrix} K'K & K'L \\ L'K & L'L \end{pmatrix} .$$

If any one of the following three conditions holds then all three hold.

$$(4.84) \qquad r\begin{pmatrix} E & F \\ F' & H \end{pmatrix} = r(E) + r(H) ,$$

$$(4.85) \qquad \begin{pmatrix} E & F \\ F' & H \end{pmatrix}^+ = \begin{pmatrix} E^+ + E^+FS^+F'E^+ & -E^+FS^+ \\ -S^+F'E^+ & S^+ \end{pmatrix} ,$$

$$(4.86) \qquad \begin{pmatrix} E & F \\ F' & H \end{pmatrix}^+ = \begin{pmatrix} T^+ & -T^+FH^+ \\ -H^+F'T^+ & H^+ + H^+F'T^+FH^+ \end{pmatrix} ,$$

where the Schur complements of E , *and of* H , *in* A ,

(4.87) $$S = H - F'E^{-}F = L'(I - KK^{+})L$$

and

(4.88) $$T = E - FH^{-}F' = K(I - LL^{+})K' ,$$

are independent of the choices of g-inverses E^{-} *and* H^{-} .

Proof. Since $F(I - H^{-}H) = K'L(I - (L'L)^{-}L'L) = 0$, it follows that

(4.89) $$r\begin{pmatrix} F \\ H \end{pmatrix} = r(H)$$

using Lemma 4.1. First, suppose (4.84) holds. Using (4.79), it follows that

(4.90) $$r(S) = r(H) = r\begin{pmatrix} F \\ H \end{pmatrix}$$

and so (4.82) holds which implies (4.85). By the reverse argument, (4.85) implies (4.84). The alternate arrangement in (4.86) follows from the "symmetry" in (4.84) with respect to E and H . (qed)

§4.4 *Inertia*.

Consider the real symmetric matrix

(4.91) $$A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix} ,$$

where E is singular. Then inertia need not be additive on the generalized Schur complement in contrast to the case where E is nonsingular (Theorem 3.1). We find, however, that under certain conditions inertia does continue to be additive on the (generalized) Schur complement. To see this we use the following:

LEMMA 4.3. *Suppose that the real symmetric* $n \times n$ *matrices* A *and* B *are rank additive:*

(4.92) $$r(A+B) = r(A) + r(B) .$$

Then

(4.93) $$In(A+B) \doteq In A + In B \ ,$$

where In denotes inertia.

Following Carlson, Haynsworth & Markham (1974, p. 172) by (4.93) we mean that $\pi(A+B) = \pi(A) + \pi(B)$ and $\nu(A+B) = \nu(A) + \nu(B)$, where $\pi(.)$ and $\nu(.)$ denote, respectively, the number of positive and negative characteristic roots.

Proof. Following (3.1) let

(4.93a) $$In A = (\pi_a, \nu_a, \delta_a) \ , \quad r(A) = \pi_a + \nu_a = r_a \ ,$$
$$In B = (\pi_b, \nu_b, \delta_b) \ , \quad r(B) = \pi_b + \nu_b = r_b \ .$$

Since A and B are both n×n real symmetric matrices, there exist real nonsingular matrices S and T such that

(4.94a) $$A = SD_aS' = (S_1, S_2, S_3)\begin{pmatrix} I_{\pi_a} & 0 & 0 \\ 0 & -I_{\nu_a} & 0 \\ 0 & 0 & 0_{\delta_a} \end{pmatrix}\begin{pmatrix} S_1' \\ S_2' \\ S_3' \end{pmatrix}$$

(4.94b) $$= S_1S_1' - S_2S_2'$$

and

(4.95a) $$B = TD_bT' = (T_1, T_2, T_3)\begin{pmatrix} I_{\pi_b} & 0 & 0 \\ 0 & -I_{\nu_b} & 0 \\ 0 & 0 & 0_{\delta_b} \end{pmatrix}\begin{pmatrix} T_1' \\ T_2' \\ T_3' \end{pmatrix}$$

(4.95b) $$= T_1T_1' - T_2T_2' \ .$$

Using (4.94b) and (4.95b), we obtain

(4.96a) $$A+B = S_1S_1' - S_2S_2' + T_1T_1' - T_2T_2'$$

(4.96b) $$= (S_1, S_2, T_1, T_2) \begin{pmatrix} I_{\pi_a} & 0 & 0 & 0 \\ 0 & -I_{\nu_a} & 0 & 0 \\ 0 & 0 & I_{\pi_b} & 0 \\ 0 & 0 & 0 & -I_{\nu_b} \end{pmatrix} \begin{pmatrix} S_1' \\ S_2' \\ T_1' \\ T_2' \end{pmatrix} .$$

From rank additivity and from (4.96b), we get

(4.97) $$r(A+B) = r_a + r_b \le r(S_1, S_2, T_1, T_2) \le \pi_a + \nu_a + \pi_b + \nu_b = r_a + r_b .$$

Hence there exists a matrix U, say, $n \times (n - r_a - r_b)$, so that $V = (S_1, T_1, S_2, T_2, U)$ is nonsingular and writing

(4.98) $$A+B = (S_1, T_1, S_2, T_2, U) \begin{pmatrix} I_{\pi_a+\pi_b} & 0 & 0 \\ 0 & -I_{\nu_a+\nu_b} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} S_1' \\ T_1' \\ S_2' \\ T_2' \\ U' \end{pmatrix}$$

completes the proof. (qed)

THEOREM 4.7 (Carlson, Haynsworth & Markham, 1974). *Consider the real symmetric matrix*

(4.99) $$A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix} .$$

Let E_r^- *be a symmetric reflexive g-inverse of* E *, and let the generalized Schur complement*

(4.100) $$S = H - F'E_r^- F = (A/E).$$

Then

(4.101) $$\mathit{In}A = \mathit{In}E + \mathit{In}(A/E) + \mathit{In}\begin{pmatrix} -XS_r^- X' & V \\ V' & 0 \end{pmatrix},$$

where

(4.102) $$X = (I - EE_r^-)F \quad \textit{and} \quad V = X(I - S_r^- S),$$

and S_r^- *is a symmetric reflexive g-inverse of* S . *Furthermore, if*

(4.103) $$r(E) = r(E, F)$$

then S *is unique and*

(4.104) $$\mathit{In}A = \mathit{In}E + \mathit{In}(A/E) .$$

Remark. The notation used in (4.101), as in (4.93), is taken to mean additivity of the numbers of positive characteristic roots *and* of the numbers of negative roots, but not necessarily of the numbers of zero roots.

Proof. We may write

(4.105) $$\begin{pmatrix} I & 0 \\ -F'E_r^- & I \end{pmatrix} \begin{pmatrix} E & F \\ F' & H \end{pmatrix} \begin{pmatrix} I & -E_r^- F \\ 0 & I \end{pmatrix} = \begin{pmatrix} E & X \\ X' & S \end{pmatrix},$$

since E_r^- is a symmetric reflexive g-inverse of E . Then by Sylvester's law of inertia, cf. (3.2), we obtain

(4.106) $$\mathit{In}A = \mathit{In}\begin{pmatrix} E & X \\ X' & S \end{pmatrix}.$$

Using Theorem 4.3 yields

$$(4.107)\qquad r\begin{pmatrix} E & X \\ X' & S \end{pmatrix} = r\begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} + r\begin{pmatrix} 0 & X \\ X' & S \end{pmatrix} ;$$

which, using Lemma 4.3, yields

$$(4.108)\qquad In A = In E + In \begin{pmatrix} 0 & X \\ X' & S \end{pmatrix} .$$

However, we may write, cf. (4.105),

$$(4.109)\qquad \begin{pmatrix} 0 & X \\ X' & S \end{pmatrix} = \begin{pmatrix} I & XS_r^- \\ 0 & I \end{pmatrix} \begin{pmatrix} U & V \\ V' & S \end{pmatrix} \begin{pmatrix} I & 0 \\ S_r^- X' & I \end{pmatrix} ,$$

where S_r^- is a symmetric reflexive g-inverse of the Schur complement $S = (A/E)$, and $U = -XS_r^-X'$. Hence, using (3.2) again, we obtain

$$(4.110)\qquad In \begin{pmatrix} 0 & X \\ X' & S \end{pmatrix} = In \begin{pmatrix} U & V \\ V' & S \end{pmatrix} = In S + In \begin{pmatrix} U & V \\ V' & 0 \end{pmatrix} ,$$

since

$$(4.111)\qquad r\begin{pmatrix} U & V \\ V' & S \end{pmatrix} = r(S) + r\begin{pmatrix} -XS_r^-X' & V \\ V' & 0 \end{pmatrix} ,$$

using Corollary 4.2 and Lemma 4.3. Thus (4.101) follows at once. If (4.103) holds then Lemma 4.1 shows $X = 0$ and so $V = 0$ and (4.104) follows. (qed)

When A is nonnegative definite then (4.103) holds, cf. (4.83), and so E and (A/E) are both nonnegative definite. Conversely, if (4.103) holds and both E and (A/E) are nonnegative definite then A is nonnegative definite (cf. Corollary 3.1). Moreover, we note that (4.103) implies (4.104) even when E_r^- in (4.100) is replaced by any E^-, for then (4.105) holds with $X = (I-EE^-)F = 0$, while $(A/E) = H - F'E^-F = H-F'E^+F$ for every E^-.

§4.5 *The quotient property and a related determinant inequality.*

The quotient property (cf. (3.17) in §3.2) may, under certain conditions, be extended using generalized Schur complements.

THEOREM 4.8 (Carlson, Haynsworth & Markham, 1974). *Consider the matrix*

$$A = \left(\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right) = \left(\begin{array}{cc|c} K & L & F_1 \\ M & N & F_2 \\ \hline G_1 & G_2 & H \end{array}\right) . \tag{4.112}$$

If

$$r(E) = r(E,\ F) = r\begin{pmatrix} E \\ G \end{pmatrix} \tag{4.113}$$

and

$$r(K) = r(K,\ L) = r\begin{pmatrix} K \\ M \end{pmatrix} \tag{4.114}$$

then

$$r(K,\ L,\ F_1) = r(K) = r\begin{pmatrix} K \\ M \\ G_1 \end{pmatrix} \tag{4.115}$$

and the generalized Schur complements (A/E), (E/K), *and* (A/K) *are uniquely determined, and*

$$(A/E) = ((A/K)/(E/K)) . \tag{4.116}$$

Proof. From (4.113) and Lemma 4.1 we may write

$$F = EE^{-}F \qquad \text{and} \qquad G = GE^{-}E \,, \tag{4.117}$$

so that

$$F_1 = (K,\ L)E^{-}F \qquad \text{and} \qquad G_1 = GE^{-}\begin{pmatrix} K \\ M \end{pmatrix}. \tag{4.118}$$

Thus

$$r(K,\ L,\ F_1) = r(K,\ L) \qquad \text{and} \qquad r\begin{pmatrix} K \\ M \\ G_1 \end{pmatrix} = r\begin{pmatrix} K \\ M \end{pmatrix}. \tag{4.119}$$

Applying (4.114) yields (4.115). The uniqueness of the Schur complements then follows, cf. remarks before Corollary 4.1. Using (4.115) we write (4.112) as

$$A = \begin{pmatrix} K & KL_o & KF_o \\ M_oK & N & F_2 \\ G_oK & G_2 & H \end{pmatrix} \tag{4.120}$$

for some matrices L_o, F_o, M_o, and G_o . To prove (4.116) we notice that

$$(A/K) = \begin{pmatrix} (E/K) & F_2 - M_oKF_o \\ G_2 - G_oKL_o & H - G_oKF_o \end{pmatrix}, \tag{4.121}$$

so that

$$((A/K)/(E/K)) = H - G_oKF_o - (G_2 - G_oKL_o)(E/K)^{-}(F_2 - M_oKF_o)\,, \tag{4.122}$$

while

$$(A/E) = H - (G_oK,\ G_2)\begin{pmatrix} K & KL_o \\ M_oK & N \end{pmatrix}^{-}\begin{pmatrix} KF_o \\ F_2 \end{pmatrix}. \tag{4.123}$$

To see that (4.122) = (4.123) we use Theorem 4.6(i) to write, cf. also (2.37) and (3.23),

$$(4.124) \qquad \begin{pmatrix} K & KL_o \\ M_oK & N \end{pmatrix}^- = \begin{pmatrix} K^- & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} K^-KL_o \\ -I \end{pmatrix} (E/K)^-(M_oKK^-, -I) \quad .$$

Substituting (4.124) into (4.123) yields (4.122). (qed)

Carlson, Haynsworth and Markham (1974) also extended Theorem 3.4 using generalized Schur complements. Let

$$(4.125) \qquad A = \begin{pmatrix} E & F \\ F' & H \end{pmatrix}$$

and

$$(4.126) \qquad B = \begin{pmatrix} K & L \\ L' & N \end{pmatrix}$$

both be symmetric $(m+n) \times (m+n)$ nonnegative definite matrices, where E and K are both $m\times m$. Then

$$(4.127) \qquad r(E) = r(E, F) \quad \text{and} \quad r(K) = r(K, L) \quad .$$

The generalized Schur complements

$$(4.128a) \qquad (A/E) = H - F'E^-F$$

$$(4.128b) \qquad (B/K) = N - L'K^-L$$

are, therefore, uniquely defined as is

$$(4.129) \qquad ((A+B)/(E+K)) = H+N - (F+L)'(E+K)^-(F+L) \quad ,$$

since $A+B$ is nonnegative definite, cf. remarks before Corollary 4.1.

THEOREM 4.9 (Carlson, Haynsworth & Markham, 1974). *Let* A *and* B *be defined as in* (4.125) *and* (4.126). Then

$$((A+B)/(E+K)) - (A/E) - (B/K) \tag{4.130}$$

is nonnegative definite.

Proof. Consider the matrices

$$P = \begin{pmatrix} E & F \\ F' & F'E^{-}F \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} K & L \\ L' & L'K^{-}L \end{pmatrix} . \tag{4.131}$$

Then P, Q and $P+Q$ are all nnd and so then is the generalized Schur complement

$$((P+Q)/(E+K)) = F'E^{-}F + L'K^{-}L - (F+L)'(E+K)^{-}(F+L) . \tag{4.132}$$

Following the proof of Theorem 3.4 we see that (4.132) = (4.130) and the proof is complete. (qed)

Carlson, Haynsworth & Markham (1974) also extended Theorem 3.5 by allowing the principal submatrices E_i and K_i to be nonnegative definite. The inequality (3.47), however, is only meaningful when the E_i and K_i are all nonsingular. If we substitute $|A|/|E_i| = |(A/E_i)|$ and $|B|/|K_i| = |(B/K_i)|$ then we obtain:

THEOREM 4.10 (Carlson, Haynsworth & Markham, 1974). *Let* A *and* B *both be* $n\times n$ *nonnegative definite matrices. Suppose further that* E_i *and* K_i, $i=1,\ldots,n$, *are the* $i\times i$ *principal submatrices in the upper left corners of the matrices* A *and* B *respectively. Then*

$$|A+B| \geq |A| + |B| + \sum_{i=1}^{n-1} [|(A/E_i)|\cdot|K_i| + |(B/K_i)|\cdot|E_i|] , \tag{4.133}$$

Proof. We will follow the proof of Theorem 3.5 and use induction on n. For $n = 2$,

$$|A+B| = |E_1+K_1| \cdot |(A+B)/(E_1+K_1)| , \tag{4.134}$$

cf. (3.48) and the remarks before Corollary 4.1. From (3.40) we have

$$|E_1+K_1| \geq |E_1| + |K_1| , \tag{4.135}$$

while

$$|(A+B)/(E_1+K_1)| \geq |(A/E_1)| + |(B/K_1)| \tag{4.136}$$

follows from (4.130). Hence

$$\begin{aligned} |A+B| &\geq (|E_1| + |K_1|)(|(A/E_1)| + |(B/K_1)|) \\ &= |E_1| \cdot |(A/E_1)| + |K_1| \cdot |(B/K_1)| + |(A/E_1)| \cdot |K_1| + |(B/K_1)| \cdot |E_1| \\ &= |A| + |B| + |(A/E_1)| \cdot |K_1| + |(B/K_1)| \cdot |E_1| . \end{aligned} \tag{4.137}$$

Thus (4.133) holds for $n = 2$.

Now assume that (4.133) holds for A and B $n \times n$. If A_1 and B_1 are $(n+1) \times (n+1)$ nonnegative definite matrices, and $A = E_n$ and $B = K_n$ are $n \times n$ submatrices of A_1 and B_1, respectively, then

$$|A_1+B_1| = |E_n+K_n| \cdot |(A_1+B_1)/(E_n+K_n)| . \tag{4.138}$$

But, by the inductive assumption,

$$|E_n+K_n| \geq \sum_{i=1}^{n} [|(E_n/E_i)| \cdot |K_i| + |(K_n/K_i)| \cdot |E_i|] \tag{4.139}$$

and by (4.130),

$$|(A_1+B_1)/(E_n+K_n)| \geq |(A_1/E_n)| + |(B_1/K_n)| \,. \tag{4.140}$$

Hence,

$$\begin{aligned} |A_1+B_1| &\geq \{\sum_{i=1}^{n} [|(E_n/E_i)|.|K_i| + |(K_n/K_i)|.|E_i|]\}\{|(A_1/E_n)| + |(B_1/K_n)|\} \\ &= \sum_{i=1}^{n} |(A_1/E_n)|.|(E_n/E_i)|.|K_i| + \sum_{i=1}^{n} |(A_1/E_n)|.|(K_n/K_i)|.|E_i| + \\ &\quad \sum_{i=1}^{n} |(B_1/K_n)|.|(E_n/E_i)|.|K_i| + \sum_{i=1}^{n} |(B_1/K_n)|.|(K_n/K_i)|.|E_i| \\ &\geq \sum_{i=1}^{n} |(A_1/E_n)|.|(E_n/E_i)|.|K_i| + |A_1| \\ &\qquad + |B_1| + \sum_{i=1}^{n} |(B_1/K_n)|.|(K_n/K_i)|.|E_i| \\ &= \sum_{i=1}^{n+1} [|(A_1/E_i)|.|K_i| + |(B_1/K_i)|.|E_i|] \,. \end{aligned} \tag{4.141}$$

Thus (4.133) holds for $(n+1) \times (n+1)$ matrices A and B and the induction proof is complete. (qed)

§4.6 *Other results.*

In this section we present a number of miscellaneous results which extend some of the theorems and corollaries presented above.

In Corollary 2.1 we proved that if F is $m{\times}n$ and G is $n{\times}m$ then, cf. (2.8),

$$\begin{vmatrix} I_m & F \\ G & I_n \end{vmatrix} = |I_m - FG| = |I_n - GF| \,. \tag{4.142}$$

Using (2.4) and (2.6) we similarly obtain

$$(4.143) \qquad \begin{vmatrix} \lambda I_m & F \\ G & I_n \end{vmatrix} = |\lambda I_m - FG| = |\lambda I_m| \cdot |I_n - GF/\lambda|$$

and so

$$(4.144) \qquad \lambda^n |\lambda I_m - FG| = \lambda^m |\lambda I_n - GF| ,$$

which shows that FG and GF have the same nonzero characteristic roots, cf. e.g., Mirsky (1955, p. 200).

Furthermore if we replace I_n in the lower right corner of (4.143) by λI_n, F by $-F$, and G by $-F'$, then

$$(4.145) \qquad \begin{vmatrix} \lambda I_m & -F \\ -F' & \lambda I_n \end{vmatrix} = |\lambda I_m| \cdot |\lambda I_n - F'F/\lambda| = \lambda^{m-n} |\lambda^2 I_n - F'F| ,$$

and so the nonzero characteristic roots of

$$(4.146) \qquad \begin{pmatrix} 0 & F \\ F' & 0 \end{pmatrix}$$

are the pairs of positive and negative singular values of F, cf. Lanczos (1958).

A similar result to (4.142), cf. Cline and Funderlic (1976), is

$$(4.147) \qquad \psi(I_m - FG) = \psi(I_n - GF),$$

where $\psi(\cdot)$ denotes (column) nullity.

To prove (4.147) we use Theorem 2.5 to write

$$(4.148) \qquad r\begin{pmatrix} I_m & F \\ G & I_n \end{pmatrix} = n + r(I_m - FG) = m + r(I_n - GF),$$

from which (4.147) follows at once.

The nullities of $I - FG$ and $I - GF$ are related to the ranks of $F - FGF$ and $G - GFG$. Using (4.28) and (4.37) we obtain

$$(4.149) \qquad r\begin{pmatrix} F & FG \\ GF & G \end{pmatrix} = r(F) + r(G - GFG) = r(G) + r(F - FGF) ,$$

and

$$(4.150) \qquad r\begin{pmatrix} I_m & F \\ FG & F \end{pmatrix} = m + r(F - FGF) = r(F) + r(I_m - FG) .$$

Hence

$$(4.151a) \qquad \psi(I - FG) = r(F) - r(F - FGF)$$

$$(4.151b) \qquad = r(G) - r(G - GFG)$$

$$(4.151c) \qquad = \psi(I - GF) ,$$

and

$$(4.152) \qquad r(F) - r(G) = r(F - FGF) - r(G - GFG).$$

If $G = A$, and $F = A^-$ is a generalized inverse of A then (4.152) yields

$$(4.153) \qquad r(A^-) = r(A) + r(A^- - A^-AA^-) \geq r(A) ,$$

and so

$$(4.154) \qquad r(A^-) = r(A) \Leftrightarrow A^- = A^-AA^- .$$

That is, a g-inverse A^- of a matrix A is reflexive if and only if the ranks of A^- and A are the same, cf. Bjerhammar (1958; 1973, p. 383).

We may extend Theorem 2.6a (Wedderburn, 1934) which showed that if $H \neq 0$ then there exist column vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ so that $\underset{\sim}{a}'H\underset{\sim}{b} \neq 0$ and

$$r(H - H\underset{\sim}{b}\underset{\sim}{a}'H/\underset{\sim}{a}'H\underset{\sim}{b}) = r(H) - 1 . \tag{4.155}$$

THEOREM 4.11. *Let the matrices* A, B *and* H *satisfy*

$$r(A'HB) = r(A'H) = r(HB) . \tag{4.156}$$

Then

$$r(H - HB(A'HB)^-A'H) = r(H) - r(A'HB) \tag{4.157}$$

for any choice of generalized inverse.

Proof. Using (4.28) and (4.37) we obtain

$$r\begin{pmatrix} H & HB \\ A'H & A'HB \end{pmatrix} = r(H) = r(A'HB) + r(M) , \tag{4.158}$$

where

$$M = \begin{pmatrix} H - HB(A'HB)^-A'H & HB[I - (A'HB)^-A'HB] \\ [I - A'HB(A'HB)^-]A'H & 0 \end{pmatrix} . \tag{4.159}$$

Using the rank cancellation rules of Marsaglia and Styan (1974a, Th. 2) we see that

$$r(A'HB) = r(HB) \Rightarrow HB[I - (A'HB)^-A'HB] = 0 \tag{4.160}$$

and

(4.161) $$r(A'HB) = r(A'H) \Rightarrow [I - A'HB(A'HB)^{-}]A'H = 0,$$

since $A'HB = A'HB(A'HB)^{-}A'HB$. Hence (4.156) ⇒ (4.157) and the proof is complete. (qed)

Rao (1973, p. 69) presents (4.157) as an exercise when A'HB is square and nonsingular; this condition clearly implies (4.156).

In a statistical study of the residuals from a linear model, Ellenberg (1973) showed that the Schur complement of a nonsingular principal submatrix in a symmetric idempotent matrix is also idempotent. We extend this result in the following:

THEOREM 4.12. *Let*

(4.162) $$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} = A^2 .$$

If

(4.163) $$r(E) = r(E,F) = r\begin{pmatrix} E \\ G \end{pmatrix}$$

then the Schur complement

(4.164) $$(A/E) = H - GE^{-}F = (A/E)^2$$

is invariant and idempotent under choice of E^{-} *and*

(4.165) $$r(A/E) = r(A) - r(E) .$$

Proof. From (4.162) we obtain

$$(4.166)\qquad \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} E^2 + FG & EF + FH \\ GE + HG & GF + H^2 \end{pmatrix},$$

while (4.163) yields, using Lemma 4.1,

$$(4.167)\qquad EE^-F = F \quad \text{and} \quad GE^-E = G$$

for every choice of E^-. Then $(A/E) = H - GE^-F$ is invariant under choice of E^-, cf. remarks before Corollary 4.1. Hence, using (4.166) and (4.167),

$$(4.168a)\qquad (A/E)^2 = H^2 + GE^-FGE^-F - HGE^-F - GE^-FH$$

$$(4.168b)\qquad = (H - GF) + GE^-(E - E^2)E^-F - (G - GE)E^-F - GE^-(F - EF)$$

$$(4.168c)\qquad = H - GF + (GE^-E)E^-F - (GE^-E)(EE^-F) - GE^-F + G(EE^-F) - GE^-F + (GE^-E)F$$

$$(4.168d)\qquad = H - GE^-F = (A/E),$$

and (4.164) is proved, while (4.165) follows using Corollary 4.3. (qed)

The special case of Theorem 4.12 considered by Ellenberg (1973) supposed that A be symmetric and E nonsingular. It is clear that when the idempotent matrix A is symmetric then it is nonnegative definite and so (4.163) always holds. Moreover E nonsingular implies (4.163) even if $A \neq A'$. When E is symmetric idempotent but A is idempotent and not symmetric then (4.164) need not hold, for let

$$(4.169)\qquad A = \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

Then $A = A_{*}^2$ and

$$(4.170) \qquad E^- = \begin{pmatrix} 1 & a & b \\ c & d & e \\ f & g & h \end{pmatrix}$$

for arbitrary scalars a,b,c,d,e,f,g and h. Then

$$(4.171) \qquad (A/E) = \begin{pmatrix} 1 & 0 \\ -e & 1 \end{pmatrix}$$

is idempotent $\Leftrightarrow e = 0$. Moreover (A/E) is not invariant under choice of E^-.

A theorem by Milliken and Akdeniz (1977), which showed that if H and E − H are both symmetric nonnegative definite matrices then the difference between the Moore-Penrose g-inverses

$$(4.172) \qquad H^+ - E^+ \text{ is nnd} \Leftrightarrow r(E) = r(H),$$

has been extended by Styan and Pukelsheim (1978), who use symmetric reflexive g-inverses rather than Moore-Penrose g-inverses. See also Theorem 3.2.

THEOREM 4.13 (Styan & Pukelsheim, 1978). *Let* H *and* E − H *be symmetric nonnegative definite matrices. Then* E *is nonnegative definite and* $r(E,H) = r(E)$. *Let* E_r^- *and* H_r^- *be symmetric reflexive g-inverses. Then*

$$(4.173) \qquad H_r^- - E_r^- \textit{ is nnd} \Leftrightarrow EE_r^- = HH_r^-$$

and then $r(E - H) = r(H_r^- - E_r^-)$.

Proof. Using Theorem 4.7 it follows that

(4.174)
$$In\begin{pmatrix} E & H \\ H & H \end{pmatrix} = In H + In(E - H) ,$$

and so E is nnd and $r(E,H) = r(E)$ since the partitioned matrix in (4.174) is nnd. Moreover

(4.175)
$$In\begin{pmatrix} E & EE_r^- \\ E_r^- E & H_r^- \end{pmatrix} = In E + In(H_r^- - E_r^-)$$

using Theorem 4.7 again and the fact that $r(E,EE_r^-) = r(E)$.

Let $EE_r^- = HH_r^-$. Then

(4.176)
$$In\begin{pmatrix} E & EE_r^- \\ E_r^- E & H_r^- \end{pmatrix} = In\begin{pmatrix} E & HH_r^- \\ H_r^- H & H_r^- \end{pmatrix} = In H_r^- + In(E - H)$$

using Theorem 4.7 and choosing $(H_r^-)_r^- = H$. Thus (4.175) = (4.176) and $H_r^- - E_r^-$ is nnd and $r(E - H) = r(H_r^- - E_r^-)$.

Now, let $H_r^- - E_r^-$ be nonnegative definite. Then the partitioned matrix in (4.175) is nnd. Thus

(4.177)
$$r(E_r^- E, H_r^-) = r(H_r^-)$$

which, in turn, implies

(4.178)
$$[I - H_r^-(H_r^-)^-]E_r^- E = 0$$

so that, choosing $(H_r^-)^- = H$,

(4.179)
$$E_r^- E = H_r^- H E_r^- E = H_r^- H ,$$

since $r(E,H) = r(E)$. Transposing yields $EE_r^- = HH_r^-$. (qed)

COROLLARY 4.8 (Milliken & Akdeniz, 1977). *Let* H *and* E - H *be symmetric nonnegative definite matrices. Then*

$$H^+ - E^+ \text{ is nnd} \Leftrightarrow r(E) = r(H) . \tag{4.180}$$

Proof. If $H^+ - E^+$ is nonnegative definite then

$$EE^+ = HH^+ \tag{4.181}$$

follows from (4.173), and

$$r(E) = r(E^+E) = r(H^+H) = r(H) . \tag{4.182}$$

Now, suppose $r(E) = r(H)$. Then Theorem 4.13 implies that $r(E) = r(H) = r(E,H)$. Then

$$E = HH^+E \quad \text{and} \quad H = EE^+H . \tag{4.183}$$

So, postmultiplying the first equation in (4.183) by E^+ yields

(4.184a) $$EE^+ = HH^+EE^+ = (HH^+EE^+)'$$

(4.184b) $$= (EE^+)'(HH^+)'$$

(4.184c) $$= EE^+HH^+ = HH^+ ,$$

and the proof is complete. (qed)

Note that $r(E) = r(H)$ does not always imply that $H_r^- - E_r^-$ is nonnegative definite. For example, let

(4.185) $$E = H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

which is nonnegative definite so that

(4.186) $$E - H = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

(4.187) $$E_r^- = \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}$$

for some scalar x and

(4.188) $$H_r^- = \begin{pmatrix} 1 & y \\ y & y^2 \end{pmatrix}$$

for some scalar y. Hence

(4.189) $$H_r^- - E_r^- = \begin{pmatrix} 0 & y-x \\ y-x & y^2-x^2 \end{pmatrix}.$$

is nonnegative definite if and only if $x = y$. But

(4.190) $$EE_r^- = HH_r^- \Leftrightarrow \begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & y \\ 0 & 0 \end{pmatrix}$$

if and only if $x = y$. From this example, we conclude that, although $r(E) = r(H)$ does not always imply that $H_r^- - E_r^-$ is nonnegative definite, the condition $EE_r^- = HH_r^-$ always does.

In a study of the existence of a nonnegative definite matrix with prescribed characteristic roots, Fiedler (1974) based his proofs on a lemma, which Dias da Silva (1976) found "interesting enough" to report in full in *Mathematical Reviews*. A rather simple proof of this lemma is possible using Schur complements.

THEOREM 4.14 (Fiedler, 1974). *Let* A *be a symmetric* m×m *matrix with characteristic roots* $\alpha_1, \alpha_2, \ldots, \alpha_m$, *and let* $\underset{\sim}{u}$ *be a normalized characteristic vector corresponding to* α_1. *Let* B *be a symmetric* n×n *matrix with characteristic roots* $\beta_1, \beta_2, \ldots, \beta_n$, *and let* $\underset{\sim}{v}$ *be a normalized characteristic vector corresponding to* β_1. *Then, for any* γ, *the matrix*

$$M = \begin{pmatrix} A & \gamma \underset{\sim}{u}\underset{\sim}{v}' \\ \gamma \underset{\sim}{v}\underset{\sim}{u}' & B \end{pmatrix} \tag{4.191}$$

has characteristic roots $\alpha_2, \ldots, \alpha_m, \beta_2, \ldots, \beta_n$, *and the characteristic roots of*

$$\begin{pmatrix} \alpha_1 & \gamma \\ \gamma & \beta_1 \end{pmatrix}. \tag{4.192}$$

Proof. The characteristic polynomial

$$p = \begin{vmatrix} A - \lambda I & \gamma \underset{\sim}{u}\underset{\sim}{v}' \\ \gamma \underset{\sim}{v}\underset{\sim}{u}' & B - \lambda I \end{vmatrix} \tag{4.193a}$$

$$= |A - \lambda I| \cdot |B - \lambda I - \gamma \underset{\sim}{v}\underset{\sim}{u}'(A - \lambda I)^{-1} \gamma \underset{\sim}{u}\underset{\sim}{v}'| , \tag{4.193b}$$

using (2.4) and so

$$p = |A - \lambda I| \cdot |B - \lambda I - \gamma^2 \underset{\sim}{v}\underset{\sim}{u}'(\alpha_1 - \lambda)^{-1} \underset{\sim}{u}\underset{\sim}{v}'| , \tag{4.194}$$

since $(A - \lambda I)\underset{\sim}{u} = (\alpha_1 - \lambda)\underset{\sim}{u}$. Hence

(4.195) $$p = |A - \lambda I| . |B - \lambda I - \gamma^2 \underset{\sim}{v}\underset{\sim}{v}'/(\alpha_1 - \lambda)| ,$$

since $\underset{\sim}{u}'\underset{\sim}{u} = 1$, and so

(4.196a) $$p = |A - \lambda I| . |B - \lambda I| . |I - \gamma^2 (B - \lambda I)^{-1} \underset{\sim}{v}\underset{\sim}{v}'/(\alpha_1 - \lambda)|$$

(4.196b) $$= |A - \lambda I| . |B - \lambda I| . |I - \gamma^2 \underset{\sim}{v}\underset{\sim}{v}'/[(\alpha_1 - \lambda)(\beta_1 - \lambda)]|$$

since $(B - \lambda I)\underset{\sim}{v} = (\beta_1 - \lambda)\underset{\sim}{v}$. Hence

(4.197a) $$p = \begin{vmatrix} A - \lambda I & \gamma \underset{\sim}{u}\underset{\sim}{v}' \\ \gamma \underset{\sim}{v}\underset{\sim}{u}' & B - \lambda I \end{vmatrix}$$

(4.197b) $$= \prod_{i=2}^{m} (\alpha_i - \lambda) . \prod_{j=2}^{n} (\beta_j - \lambda) . [(\alpha_1 - \lambda)(\beta_1 - \lambda) - \gamma^2] .$$

(qed)

Theorem 4.14 may be used to find the characteristic roots of a special correlation matrix structure, cf. the remarks after (6.70).

CHAPTER V

NUMERICAL MATRIX INVERSION USING SCHUR COMPLEMENTS

A number of algorithms for matrix inversion use Schur complements. The earliest of these is probably the "bordering method" published in the book by Frazer, Duncan and Collar (1938).

§5.1 *The bordering method.*

Consider the matrix

$$(5.1)\qquad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

The method proposed by Frazer, Duncan and Collar (1938, p. 114) considers the principal leading submatrices

$$(5.2)\qquad E_i = \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix}, \quad i=1,2,\ldots,n,$$

$$(5.3)\qquad = \left(\begin{array}{ccc|c} & E_{i-1} & & \begin{matrix} a_{1i} \\ \cdots \\ a_{i-1,i} \end{matrix} \\ \hline a_{i1} & \cdots & a_{i,i-1} & a_{ii} \end{array}\right), \quad i=2,3,\ldots,n,$$

$$(5.4)\qquad = \begin{pmatrix} E_{i-1} & \underset{\sim}{f}_{i-1} \\ \underset{\sim}{g}'_{i-1} & a_{ii} \end{pmatrix}, \quad i=2,3,\ldots,n,$$

where $\underset{\sim}{f}_{i-1} = (a_{1i},\ldots,a_{i-1,i})'$ and $\underset{\sim}{g}'_{i-1} = (a_{i1},\ldots,a_{i,i-1})$; $E_1 \equiv a_{11}$. Assuming E_{i-1}^{-1} known, we can apply (2.37). Then

$$(5.5)\qquad E_i^{-1} = \begin{pmatrix} E_{i-1}^{-1} + E_{i-1}^{-1}\, \underset{\sim}{f}_{i-1}\, \underset{\sim}{g}'_{i-1}\, E_{i-1}^{-1}/s_{i-1} & -E_{i-1}^{-1}\, \underset{\sim}{f}_{i-1}/s_{i-1} \\ -\underset{\sim}{g}'_{i-1}\, E_{i-1}^{-1}/s_{i-1} & 1/s_{i-1} \end{pmatrix},$$

where $s_{i-1} = a_{ii} - \underset{\sim}{g}'_{i-1}\, E_{i-1}^{-1}\, \underset{\sim}{f}_{i-1}$; $i=2,3,\ldots,n$. Define $s_o \equiv a_{11}$. This method requires that all the E_i's $(i=1,2,\ldots,n)$ be nonsingular. Hence, if one or more of the E_i's is singular, $i=1,2,\ldots,n$ (i.e., when at least one Schur complement $s_{i-1} = 0$), then we can find a permutation matrix

$$(5.6)\qquad \Pi = (\underset{\sim}{e}_{i_1}, \underset{\sim}{e}_{i_2}, \ldots, \underset{\sim}{e}_{i_n}),$$

where $\{i_1, i_2, \ldots, i_n\}$ is a permutation of $\{1,2,\ldots,n\}$, such that all the principal submatrices in ΠA are nonsingular. Having obtained $(\Pi A)^{-1}$, we postmultiply by Π to obtain A^{-1}.

Example 5.1. To find the inverse of

$$(5.7)\qquad A = \begin{pmatrix} 5 & 3 & -3 \\ 2 & -4 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

we write

$$(5.8)\qquad A = \begin{pmatrix} E_2 & \underset{\sim}{f}_2 \\ \underset{\sim}{g}'_2 & a_{33} \end{pmatrix},$$

where

$$(5.9)\qquad E_2 = \begin{pmatrix} 5 & 3 \\ 2 & -4 \end{pmatrix}$$

$\underset{\sim}{f}_2 = (-3,4)'$, $\underset{\sim}{g}'_2 = (3,2)$ and $a_{33} = 1$.

To apply (5.5) we compute:

$$(5.10)\qquad E_2^{-1} = \frac{1}{26}\begin{pmatrix} 4 & 3 \\ 2 & -5 \end{pmatrix} = \begin{pmatrix} \frac{2}{13} & \frac{3}{26} \\ \frac{1}{13} & \frac{-5}{26} \end{pmatrix},$$

$$(5.11)\qquad E_2^{-1}\underset{\sim}{f}_2 = \frac{1}{26}\begin{pmatrix} 4 & 3 \\ 2 & -5 \end{pmatrix}\begin{pmatrix} -3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

$$(5.12)\qquad \underset{\sim}{g}_2' E_2^{-1} = \frac{1}{26}\,(3 \quad 2)\begin{pmatrix} 4 & 3 \\ 2 & -5 \end{pmatrix} = \frac{1}{26}\,(16 \quad -1)$$

$$= \left(\frac{8}{13} \quad \frac{-1}{26}\right),$$

$$(5.13)\qquad s_2 = a_{33} - \underset{\sim}{g}_2' E_2^{-1}\underset{\sim}{f}_2 = 1 - (3 \quad 2)\begin{pmatrix} 0 \\ -1 \end{pmatrix} = 3,$$

$$(5.14)\qquad E_2^{-1} + E_2^{-1}\underset{\sim}{f}_2\,\underset{\sim}{g}_2' E_2^{-1}/s_2 = \frac{1}{26}\begin{pmatrix} 4 & 3 \\ 2 & -5 \end{pmatrix} + \frac{1}{26}\begin{pmatrix} 0 \\ -1 \end{pmatrix}(16 \quad -1)/3$$

$$= \frac{1}{78}\begin{pmatrix} 12 & 9 \\ -10 & -14 \end{pmatrix} = \begin{pmatrix} \frac{2}{13} & \frac{3}{26} \\ -\frac{5}{39} & -\frac{7}{39} \end{pmatrix}.$$

Hence

$$(5.15)\qquad A^{-1} = \frac{1}{78}\begin{pmatrix} 12 & 9 & 0 \\ -10 & -14 & 26 \\ -16 & 1 & 26 \end{pmatrix} = \begin{pmatrix} \frac{2}{13} & \frac{3}{26} & 0 \\ -\frac{5}{39} & -\frac{7}{39} & \frac{1}{3} \\ -\frac{8}{39} & \frac{1}{78} & \frac{1}{3} \end{pmatrix}.$$

A variant of the above method was given by Jossa (1940), who showed that, when E_{i-1}^{-1} is known, then the following operations yield $E_i^{-1} = \{a_i^{k,l}\}$, $i=2,3,\ldots,n$. Define $a_1^{1,1} = 1/a_{11}$. For $i=2,3,\ldots,n$ set

$$r_{ki} = -\sum_{h=1}^{i-1} a_{i-1}^{k,h} a_{hi} \; ; \qquad k=1,2,\ldots,i-1, \tag{5.16}$$

$$a_i^{i,i} = 1/(a_{ii} + \sum_{h=1}^{i-1} a_{ih} r_{hi}) \; ; \tag{5.17}$$

$$a_i^{k,i} = a_i^{i,i} r_{ki} \; ; \qquad k=1,2,\ldots,i-1, \tag{5.18}$$

$$a_i^{i,\ell} = -a_i^{i,i}\left(\sum_{h=1}^{i-1} a_{ih} a_{i-1}^{h,\ell}\right); \qquad \ell=1,2,\ldots,i-1, \tag{5.19}$$

$$a_i^{k,\ell} = a_{i-1}^{k,\ell} + r_{ki} a_i^{i,\ell}; \qquad k,\ell=1,2,\ldots,i-1. \tag{5.20}$$

The above equations may be obtained simply by rewriting (5.5) in scalar notation.

Example 5.1 (reprise). Using Jossa's method, we obtain

$$a_1^{1,1} = 1/5$$

(5.16) $r_{12} = -a_1^{1,1} a_{12} = -3/5$

(5.17) $a_2^{2,2} = 1/(a_{22} + a_{21} r_{12}) = -5/26$

(5.18) $a_2^{1,2} = a_2^{2,2} r_{12} = 3/26$

(5.19) $a_2^{2,1} = -a_2^{2,2} a_{21} a_1^{1,1} = 2/26$

(5.20) $a_2^{1,1} = a_1^{1,1} + r_{12} a_2^{2,1} = 4/26$

(5.16) $\begin{cases} r_{13} = -(a_2^{1,1} a_{13} + a_2^{1,2} a_{23}) = 0 \\ r_{23} = -(a_2^{2,1} a_{13} + a_2^{2,2} a_{23}) = 1 \end{cases}$

(5.17) $a_3^{3,3} = 1/(a_{33} + a_{31} r_{13} + a_{32} r_{23}) = 1/3$

(5.18) $\begin{cases} a_3^{1,3} = a_3^{3,3} r_{13} = 0 \\ a_3^{2,3} = a_3^{3,3} r_{23} = 1/3 \end{cases}$

(5.19) $\begin{cases} a_3^{3,1} = -a_3^{3,3}(a_{31} a_2^{1,1} + a_{32} a_2^{2,1}) = -16/78 \\ a_3^{3,2} = -a_3^{3,3}(a_{31} a_2^{1,2} + a_{32} a_2^{2,2}) = 1/78 \end{cases}$

(5.20) $\begin{cases} a_3^{1,1} = a_2^{1,1} + r_{13} a_3^{3,1} = 4/26 = 12/78 \\ a_3^{1,2} = a_2^{1,2} + r_{13} a_3^{3,2} = 3/26 = 9/78 \\ a_3^{2,1} = a_2^{2,1} + r_{23} a_3^{3,1} = -10/78 \\ a_3^{2,2} = a_2^{2,2} + r_{23} a_3^{3,2} = -14/78 \end{cases}$.

Hence, we obtain (5.15).

Louis Guttman (1946) called this bordering method "first order enlargement" in view of the partitioning (5.3) adding a single row and column to E_{i-1}. We now consider the partitioning with E_{i-2} bordered by 2 rows and 2 columns:

$$(5.21)\qquad E_i = \left(\begin{array}{ccc|cc} & & & a_{1,i-1} & a_{1,i} \\ & E_{i-2} & & \cdots\cdots & \cdots\cdots \\ & & & a_{i-2,i-1} & a_{i-2,i} \\ \hline a_{i-1,1} & \cdots & a_{i-1,i-2} & a_{i-1,i-1} & a_{i-1,i} \\ a_{i,1} & \cdots & a_{i,i-2} & a_{i,i-1} & a_{i,i} \end{array}\right)$$

$$= \begin{pmatrix} E_{i-2} & F_{i-2} \\ G_{i-2} & H_{i-2} \end{pmatrix}; \qquad i = 3,4,\ldots,n\ ,$$

where

$$(5.22)\qquad F_{i-2} = \begin{pmatrix} a_{1,i-1} & a_{1,i} \\ \vdots & \vdots \\ a_{i-2,i-1} & a_{i-2,i} \end{pmatrix} \qquad : (i-2) \times 2$$

$$(5.23)\qquad G_{i-2} = \begin{pmatrix} a_{i-1,1} & \cdots & a_{i-1,i-2} \\ a_{i,1} & \cdots & a_{i,i-2} \end{pmatrix} \qquad : \ 2 \times (i-2)$$

$$(5.24)\qquad H_{i-2} = \begin{pmatrix} a_{i-1,i-1} & a_{i-1,i} \\ a_{i,i-1} & a_{ii} \end{pmatrix} \qquad : \ 2 \times 2.$$

Assuming E_{i-2}^{-1} known, we can apply (2.37).

Example 5.2. To find the inverse of the matrix A defined by (5.7) using the method described above we write

$$A = \begin{pmatrix} e_1 & \underset{\sim}{f}_1' \\ \underset{\sim}{g}_1 & H_1 \end{pmatrix}, \tag{5.25}$$

where $e_1 = 5$, $\underset{\sim}{f}_1' = (3,-3)$, $\underset{\sim}{g}_1 = (2,3)'$ and

$$H_1 = \begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix}. \tag{5.26}$$

The Schur complement $S = (A/e_1)$ is given by

$$S = \begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 2 \\ 3 \end{pmatrix} (3 \quad -3) = \begin{pmatrix} -\frac{26}{5} & \frac{26}{5} \\ \frac{1}{5} & \frac{14}{5} \end{pmatrix}. \tag{5.27}$$

Applying (2.37), we obtain

$$A^{-1} = \begin{pmatrix} 1/e_1 + (1/e_1^2)\underset{\sim}{f}_1' S^{-1} \underset{\sim}{g}_1 & -\underset{\sim}{f}_1' S^{-1}/e_1 \\ -S^{-1}\underset{\sim}{g}_1/e_1 & S^{-1} \end{pmatrix}. \tag{5.28}$$

To apply (5.28) we compute:

$$S^{-1} = \frac{-5}{78} \begin{pmatrix} \frac{14}{5} & -\frac{26}{5} \\ -\frac{1}{5} & -\frac{26}{5} \end{pmatrix} = \begin{pmatrix} -\frac{7}{39} & \frac{1}{3} \\ \frac{1}{78} & \frac{1}{3} \end{pmatrix} \tag{5.29}$$

$$-\underset{\sim}{f}_1' S^{-1}/e_1 = \frac{1}{78} (3 \quad -3) \begin{pmatrix} \frac{14}{5} & -\frac{26}{5} \\ -\frac{1}{5} & -\frac{26}{5} \end{pmatrix} = (\tfrac{9}{78} \quad 0) = (\tfrac{3}{26} \quad 0) \tag{5.30}$$

$$-S^{-1}\underset{\sim}{g}_1/e_1 = \frac{1}{78} \begin{pmatrix} \frac{14}{5} & -\frac{26}{5} \\ -\frac{1}{5} & -\frac{26}{5} \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -\frac{10}{78} \\ -\frac{16}{78} \end{pmatrix} = \begin{pmatrix} -\frac{5}{39} \\ -\frac{8}{39} \end{pmatrix} \tag{5.31}$$

$$1/e_1 + (1/e_1^2)\underset{\sim}{f}_1' S^{-1} \underset{\sim}{g}_1 = \frac{1}{5} - \frac{1}{5} (\tfrac{9}{78} \quad 0) \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \frac{12}{78} = \frac{2}{13}. \tag{5.32}$$

Hence, we obtain (5.15).

§5.2 *Geometric enlargement*.

The method of "geometric enlargement" due to Louis Guttman (1946) allows the inverse of the matrix A in (5.1) to be obtained by successively constructing the inverses of the principal submatrices

$$E_{2i} = \begin{pmatrix} a_{11} & \cdots & a_{1,2i} \\ \cdots & \cdots & \cdots \\ a_{2i,1} & \cdots & a_{2i,2i} \end{pmatrix} \quad : \; 2i \times 2i \tag{5.33}$$

$$= \begin{pmatrix} E_i & F_i \\ G_i & H_i \end{pmatrix} \quad ; \; i=1,2,\ldots,[\tfrac{n}{2}] . \tag{5.34}$$

Assuming E_i^{-1} is known, we can apply (2.37). If $i=1$, then

$$E_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \; ; \tag{5.35}$$

setting $E_1 = a_{11}$, we see that "geometric enlargement" reduces to "first-order enlargement". Also, if $i=2$, we have

$$E_4 = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \tag{5.36}$$

and E_2 is defined by (5.35); here, the "geometric enlargement" reduces to "second-order enlargement".

§5.3 *Partitioned Schur complements*.

We begin by partitioning the $n \times n$ nonsingular matrix

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} , \tag{5.37}$$

where E is $n_o \times n_o$, nonsingular, and readily invertible (e.g., $n_o = 1$ or 2, E diagonal). We then compute the Schur complement $S = (A/E)$ and if it is easily invertible, then we compute A^{-1} using the Schur-Banachiewicz formula (2.37). Otherwise we partition the Schur complement

$$S = \begin{pmatrix} E_1 & F_1 \\ G_1 & H_1 \end{pmatrix}, \tag{5.38}$$

where E_1 is $n_1 \times n_1$, nonsingular, and readily invertible. We now compute the Schur complement $S_1 = (S/E_1)$, and if it is easily invertible then we compute S^{-1} using (2.37), from which A^{-1} follows using (2.37) again. Otherwise we partition

$$S_1 = \begin{pmatrix} E_2 & F_2 \\ G_2 & H_2 \end{pmatrix}, \tag{5.39}$$

where E_2 is $n_2 \times n_2$, nonsingular, and readily invertible. We compute $S_2 = (S_1/E_2)$ and repeat the procedure performed with S_1. And so on. Writing

$$S_k = \begin{pmatrix} E_{k+1} & F_{k+1} \\ G_{k+1} & H_{k+1} \end{pmatrix} \quad ; \; k=1,2,\ldots,m-1, \tag{5.40}$$

the forward part of this algorithm stops at $k = m$ when $S_m = (S_{m-1}/E_m)$ is easily invertible. At most $m = n-2$. Some $m+1$ E-matrices will have been inverted. We now invert S_m and proceed backwards computing in turn each inverse S_k^{-1}; $k=m-1,m-2,\ldots,3,2,1,0$, using S_{k+1}^{-1} and (2.37), with $S = S_o$. A^{-1} follows using (2.37) again. Louis Guttman (1946) sketched the above algorithm with $n_k = 1$ or 2; $k=0,1,\ldots,m$. Zlobec and Chan (1974) gave full details with all $n_k = 1$; they also state that their "program in APL consists of only seven lines".

Example 5.3. Find the inverse of

$$(5.41)\qquad A = \begin{pmatrix} 5 & 3 & -3 \\ 2 & -4 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

using partitioned Schur complements.

The forward part.

We may write

$$(5.42)\qquad A = \begin{pmatrix} e & \underset{\sim}{f}' \\ \underset{\sim}{g} & H \end{pmatrix}$$

where $e = 5(\neq 0)$, $\underset{\sim}{f}' = (3,-3)$, $\underset{\sim}{g}' = (2,3)$ and

$$(5.43)\qquad H = \begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix} .$$

$$(5.44)\qquad S = \begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix} - \frac{1}{5}\begin{pmatrix} 2 \\ 3 \end{pmatrix} (3,-3)$$

$$(5.45)\qquad = \begin{pmatrix} -\frac{26}{5} & \frac{26}{5} \\ \frac{1}{5} & \frac{14}{5} \end{pmatrix} .$$

Partition S as follows

$$(5.46)\qquad S = \begin{pmatrix} e_1 & f_1 \\ g_1 & h_1 \end{pmatrix}$$

where $e_1 = -\frac{26}{5}$ $(\neq 0)$, $f_1 = \frac{26}{5}$, $g_1 = \frac{1}{5}$, $h_1 = \frac{14}{5}$.

Hence,

$$(5.47)\qquad s_1 = 14/5 - (1/5)(-26/5)^{-1}(26/5) = 3.$$

The backward part.

$1/s_1 = 1/3$. Applying (2.37), we obtain

(5.48)
$$S^{-1} = \begin{pmatrix} -\frac{14}{78} & \frac{1}{3} \\ \frac{1}{78} & \frac{1}{3} \end{pmatrix} .$$

Applying (2.37) again, we obtain (5.15).

§5.4 *Rank annihilation.*

We express the $n \times n$ nonsingular matrix A as the sum of a nonsingular matrix D and the sum of h matrices each of rank one (cf. Wilf, 1959),

(5.49)
$$A = D + \sum_{i=1}^{h} \underset{\sim}{f}_i \underset{\sim}{g}_i' .$$

The matrix D is easily invertible, e.g., diagonal. Clearly $h \leq n$. Let us write

(5.50)
$$\begin{cases} E_0 = D \\ E_1 = D + \underset{\sim}{f}_1 \underset{\sim}{g}_1' \\ \vdots \\ E_j = E_{j-1} + \underset{\sim}{f}_j \underset{\sim}{g}_j' ; \quad j = 1,2,\ldots,h-1 \\ \vdots \\ E_h = A . \end{cases}$$

Then we compute, in turn, $E_0^{-1}, E_1^{-1}, \ldots, E_h^{-1} = A^{-1}$ using (2.59),

(5.51)
$$E_j^{-1} = E_{j-1}^{-1} - E_{j-1}^{-1} \underset{\sim}{f}_j \underset{\sim}{g}_j' E_{j-1}^{-1} / (1 + \underset{\sim}{g}_j' E_{j-1}^{-1} \underset{\sim}{f}_j) ,$$

where

(5.52) $$1 + \underset{\sim}{g}_j' E_{j-1}^{-1} \underset{\sim}{f}_j \neq 0 \; ; \qquad j = 1,\ldots,h \; .$$

This method requires that all the E_j's ($j = 0,1,\ldots,h$) be nonsingular; $A^{-1} = E_h^{-1}$.

Edelblute (1966) considered the special case of (5.49) with $D = I$, $\underset{\sim}{f}_j = (A - I)\underset{\sim}{e}_j$, $\underset{\sim}{g}_i = \underset{\sim}{e}_i$, and $h = n$. Then

(5.53) $$\begin{cases} A = I + \sum_{j=1}^{n} (A - I)\underset{\sim}{e}_j \underset{\sim}{e}_j' \\ E_0 = I \\ E_1 = I + (A - I)\underset{\sim}{e}_1 \underset{\sim}{e}_1' \\ \vdots \\ E_j = E_{j-1} + (A - I)\underset{\sim}{e}_j \underset{\sim}{e}_j' \; ; \qquad j = 1,\ldots,n \\ \vdots \\ E_n = A \; . \end{cases}$$

Hence

(5.54a) $$E_1^{-1} = I - (A - I)\underset{\sim}{e}_1 \underset{\sim}{e}_1' / a_{11} \; ,$$

(5.54b) $$E_j^{-1} = E_{j-1}^{-1} - E_{j-1}^{-1}(A - I)\underset{\sim}{e}_j \underset{\sim}{e}_j' E_{j-1}^{-1} \,/\, (1 + \underset{\sim}{e}_j' E_{j-1}^{-1}(A - I)\underset{\sim}{e}_j) \; , \quad j = 2,\ldots,n.$$

Example 5.4. Find the inverse of

(5.55) $$A = \begin{pmatrix} 5 & 3 & -3 \\ 2 & -4 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

by rank annihilation.

We write

$$A = I + \sum_{i=1}^{3} f_i e_i' , \tag{5.56}$$

where $f_1 = (4,2,3)'$, $f_2 = (3,-5,2)'$, and $f_3 = (-3,4,0)'$. Then

$$E_1^{-1} = I - \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix} (1 \quad 0 \quad 0)/5 \tag{5.57}$$

$$= \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ -\frac{2}{5} & 1 & 0 \\ -\frac{3}{5} & 0 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 5 & 0 \\ -3 & 0 & 5 \end{pmatrix} . \tag{5.58}$$

Thus, using (5.51),

$$E_2^{-1} = E_1^{-1} - E_1^{-1}(A - I) e_2 e_2' E_1^{-1}/a , \tag{5.59}$$

where

$$a = 1 + (-\tfrac{2}{5}, 1, 0) \begin{pmatrix} 3 \\ -5 \\ 2 \end{pmatrix} = -\frac{26}{5} \tag{5.60}$$

so that

$$E_2^{-1} = E_1^{-1} + \frac{5}{26} E_1^{-1} \begin{pmatrix} 3 \\ -5 \\ 2 \end{pmatrix} (-\tfrac{2}{5}, 1, 0) = \tag{5.61}$$

(5.62)
$$= E_1^{-1} + \frac{1}{26}\begin{pmatrix} 3 \\ -31 \\ 1 \end{pmatrix}(-\tfrac{2}{5},1,0)$$

(5.63)
$$= \frac{1}{130}\begin{pmatrix} 26 & 0 & 0 \\ -52 & 130 & 0 \\ -78 & 0 & 130 \end{pmatrix} + \frac{1}{130}\begin{pmatrix} -6 & 15 & 0 \\ 62 & -155 & 0 \\ -2 & 5 & 0 \end{pmatrix}$$

(5.64)
$$= \frac{1}{26}\begin{pmatrix} 4 & 3 & 0 \\ 2 & -5 & 0 \\ -16 & 1 & 26 \end{pmatrix} .$$

Hence

(5.65)
$$A^{-1} = E_3^{-1} = E_2^{-1} - E_2^{-1}\begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix}(0,0,1)E_2^{-1}/b ,$$

where

(5.66)
$$b = 1 + (0,0,1)E_2^{-1}\begin{pmatrix} -3 \\ 4 \\ 0 \end{pmatrix} = 3 .$$

Thus

(5.67)
$$A^{-1} = E_2^{-1} - \frac{1}{78}\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}(-16,1,26) =$$

$$(5.68)\qquad = \frac{1}{78}\begin{pmatrix} 12 & 9 & 0 \\ -10 & -14 & 26 \\ -16 & 1 & 26 \end{pmatrix},$$

cf. (5.15).

§5.5 *Complex matrices.*

Let the $n\times n$ complex matrix $E + iF$, where E, F are both real matrices, be nonsingular, and let us write its inverse as $K + iL$, where K and L are real. Then, cf. Newman (1962), Westlake (1968, p. 31),

$$(5.69)\qquad \begin{pmatrix} E & F \\ -F & E \end{pmatrix}\begin{pmatrix} K & L \\ -L & K \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Thus, we note that $E + iF$ is nonsingular if and only if

$$(5.70)\qquad A = \begin{pmatrix} E & F \\ -F & E \end{pmatrix}$$

is nonsingular. If E is nonsingular, then $E + iF$ is nonsingular if and only if the Schur complement $(A/E) = E + FE^{-1}F$ is nonsingular, cf. (2.4), and then

$$(5.71)\qquad K = (E + FE^{-1}F)^{-1}, \quad L = -E^{-1}F(E + FE^{-1}F)^{-1},$$

cf. (2.37).

If E is singular we may rearrange the columns of A in order to obtain a submatrix in the top left-hand corner which is nonsingular. This is possible since (5.69) implies that $r(E,F) = n$. But by this rearrangement, the nice pattern in (5.70) would usually be lost.

§5.6 *Generalized inversion by partitioned Schur complements.*

We begin as in §5.3 by partitioning the rectangular or singular matrix

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix}, \tag{5.72}$$

where E is $n_0 \times n_0$, nonsingular, and readily invertible (e.g., $n_0 = 1$ or 2, E diagonal). We then compute the Schur complement $S = (A/E)$ and if it is easy to find a g-inverse of S, then we compute (cf. Zlobec, 1975) a g-inverse A^- using, cf. (4.28) and (4.61).

THEOREM 5.1. *If*

$$A = \begin{pmatrix} E & F \\ G & H \end{pmatrix} \tag{5.73}$$

and E *is nonsingular then*

$$\begin{pmatrix} E^{-1} + E^{-1}FS^-GE^{-1} & -E^{-1}FS^- \\ -S^-GE^{-1} & S^- \end{pmatrix} = A^- \tag{5.74}$$

where $S = H - GE^{-1}F$.

Otherwise we partition the Schur complement

$$S = \begin{pmatrix} E_1 & F_1 \\ G_1 & H_1 \end{pmatrix}, \tag{5.75}$$

where E_1 is $n_1 \times n_1$, nonsingular, and readily invertible. We now compute the Schur complement $S_1 = (S/E_1)$, and if it is easy to find an S_1^-

then we compute S^- using Theorem 5.1, from which A^- follows using Theorem 5.1 again. Otherwise we partition

$$(5.76)\qquad S_1 = \begin{pmatrix} E_2 & F_2 \\ G_2 & H_2 \end{pmatrix},$$

where E_2 is $n_2 \times n_2$, nonsingular, and readily invertible. We compute $S_2 = (S_1/E_2)$ and repeat the procedure performed with S_1. And so on. Writing

$$(5.77)\qquad S_k = \begin{pmatrix} E_{k+1} & F_{k+1} \\ G_{k+1} & H_{k+1} \end{pmatrix}; \quad k = 1,2,\ldots,r-1 ,$$

the forward part of this algorithm stops at $k = r$ when $S_r = (S_{r-1}/E_r)$ has a g-inverse S_r^- which is easy to find. At most $r = r(A)$. We now proceed backwards computing in turn each g-inverse S_k^-; $k = r-1, r-2, \ldots, 3,2,1,0$, using S_{k+1}^- and Theorem 5.1, with $S = S_0$. A^- follows using Theorem 5.1 again.

Example 5.5. Find a g-inverse of

$$(5.78)\qquad A = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix}$$

using partitioned Schur complements. The matrix (5.78) is clearly singular and has rank 2. We may partition (5.78) as

$$(5.79)\qquad A = \begin{pmatrix} e & \underset{\sim}{f}' \\ \underset{\sim}{g} & H \end{pmatrix},$$

where $e = 2$, $\underset{\sim}{f}' = (1,2)$, $\underset{\sim}{g} = (2,0)'$ and

(5.80) $$H = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

Then

(5.81) $$S = (A/E) = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 2 \\ 0 \end{pmatrix}(1,2) = \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}.$$

Noting that

(5.82) $$S = \begin{pmatrix} -1 \\ 1 \end{pmatrix}(1 \quad 0),$$

we see at once that

(5.83) $$\frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}(-1 \quad 1) = S^+,$$

the Moore-Penrose g-inverse of S. Hence we use Theorem 5.1 to compute

(5.84) $$A^- = \begin{pmatrix} \frac{1}{2} + \frac{1}{4}(1,2)S^+(2,0)' & -\frac{1}{2}(1,2)S^+ \\ -\frac{1}{2}S^+(2,0)' & S^+ \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}.$$

If we had not noticed the factorization (5.82) we could partition (5.81) as

$$S = \left(\begin{array}{c|c} -1 & 0 \\ \hline 1 & 0 \end{array}\right) \tag{5.85}$$

and compute, using Theorem 5.1, the g-inverse

$$S^{\sim} = \left(\begin{array}{c|c} -1 & 0 \\ \hline 1 & 1 \end{array}\right). \tag{5.86}$$

Hence, again using Theorem 5.1, we obtain the alternate g-inverse

$$A^{\sim} = \begin{pmatrix} \frac{1}{2} + \frac{1}{4}(1,2)S^{\sim}(2,0)' & -\frac{1}{2}(1,2)S^{\sim} \\ -\frac{1}{2}S^{\sim}(2,0)' & S^{\sim} \end{pmatrix} \tag{5.87}$$

$$= \begin{pmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}.$$

A third g-inverse of A may be found by noting that any g-inverse of S must have the form

$$S^{-} = \begin{pmatrix} a & 1+a \\ b & c \end{pmatrix}, \tag{5.88}$$

where a, b and c are arbitrary scalars. Thus $a = -1/2$ and $b = c = 0$ yields S^{+}, while $a = -1$ and $b = c = 1$ yields $S^{\sim}$. Letting $a = b = c = 0$

we obtain

$$S^- = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{5.89}$$

and using Theorem 5.1 again yields

$$A^- = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.90}$$

§5.7 *Generalized inversion by rank annihilation.*

When $1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{f} \neq 0$,

$$(E + \underset{\approx}{fg}')^{-1} = E^{-1} - E^{-1}\underset{\approx}{fg}'E^{-1}/(1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{f}) , \tag{5.91}$$

cf. (2.59) and (5.51), was used repeatedly in §5.4 to find the inverse by rank annihilation. When $1 + \underset{\sim}{g}'E^{-1}\underset{\sim}{f} = 0$, it follows that (cf. Ahsanullah & Rahman, 1973) E^{-1} is a g-inverse of $E + \underset{\approx}{fg}'$, as is easily verified.

We express the $n{\times}n$ matrix A as in (5.53) and we write

$$A = I + (A - I)\begin{pmatrix} I_a & 0 \\ 0 & 0 \end{pmatrix} + (A - I)\begin{pmatrix} 0 & 0 \\ 0 & I_{n-a} \end{pmatrix} \tag{5.92}$$

with

$$E_a = I + (A - I)\begin{pmatrix} I_a & 0 \\ 0 & 0 \end{pmatrix} = \tag{5.93a}$$

(5.93b) $$= I + \begin{pmatrix} E - I_a & F \\ G & H - I_{n-a} \end{pmatrix} \begin{pmatrix} I_a & 0 \\ 0 & 0 \end{pmatrix}$$

(5.93c) $$= \begin{pmatrix} E & 0 \\ G & I_{n-a} \end{pmatrix},$$

where A is partitioned as usual, cf. e.g. (2.2), with E $a \times a$. Then, cf. Ahsanullah & Rahman (1973, p.3),

(5.94) $$|E_a| = 0 ; \quad r(A) < a < n-1 ?$$

To prove (5.94) we note that

(5.95a) $$r(E_a) = r(E) + n - a$$

(5.95b) $$\le r(A) + n - a$$

(5.95c) $$< a + n - a = n .$$

If $r(A) = n - 1$ and E_j is nonsingular for all j such that $1 \le j \le n - 1$ then $E_{n-1}^{-1} = A^-$. More generally, let

(5.96) $$A = D + F_1G_1' + F_2G_2' ,$$

where F_1 and G_1 are $n \times r$, F_2 and G_2 are $n \times (n-r)$ and $r(A) = r$. If $D + F_1G_1'$ ($= E_r$ say) is nonsingular then

(5.97) $$(D + F_1G_1')^{-1} = A^- .$$

To prove this, consider

(5.98) $$M = \begin{pmatrix} D + F_1G_1' & F_2 \\ G_2' & -I_{n-r} \end{pmatrix}.$$

Then

$$(5.99a) \qquad r(M) = n - r + r(D + F_1G_1' + F_2G_2')$$

$$(5.99b) \qquad = n - r + r(A) = n$$

$$(5.99c) \qquad = r(D + F_1G_1') + r(-I_{n-r} - G_2'(D + F_1G_1')^{-1}F_2)$$

$$(5.99d) \qquad = n + r(I_{n-r} + G_2'(D + F_1G_1')^{-1}F_2) \ .$$

Thus

$$(5.100) \qquad I_{n-r} + G_2'(D + F_1G_1')^{-1}F_2 = 0 \ .$$

Now let $E_r = D + F_1G_1'$. Then

$$(5.101a) \qquad AE_r^{-1}A = (E_r + F_2G_2')E_r^{-1}(E_r + F_2G_2')$$

$$(5.101b) \qquad = (I_r + F_2G_2'E_r^{-1})(E_r + F_2G_2')$$

$$(5.101c) \qquad = E_r + F_2G_2' + F_2G_2' + F_2G_2'E_r^{-1}F_2G_2' \ .$$

It follows that

$$(5.102) \qquad AE_r^{-1}A = A \Leftrightarrow F_2(I + G_2'E_r^{-1}F_2)G_2' = 0 \ ,$$

which is implied by (5.100). Hence the proof is complete.

Thus if

$$(5.103) \qquad A = D + \sum_{i=1}^{r} \underset{\sim}{f}_i \underset{\sim}{g}_i' + \sum_{i=r+1}^{n} \underset{\sim}{f}_i \underset{\sim}{g}_i'$$

so that $D + \sum_{i=1}^{r} \underset{\sim}{f}_i \underset{\sim}{g}_i'$ is nonsingular and $r = r(A)$ then

$$(5.104) \qquad (D + \sum_{i=1}^{r} \underset{\sim}{f}_i \underset{\sim}{g}_i')^{-1} = A^{-} \ .$$

If $D = I$, $\underset{\sim}{f}_i = (A - I)\underset{\sim}{e}_i$, $\underset{\sim}{g}_i = \underset{\sim}{e}_i$, (5.103) may not be possible. For example, if

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{pmatrix}, \quad \text{rank} = 2 \tag{5.105}$$

$$= I + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \underset{\sim}{e}_1' + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \underset{\sim}{e}_2' + \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \underset{\sim}{e}_3' . \tag{5.106}$$

We may write

$$E_1 = \begin{pmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \tag{5.107}$$

and its inverse

$$E_1^{-1} = I - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \underset{\sim}{e}_1'/2 \tag{5.108}$$

$$= \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -1 & 1 & 0 \\ -\frac{3}{2} & 0 & 1 \end{pmatrix} \tag{5.109}$$

using (5.91). Then $\underset{\sim}{g}_i' E_1^{-1} \underset{\sim}{f}_i = -1$; $i = 2,3$, as

(5.110) $$\underset{\sim}{e}_2' E_1^{-1}(A - I)\underset{\sim}{e}_2 = (-1,1,0)\begin{pmatrix}1\\0\\0\end{pmatrix} = -1$$

and

(5.111) $$\underset{\sim}{e}_3' E_1^{-1}(A - I)\underset{\sim}{e}_3 = (-\tfrac{3}{2},0,1)\begin{pmatrix}2\\2\\2\end{pmatrix} = -1 .$$

Moreover,

(5.112) $$\underset{\sim}{e}_2' E_1^{-1}(A - I)\underset{\sim}{e}_3 = (-1,1,0)\begin{pmatrix}2\\2\\2\end{pmatrix} = 0$$

and

(5.113) $$\underset{\sim}{e}_3' E_1^{-1}(A - I)\underset{\sim}{e}_2 = (-\tfrac{3}{2},0,1)\begin{pmatrix}1\\0\\0\end{pmatrix} = -\tfrac{3}{2} \neq 0 .$$

Hence $E_1^{-1} \neq A^-$. In fact,

(5.114) $$AE_1^{-1}A = \begin{pmatrix}2 & 1 & 2\\2 & 1 & 2\\3 & 0 & 3\end{pmatrix}\begin{pmatrix}\tfrac{1}{2} & 0 & 0\\-1 & 1 & 0\\-\tfrac{3}{2} & 0 & 1\end{pmatrix}A$$

and so

$$AE_1^{-1}A = \begin{pmatrix} -3 & 1 & 2 \\ -3 & 1 & 2 \\ -3 & 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 2 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{pmatrix} \tag{5.115}$$

$$= \begin{pmatrix} 2 & -2 & 2 \\ 2 & -2 & 2 \\ 3 & -3 & 3 \end{pmatrix} \neq A = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{pmatrix},$$

cf. (5.105).

CHAPTER VI

STATISTICAL APPLICATIONS OF SCHUR COMPLEMENTS

The Schur complement arises in a number of different areas of mathematical statistics. As observed by Cottle (1974, p. 192) "the multivariate normal distribution provides a magnificent example of how the Schur complement arises naturally".

§6.1 *The multivariate normal distribution.*

Let the random vector

$$\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \end{pmatrix} \tag{6.1}$$

follow a p-variate normal distribution with mean vector

$$\underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix} \tag{6.2}$$

and covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} , \tag{6.3}$$

where Σ_{22} is positive definite. Then the conditional distribution of $\underset{\sim}{x}_1$ given $\underset{\sim}{x}_2$ is multivariate normal with mean vector

$$\underset{\sim}{\nu}_1 = \underset{\sim}{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2) \tag{6.4}$$

and covariance matrix the Schur complement of Σ_{22} in Σ,

$$(\Sigma/\Sigma_{22}) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} . \tag{6.5}$$

To prove this result we note first that the joint distribution of

$$\begin{pmatrix} \underset{\sim}{x}_1 - \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{x}_2 \\ \underset{\sim}{x}_2 \end{pmatrix} \tag{6.6}$$

is multivariate normal with mean vector

$$\begin{pmatrix} \underset{\sim}{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{\mu}_2 \\ \underset{\sim}{\mu}_2 \end{pmatrix} \tag{6.7}$$

and covariance matrix

$$\begin{pmatrix} (\Sigma/\Sigma_{22}) & 0 \\ 0 & \Sigma_{22} \end{pmatrix} , \tag{6.8}$$

cf. (2.26). Hence $\underset{\sim}{x}_1 - \Sigma_{12}\Sigma_{22}^{-1}\underset{\sim}{x}_2$ is distributed independently of $\underset{\sim}{x}_2$, and so its conditional distribution given $\underset{\sim}{x}_2$ is the same as its unconditional distribution. Thus $\underset{\sim}{x}_1$ given $\underset{\sim}{x}_2$ is multivariate normal with mean vector (6.4) and covariance matrix (6.5).

Consider now the density function of the multivariate normal distribution

$$\phi(\underset{\sim}{x}) = (2\pi)^{-\frac{1}{2}p}|\Sigma|^{-\frac{1}{2}} \, exp\{-\tfrac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\} , \tag{6.9}$$

cf. Anderson (1958, p. 17). Then the above result concerning the conditional distribution of $\underset{\sim}{x}_1$ given $\underset{\sim}{x}_2$ yields

$$\phi(\underset{\sim}{x}) = \phi(\underset{\sim}{x}_1|\underset{\sim}{x}_2)\ \phi(\underset{\sim}{x}_2) \quad ; \tag{6.10}$$

thus

$$|\Sigma| = |(\Sigma/\Sigma_{22})| . |\Sigma_{22}| , \tag{6.11}$$

cf. (2.6), and

$$(6.12)\qquad (\underset{\sim}{x}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu}) = (\underset{\sim}{x}_1-\underset{\sim}{\nu}_1)'(\Sigma/\Sigma_{22})^{-1}(\underset{\sim}{x}_1-\underset{\sim}{\nu}_1) + (\underset{\sim}{x}_2-\underset{\sim}{\mu}_2)'\Sigma_{22}^{-1}(\underset{\sim}{x}_2-\underset{\sim}{\mu}_2) \ .$$

To verify (6.12) directly we use (2.40) to write

$$(6.13)\qquad \begin{pmatrix} I \\ -\Sigma_{22}^{-1}\Sigma_{21} \end{pmatrix}(\Sigma/\Sigma_{22})^{-1}(I, \ -\Sigma_{12}\Sigma_{22}^{-1}) + \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \ .$$

Substituting (6.13) into the left-hand side of (6.12) yields the right-hand side directly, since

$$(6.14)\qquad (I, \ -\Sigma_{12}\Sigma_{22}^{-1})(\underset{\sim}{x}-\underset{\sim}{\mu}) = \underset{\sim}{x}_1-\underset{\sim}{\nu}_1 \ .$$

When Σ_{22} is positive semidefinite and singular, the covariance matrix of Σ is also singular, cf. Corollary 4.5, and so $\underset{\sim}{x}$ does not have a density function. Using generalized inverses, however, we may evaluate (cf. Rao, 1973, pp. 522-523) the joint distribution of

$$(6.15)\qquad \begin{pmatrix} \underset{\sim}{x}_1 - \Sigma_{12}\Sigma_{22}^{-}\underset{\sim}{x}_2 \\ \underset{\sim}{x}_2 \end{pmatrix} ,$$

cf. (6.6), as multivariate normal with mean vector

$$(6.16)\qquad \begin{pmatrix} \underset{\sim}{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-}\underset{\sim}{\mu}_2 \\ \underset{\sim}{\mu}_2 \end{pmatrix}$$

and covariance matrix (6.8), where

$$(6.17)\qquad (\Sigma/\Sigma_{22}) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21}$$

is the generalized Schur complement of Σ_{22} in Σ ; cf. (4.2). From (4.52) we see that (6.17) is unique for all choices of g-inverse Σ_{22}^{-} . It should be noted that

$$(6.18)\qquad \Sigma_{22}\Sigma_{22}^{-}\Sigma_{12} = \Sigma_{12}$$

is needed to establish that the off-diagonal blocks in (6.8) are still 0 . The equation (6.18) is equivalent to

(6.19) $$r(\Sigma_{12},\Sigma_{22}) = r(\Sigma_{22}) ,$$

in view of Lemma 4.1, and (6.19) holds because of the nonnegative definiteness of Σ . It is interesting to note that (6.19) is just the condition for consistency of

(6.20) $$A\Sigma_{22} = \Sigma_{12} ,$$

which is analogous to the "normal equations" in regression analysis.

It follows at once that the conditional distribution of $\underset{\sim}{x}_1$ given $\underset{\sim}{x}_2$ is multivariate normal with mean vector

(6.21) $$\underset{\sim}{\nu}_1 = \underset{\sim}{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-}(\underset{\sim}{x}_2-\underset{\sim}{\mu}_2) ,$$

cf. (6.4), and covariance matrix (6.17). The mean vector (6.21) is unique provided $\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2$ lies in the column space of Σ_{22} (with probability 1), in view of (6.18). This is assured by the distribution of $\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2$ being multivariate normal with mean vector $\underset{\sim}{0}$ and covariance matrix Σ_{22} .

Cottle (1974, p. 195) gives an interesting interpretation of the "quotient property" for the multivariate normal distribution. See also Anderson (1958, p. 33). Let the random vector

(6.22) $$\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \\ \underset{\sim}{x}_3 \end{pmatrix} .$$

Suppose that we have the conditional distribution of $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ given $\underset{\sim}{x}_3$. How do we find the conditional distribution of $\underset{\sim}{x}_1$ given $\underset{\sim}{x}_2$ and $\underset{\sim}{x}_3$? Let us partition the covariance matrix of $\underset{\sim}{x}$ as

$$(6.23) \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix},$$

and write

$$(6.24) \qquad \Sigma_{2\&3} = \begin{pmatrix} \Sigma_{22} & \Sigma_{23} \\ \Sigma_{32} & \Sigma_{33} \end{pmatrix}.$$

Then (3.17) yields

$$(6.25) \qquad (\Sigma/\Sigma_{2\&3}) = ((\Sigma/\Sigma_{33})/(\Sigma_{2\&3}/\Sigma_{33})).$$

Thus the conditional distribution of $\underset{\sim}{x}_1$ given $\underset{\sim}{x}_2$ and $\underset{\sim}{x}_3$ is the conditional distribution of $\begin{pmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \end{pmatrix} \mid \underset{\sim}{x}_3$ given $(\underset{\sim}{x}_2 | \underset{\sim}{x}_3)$. In other words we may condition sequentially.

§6.2 *Partial correlation coefficients.*

In Section 6.1 we saw that (Σ/Σ_{22}), the Schur complement of Σ_{22} in the covariance matrix

$$(6.26) \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

is also a covariance matrix. Anderson (1958, p. 29) defines the elements of (Σ/Σ_{22}) to be *partial covariances*.

Writing

$$(6.27) \qquad (\Sigma/\Sigma_{22}) = \{\sigma_{ij}^{(2)}\} \quad ,$$

we may define the *partial correlation coefficient* as

$$(6.28) \qquad \rho_{ij}^{(2)} = \sigma_{ij}^{(2)} / (\sigma_{ii}^{(2)}\sigma_{jj}^{(2)})^{\frac{1}{2}} \quad ,$$

provided $\sigma_{ii}^{(2)} > 0$ for all i (which is assured when Σ is positive definite).

The diagonal elements of the covariance matrix Σ are the variances of the components of the underlying random vector $\underset{\sim}{x}$. When these variances are all positive we may form the correlation matrix of $\underset{\sim}{x}$ as

$$(6.29) \qquad R = \Delta^{-1}\Sigma\Delta^{-1} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} ,$$

where

$$(6.30) \qquad \Delta = diag(\sigma_{ii}^{\frac{1}{2}}) = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix} \quad ,$$

say, is the diagonal matrix of standard deviations. If the Schur complement

$$(6.31) \qquad (R/R_{22}) = \{r_{ij}^{(2)}\} \, ,$$

then

$$(6.32) \qquad \rho_{ij}^{(2)} = r_{ij}^{(2)}/(r_{ii}^{(2)}r_{jj}^{(2)})^{\frac{1}{2}} \quad ,$$

i.e., the matrix of partial correlation coefficients is also the correlation matrix formed from the Schur complement in the original correlation matrix. To prove (6.32) notice that

$$(6.33) \qquad R_{11} - R_{12}R_{22}^{-}R_{21} = \Delta_1^{-1}(\Sigma_{11} - \Sigma_{12}\Delta_2^{-1}(\Delta_2^{-1}\Sigma_{22}\Delta_2^{-1})^{-}\Delta_2^{-1}\Sigma_{21})\Delta_1^{-1}$$
$$= \Delta_1^{-1}(\Sigma/\Sigma_{22})\Delta_1^{-1} \quad .$$

We may exploit the quotient property, cf. (6.25), to obtain a recursion formula for partial correlation coefficients (cf. Anderson, 1958, p. 34). Partition the random vector

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{6.34}$$

where x_2 is a scalar, and write

$$R = \begin{pmatrix} R_{11} & \rho_{12} & R_{13} \\ \rho_{12}' & 1 & \rho_{23}' \\ R_{13}' & \rho_{23} & R_{33} \end{pmatrix}. \tag{6.35}$$

Then, using (6.25), we obtain

$$\{\rho_{ij}^{(2\&3)}\} = R_{11} - R_{13}R_{33}^{-} R_{13}' - \frac{(\rho_{12} - R_{13}R_{33}^{-}\rho_{23})(\rho_{12} - R_{13}R_{33}^{-}\rho_{23})'}{1 - \rho_{23}'R_{33}^{-}\rho_{23}} \tag{6.36}$$

$$= \left\{ r_{ij}^{(3)} - \frac{r_{i2}^{(3)} r_{j2}^{(3)}}{r_{22}^{(3)}} \right\}. \tag{6.37}$$

Hence

$$\rho_{ij}^{(2\&3)} = \frac{\rho_{ij}^{(3)} - \rho_{i2}^{(3)} \rho_{j2}^{(3)}}{(1 - [\rho_{i2}^{(3)}]^2)^{\frac{1}{2}}(1 - [\rho_{j2}^{(3)}]^2)^{\frac{1}{2}}}, \tag{6.38}$$

cf. (34) in Anderson (1958, p. 34).

Now suppose that $\underset{\sim}{x}_1$ in (6.34) is also a scalar, and partition

$$(6.39) \qquad R = \left(\begin{array}{cc|c} 1 & \rho_{12} & \underset{\sim}{\rho}'_{13} \\ \rho_{12} & 1 & \underset{\sim}{\rho}'_{23} \\ \hline \underset{\sim}{\rho}_{13} & \underset{\sim}{\rho}_{23} & R_3 \end{array}\right).$$

Then the Schur complement

$$(6.40) \qquad (R/R_3) = \begin{pmatrix} 1 - \underset{\sim}{\rho}'_{13} R_3^- \underset{\sim}{\rho}_{13} & \rho_{12} - \underset{\sim}{\rho}'_{13} R_3^- \underset{\sim}{\rho}_{23} \\ \rho_{12} - \underset{\sim}{\rho}'_{13} R_3^- \underset{\sim}{\rho}_{23} & 1 - \underset{\sim}{\rho}'_{23} R_3^- \underset{\sim}{\rho}_{23} \end{pmatrix},$$

and so

$$(6.41) \qquad \rho_{12}^{(3)} = \frac{\rho_{12} - \underset{\sim}{\rho}'_{13} R_3^- \underset{\sim}{\rho}_{23}}{(1 - \underset{\sim}{\rho}'_{13} R_3^- \underset{\sim}{\rho}_{13})^{\frac{1}{2}} (1 - \underset{\sim}{\rho}'_{23} R_3^- \underset{\sim}{\rho}_{23})^{\frac{1}{2}}}.$$

When R is nonsingular, we may obtain an alternate formula for $\rho_{12}^{(3)}$, using R^{-1}. From (2.40) we may write

$$(6.42) \qquad \{\rho^{ij}\} = R^{-1} = \begin{pmatrix} (R/R_3)^{-1} & \cdot \\ \cdot & \cdot \end{pmatrix},$$

and

$$(6.43) \qquad (R/R_3)^{-1} = \begin{pmatrix} 1 - \underset{\sim}{\rho}'_{23} R_3^{-1} \underset{\sim}{\rho}_{23} & -\rho_{12} + \underset{\sim}{\rho}'_{13} R_3^{-1} \underset{\sim}{\rho}_{23} \\ -\rho_{12} + \underset{\sim}{\rho}'_{13} R_3^{-1} \underset{\sim}{\rho}_{23} & 1 - \underset{\sim}{\rho}'_{13} R_3^{-1} \underset{\sim}{\rho}_{13} \end{pmatrix} \Big/ |(R/R_3)|.$$

Hence

$$(6.44) \qquad \rho_{12}^{(3)} = \frac{-\rho^{12}}{(\rho^{11}\rho^{22})^{\frac{1}{2}}},$$

the negative of the corresponding correlation coefficient in R^{-1}, (note that the minus sign has been dropped in (4g. 2.8) in Rao, 1973, p. 270).

§6.3 *Special covariance and correlation structures.*

There are several special covariance and correlation structures that arise in statistical applications. For example, consider the following correlation structure

$$R = (1-\rho)I_n + \rho \underset{\sim}{e}\underset{\sim}{e}' , \tag{6.45}$$

which arises, for example, in the one-way random-effects analysis of variance (cf. e.g., Scheffé, 1959, p.225). Consider the model

$$y_{ij} = \mu + a_i + u_{ij} \ ; \ j = 1,\ldots,n_i, \ \ i = 1,\ldots,k \ , \tag{6.46}$$

with

$$n = \sum_{i=1}^{k} n_i \ . \tag{6.47}$$

We assume that the k+n random variables $a_1,\ldots,a_k,u_{11},u_{12},\ldots,u_{kn_k}$ all have zero mean and are uncorrelated, and that

$$\mathcal{V}(a_i) = \sigma_a^2 , \ \ i = 1,\ldots,k \tag{6.48a}$$

$$\mathcal{V}(u_{ij}) = \sigma^2, \ \ j = 1,\ldots,n_i, \ \ i = 1,\ldots,k \ . \tag{6.48b}$$

Let $\underset{\sim}{y}_i = \{y_{ij}\}_{j=1,\ldots,n_i}$ and $\underset{\sim}{y} = \{\underset{\sim}{y}_i\}_{i=1,\ldots,k}$. Then the covariance matrix of $\underset{\sim}{y}$ is

$$diag(\sigma^2 I_{n_i} + \sigma_a^2 \underset{\sim}{e}^{(n_i)} \underset{\sim}{e}^{(n_i)'})_{i=1,\ldots,k} , \tag{6.49}$$

where $\underset{\sim}{e}^{(n_i)}$ is the n_ix1 vector of ones. The correlation matrix of $\underset{\sim}{y}_i$ is, therefore, of the type (6.45), with

$$\rho = \sigma_a^2/(\sigma^2 + \sigma_a^2) \ ; \tag{6.50}$$

this is called the "intraclass" correlation between y_{ij} and $y_{ij'}$, where $j \neq j'$. If $n_1 = n_2 = \ldots = n_k = m$ then (6.49) becomes

$$I_k \otimes (\sigma^2 I_m + \sigma_a^2 e^{(m)} e^{(m)'}) \; , \tag{6.51}$$

where $\otimes$ is the Kronecker product.

It is of interest to obtain, in closed form, expressions for the determinant, inverse, and characteristic roots of a correlation matrix with structure like (6.45). The determinant and inverse, for example, occur in the density function of the multivariate normal distribution, cf. (6.9).

The determinant of the $n \times n$ matrix R given by (6.45) is

$$|R| = |(1-\rho)I_n + \rho ee'| = (1-\rho)^n |I_n + \rho ee'/(1-\rho)| \tag{6.52a}$$

$$= (1-\rho)^n [1 + \rho n/(1-\rho)] \tag{6.52b}$$

$$= (1-\rho)^{n-1} [1 + \rho(n-1)] \; , \tag{6.52c}$$

using (2.8). Thus R is nonsingular provided $\rho \neq 1$ or $-1/(n-1)$, and then we may compute the inverse R^{-1} using the formula (2.59), i.e.,

$$R^{-1} = [(1-\rho)I_n + \rho ee']^{-1} \tag{6.53a}$$

$$= \frac{1}{1-\rho} I_n - \frac{\rho}{(1-\rho)^2 [1 + \rho n/(1-\rho)]} ee' \tag{6.53b}$$

$$= \frac{1}{1-\rho} \{I_n - \rho ee'/[1 + \rho(n-1)]\} \; . \tag{6.53c}$$

We may find the characteristic roots by solving

$$|R - \lambda I_n| = |(1-\rho-\lambda)I_n + \rho ee'| = 0 \; . \tag{6.54}$$

Using (2.8), we obtain .

$$|R - \lambda I_n| = (1-\rho-\lambda)^{n-1}(1-\rho-\lambda+n\rho) , \tag{6.55}$$

and so the characteristic roots are $1-\rho$ with multiplicity $n-1$ and $1+\rho(n-1)$ with multiplicity 1.

The matrix R defined by (6.45) is positive definite if and only if all the characteristic roots are positive, i.e.,

$$-\frac{1}{n-1} < \rho < 1 . \tag{6.56}$$

As $n\to\infty$ the region of allowable negative values of ρ decreases to 0. For intraclass correlation, however, $\rho > 0$, cf. (6.50).

Another special correlation structure, called the multivariate extension of intraclass correlation by Sampson (1978), is

$$R = \begin{pmatrix} I_m & \rho \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(n)'} \\ \rho \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(m)'} & I_n \end{pmatrix}, \tag{6.57}$$

which arises, e.g., in the two-way balanced fixed effects analysis of variance. Assuming one observation per cell the design matrix may be written as

$$X = \left(\begin{array}{c|c} \underset{\sim}{e}^{(n)} \underset{\sim}{e}_1^{(m)'} & I_n \\ \vdots & \vdots \\ \underset{\sim}{e}^{(n)} \underset{\sim}{e}_m^{(m)'} & I_n \end{array}\right), \tag{6.58}$$

cf. e.g., Scheffé (1959, p. 100), where $\underset{\sim}{e}_i^{(k)}$ is the $k\times 1$ vector with 1 in the ith cell and 0 elsewhere. The matrix X is $mn\times(m+n)$, where m is the number of rows and n the number of columns in the experimental design. Hence

$$(6.59) \qquad X'X = \begin{pmatrix} nI_m & \underset{\sim}{e}^{(m)}\underset{\sim}{e}^{(n)'} \\ \underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(m)'} & mI_n \end{pmatrix} .$$

When the vector $\underset{\sim}{y}$ of observations on the dependent variable has covariance matrix $\sigma^2 I$, then $X'\underset{\sim}{y}$ has covariance matrix $\sigma^2 X'X$. The corresponding correlation matrix has the structure (6.57) with $\rho = (mn)^{-\frac{1}{2}}$, which is the maximum value of ρ so that (6.57) be nonnegative definite, cf. (6.70) below.

The determinant of (6.57) is

$$(6.60a) \qquad \begin{vmatrix} I_m & \rho\underset{\sim}{e}^{(m)}\underset{\sim}{e}^{(n)'} \\ \rho\underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(m)'} & I_n \end{vmatrix} = |I_n - m\rho^2\underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(n)'}|$$

$$(6.60b) \qquad = 1 - mn\rho^2 ,$$

using (2.4) and (2.8). Thus (6.57) is singular $\Leftrightarrow \rho^2 = 1/(mn)$, and so (6.59) is singular. Using (2.25), moreover, we see that

$$(6.61a) \qquad r(X) = r(X'X) = m + r(mI_n - m\underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(n)'}/n)$$

$$(6.61b) \qquad = m + r(I_n - \underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(n)'}/n) .$$

The matrix $C_n = I_n - \underset{\sim\sim}{ee}'/n$ may be called the "centering matrix", cf. Sharpe & Styan (1965). The corresponding correlation matrix is the intraclass correlation matrix (6.45) with $\rho = -1/(n-1)$; this value of ρ is the lowest so that (6.45) remains nonnegative definite, cf. (6.56). Using (4.147), however, we see that the centering matrix has nullity 1 and hence has rank $n-1$. Thus the design matrix (6.58) has rank

$$(6.62) \qquad r(X) = m + n - 1 .$$

To compute the inverse of

$$(6.63) \qquad R = \begin{pmatrix} I_m & \rho \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(n)'} \\ \rho \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(m)'} & I_n \end{pmatrix}, \qquad \rho^2 \neq 1/(mn) ,$$

we use (2.41) and the Schur complements :

$$(6.64a) \qquad S = (R/I_m) = I_n - m\rho^2 \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(n)'} ,$$

$$(6.64b) \qquad T = (R/I_n) = I_m - n\rho^2 \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(m)'} ,$$

and their inverses :

$$(6.65a) \qquad S^{-1} = I_n + m\rho^2 \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(n)'} /(1 - mn\rho^2)$$

$$(6.65b) \qquad T^{-1} = I_m + n\rho^2 \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(m)'} /(1 - mn\rho^2) ,$$

which may be found using (2.59). Hence

$$(6.66) \qquad R^{-1} = \begin{pmatrix} I_m + n\rho^2 \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(m)'} /(1-mn\rho^2) & -\rho \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(n)'} /(1-mn\rho^2) \\ -\rho \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(m)'} /(1-mn\rho^2) & I_n + m\rho^2 \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(n)'} /(1-mn\rho^2) \end{pmatrix} .$$

To compute a generalized inverse of (6.59),

$$(6.67) \qquad X'X = \begin{pmatrix} nI_m & \underset{\sim}{e}^{(m)} \underset{\sim}{e}^{(n)'} \\ \underset{\sim}{e}^{(n)} \underset{\sim}{e}^{(m)'} & mI_n \end{pmatrix},$$

we may use (4.61) since (6.67) is nonnegative definite. The Schur complement

$$(6.68a) \qquad (X'X/nI_m) = m(I_n - \underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(n)'}/n) = mC_n ,$$

where C_n is the centering matrix, cf. (6.61). Since C_n is idempotent it follows that $C_n = C_n^-$ and so $C_n/m = (X'X/nI_m)^-$. Hence

$$(6.68b) \qquad \begin{pmatrix} I_m/n & 0 \\ 0 & C_n/m \end{pmatrix} = (X'X)^- .$$

The characteristic roots of (6.57) may be obtained from

$$(6.69a) \qquad |R - \lambda I_{m+n}| = \begin{vmatrix} (1-\lambda)I_m & \rho\underset{\sim}{e}^{(m)}\underset{\sim}{e}^{(n)'} \\ \rho\underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(m)'} & (1-\lambda)I_n \end{vmatrix}$$

$$(6.69b) \qquad = (1-\lambda)^m|(1-\lambda)I_n - \rho^2 m\underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(n)'}/(1-\lambda)|$$

$$(6.69c) \qquad = (1-\lambda)^{m+n}|I_n - \rho^2 m\underset{\sim}{e}^{(n)}\underset{\sim}{e}^{(n)'}/(1-\lambda)^2|$$

$$(6.69d) \qquad = (1-\lambda)^{m+n-2}[(1-\lambda)^2 - \rho^2 mn] ,$$

using (2.4) and (2.8). Hence, the characteristic roots of (6.57) are 1 with multiplicity $m+n-2$, and $1 \pm \rho\sqrt{mn}$, each with multiplicity 1. Thus (6.57) is positive definite if and only if

$$(6.70) \qquad -(mn)^{-\frac{1}{2}} < \rho < (mn)^{-\frac{1}{2}} .$$

We note that the correlation structure (6.57) is a special case of that considered in Theorem 4.14. In (4.191) set $A = I_m$, $B = I_n$; then $\underset{\sim}{u} = m^{-\frac{1}{2}}\underset{\sim}{e}^{(m)}$ is a normalized characteristic vector of A corresponding to a unit root. Similarly $\underset{\sim}{v} = n^{-\frac{1}{2}}\underset{\sim}{e}^{(n)}$ for B . Hence put $\gamma = \rho(mn)^{\frac{1}{2}}$. Then the characteristic roots of (6.57) are 1 with multiplicity $m+n-2$ and the two roots of

$$\begin{pmatrix} 1 & \rho(mn)^{\frac{1}{2}} \\ \rho(mn)^{\frac{1}{2}} & 1 \end{pmatrix},$$

cf. (4.192).

§6.4 *The chi-squared and Wishart distributions.*

In this section, we will discuss results pertaining to distributions of certain statistics which Rao (1973, p. 189) states as being "fundamental to the theory of least squares".

Consider the general linear model with normality

$$(6.71) \qquad \underset{\sim}{y} \sim N(X\underset{\sim}{\gamma}, \sigma^2 I_n) ,$$

where X has rank r . The residual sum of squares

$$(6.72) \qquad S_e = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{y}'X(X'X)^- X'\underset{\sim}{y} = \underset{\sim}{y}'[I - X(X'X)^- X']\underset{\sim}{y}$$

is the Schur complement of $X'X$ in the matrix

$$(6.73) \qquad \begin{pmatrix} X'X & X'\underset{\sim}{y} \\ \underset{\sim}{y}'X & \underset{\sim}{y}'\underset{\sim}{y} \end{pmatrix} = \begin{pmatrix} X' \\ \underset{\sim}{y}' \end{pmatrix} (X, \underset{\sim}{y}) .$$

Hence,

$$(6.74) \qquad S_e \sim \sigma^2 \chi^2_{n-r} ,$$

central chi-squared with $n-r$ degrees of freedom, cf. Rao (1973, p. 189).

Now consider the multivariate general linear model with normality,

$$(6.75) \qquad \underset{\sim}{Y} = X\Gamma + U ,$$

where Y and U are $n \times p$ with rows following independent p-variate normal distributions with covariance matrix Σ. The residual matrix of sums of squares and cross-products

(6.76) $$S_e = Y'Y - Y'X(X'X)^{-}X'Y = Y'[I - X(X'X)^{-}X']Y$$

is the Schur complement of $X'X$ in the matrix

(6.77) $$\begin{pmatrix} X'X & X'Y \\ Y'X & Y'Y \end{pmatrix} = \begin{pmatrix} X' \\ Y' \end{pmatrix} (X, Y)$$

and

(6.78) $$S_e \sim \mathcal{W}_p(n-r, \Sigma) ,$$

the p-variate central Wishart distribution with $n-r$ degrees of freedom and scale parameter Σ, cf. Rao (1973, p. 534). When $p = 1$, then $\Sigma = \sigma^2$ and (6.78) reduces to (6.74).

To prove (6.78) we may use the following result, cf. Rao (1973, p. 536). Let the random $n \times p$ matrix Z have independent rows each normally distributed with covariance matrix Σ. Suppose $\mathcal{E}(Z) = \Omega$. If A is a nonrandom symmetric $n \times n$ matrix then

(6.79) $$W = Z'AZ \sim \mathcal{W}_p(f, \Sigma)$$

if and only if $A = A^2$ and $A\Omega = 0$, and then $f = r(A)$. Clearly $I - X(X'X)^{-}X' = M = M^2$ and $r(M) = n - r(X) = n-r$. Since $\Omega = X\Gamma$, then $A\Omega = MX\Gamma = 0$.

A somewhat different result concerning the Wishart distribution of a Schur complement may be obtained from (6.79) by setting $A = I$ and partitioning

(6.80) $$Z = (Z_1, Z_2) ,$$

where Z_1 is $n \times p_1$ and Z_2 is $n \times p_2$, and $n > p$. Then

(6.81) $$W = \begin{pmatrix} Z_1'Z_1 & Z_1'Z_2 \\ Z_2'Z_1 & Z_2'Z_2 \end{pmatrix} .$$

Partition Σ similarly so that

(6.82) $$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

and suppose $\Omega = E(Z) = 0$. Then

(6.83) $$(W/Z_2'Z_2) \sim \mathcal{W}_{p_1}(n-r_2,(\Sigma/\Sigma_{22})) ,$$

where $r_2 = r(\Sigma_{22})$.

To prove (6.83) we consider

(6.84) $$(W/Z_2'Z_2) = Z_1'[I - Z_2(Z_2'Z_2)^- Z_2']Z_1 .$$

Moreover given Z_2 the rows of Z_1 are independently normally distributed with covariance matrix (Σ/Σ_{22}), while

(6.85) $$E(Z_1|Z_2) = Z_2\Sigma_{22}^-\Sigma_{21} ,$$

cf. (6.17) and (6.21). Then (6.83) follows at once since Z_2 has rank $r(\Sigma_{22})$ with probability 1. Rao (1973, p. 539) proves (6.83) when Σ_{22} is positive definite, while Ruohonen (1976) establishes (6.83) using Moore-Penrose g-inverses.

Mitra (1970) derives a result analogous to (6.83) for the matrix-variate beta distribution. Let W_1 and W_2 be independent $p \times p$ random matrices such that

(6.86) $$W_i \sim \mathcal{W}_p(k_i,\Sigma) ; \quad i = 1,2 ,$$

and $k_1 + k_2 \geq p$. Then $W = W_1 + W_2$ is positive definite with probability 1 and we may define

$$B = W^{-\frac{1}{2}} W_1 W^{-\frac{1}{2}} \sim B_p(k_1, k_2) , \tag{6.87}$$

the p-variate beta distribution with k_1 , k_2 degrees of freedom. If we partition

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} , \tag{6.88}$$

where B_{11} is $p_1 \times p_1$ and B_{22} is $p_2 \times p_2$, then the Schur complement

$$(B/B_{22}) \sim B_{p_1}(k_1, k_2 - r_2) , \tag{6.89}$$

the p_1-variate beta distribution with k_1 , $k_2 - r_2$ degrees of freedom.

§6.5 *The Cramér-Rao inequality.*

Let $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n$ be independently and identically distributed as the random vector $\underset{\sim}{x}$, whose distribution depends on the unknown parameter vector $\underset{\sim}{\theta}$. Then the score vector is defined as

$$\underset{\sim}{s} = \partial \log L \,/\, \partial \underset{\sim}{\theta} , \tag{6.90}$$

where L denotes the likelihood function of $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n$. Let $\underset{\sim}{t}$ be an unbiased estimator for $\underset{\sim}{\theta}$, i.e.,

$$E(\underset{\sim}{t}) = \underset{\sim}{\theta} . \tag{6.91}$$

Then the random vector

$$\underset{\sim}{u} = \begin{pmatrix} \underset{\sim}{s} \\ \\ \underset{\sim}{t} \end{pmatrix} \tag{6.92}$$

has, under certain regularity conditions, mean vector

$$\underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{0} \\ \\ \underset{\sim}{\theta} \end{pmatrix} \tag{6.93}$$

and covariance matrix structure

$$\Sigma = \begin{pmatrix} \Sigma_{11} & I \\ I & \Sigma_{22} \end{pmatrix} . \tag{6.94}$$

If Σ_{11} is positive definite then it follows from Theorem 3.2 that the Schur complement

$$(\Sigma/\Sigma_{11}) = \Sigma_{22} - \Sigma_{11}^{-1} \tag{6.95}$$

is nonnegative definite. If, then, an unbiased statistic $\underset{\sim}{t}_0$, say, can be found with covariance matrix $\Sigma_{11}^{-1} = [V(\partial \log L/\ \partial \underset{\sim}{\theta})]^{-1}$ then $\underset{\sim}{t}_0$ is the *minimum variance unbiased* or *Markov* estimator of $\underset{\sim}{\theta}$. This result is usually called the Cramér-Rao inequality, though Sverdrup (1967, p. 72) and Savage (1972, p. 238) claim that it is due to Fréchet (1943).

To prove $E(\underset{\sim}{s}) = \underset{\sim}{0}$ and $cov(\underset{\sim}{s},\underset{\sim}{t}) = I$ we note first that

$$\int L \, d\underset{\sim}{x}_1 \ldots d\underset{\sim}{x}_n = 1 , \tag{6.96}$$

which implies that, under appropriate regularity conditions,

$$E(\underset{\sim}{s}) = \int \frac{\partial \log L}{\partial \underset{\sim}{\theta}} \cdot L \cdot d\underset{\sim}{z} = \int \frac{\partial L}{\partial \underset{\sim}{\theta}} \cdot d\underset{\sim}{z} \tag{6.97a}$$

$$= \frac{\partial}{\partial \underset{\sim}{\theta}} \int L \cdot d\underset{\sim}{z} = \underset{\sim}{0} \,, \tag{6.97b}$$

where $d\underset{\sim}{z} = d\underset{\sim}{x}_1 \, d\underset{\sim}{x}_2 \ldots d\underset{\sim}{x}_n$. Moreover,

$$cov(\underset{\sim}{s},\underset{\sim}{t}) = E(\underset{\sim}{s}\underset{\sim}{t}') = \{E(s_i t_j)\} \tag{6.98a}$$

$$= \left\{ \int \frac{\partial \log L}{\partial \theta_i} \cdot t_j \cdot L \cdot d\underset{\sim}{z} \right\} \tag{6.98b}$$

$$= \left\{ \int \frac{\partial L}{\partial \theta_i} \cdot t_j \cdot d\underset{\sim}{z} \right\} \tag{6.98c}$$

$$= \left\{ \frac{\partial}{\partial \theta_i} \int t_j \, L \, d\underset{\sim}{z} \right\} \tag{6.98d}$$

$$= \left\{ \frac{\partial}{\partial \theta_i} E(t_j) \right\} = \left\{ \frac{\partial \theta_j}{\partial \theta_i} \right\} = \{\delta_{ij}\} = I \,. \tag{6.98e}$$

In particular, if

$$\underset{\sim}{y} \sim N(X\underset{\sim}{\gamma}, \sigma^2 I) \,, \tag{6.99}$$

where X has full column rank, then

$$L = (2\pi\sigma^2)^{-\frac{1}{2}n} \exp\{-\tfrac{1}{2}(\underset{\sim}{y} - X\underset{\sim}{\gamma})'(\underset{\sim}{y} - X\underset{\sim}{\gamma})/\sigma^2\} \,, \tag{6.100}$$

$$\log L = -\tfrac{1}{2}n \log 2\pi - n \log \sigma - \tfrac{1}{2}(\underset{\sim}{y} - X\underset{\sim}{\gamma})'(\underset{\sim}{y} - X\underset{\sim}{\gamma})/\sigma^2 \,, \tag{6.101}$$

$$\underset{\sim}{s} = \partial \log L/\partial \underset{\sim}{\gamma} = X'(\underset{\sim}{y} - X\underset{\sim}{\gamma}) = -(X'X\underset{\sim}{\gamma} - X'\underset{\sim}{y}) \,, \tag{6.102}$$

$$V(\underset{\sim}{s}) = X'X/\sigma^2 \,, \tag{6.103}$$

$$\Sigma_{11}^{-1} = \sigma^2(X'X)^{-1} \quad . \tag{6.104}$$

The maximum likelihood estimator of $\underset{\sim}{\gamma}$ is

$$\hat{\underset{\sim}{\gamma}} = (X'X)^{-1}X'\underset{\sim}{y} \tag{6.105}$$

and this has covariance matrix (6.104). Hence, $\hat{\underset{\sim}{\gamma}}$ is the minimum variance unbiased or Markov estimator of $\underset{\sim}{\gamma}$.

REFERENCES AND BIBLIOGRAPHY

Ahsanullah, M., & Rahman, M. (1973). A singular inverse of a matrix by rank annihilation. *Canadian Mathematical Bulletin*, 16, 1-4.

Aitken, A.C. (1939). *Determinants and Matrices*. Oliver & Boyd, Edinburgh.

Anderson, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

Ando, T. (1978). Commutativity formulas for generalized Schur complements. Submitted for publication.

Banachiewicz, T. (1937). Zur Berechnung der Determinanten, wie auch der Inversen, und zur darauf basierten Auflösung der Systeme linearer Gleichungen. *Acta Astronomica, Série C*, 3, 41-67.

Bartlett, M.S.(1951). An inverse matrix adjustment arising in discriminant analysis. *The Annals of Mathematical Statistics*, 22, 107-111.

Bhimasankaram, P. (1971). On generalized inverses of partitioned matrices. *Sankhyā, Series A*, 33, 311-314.

Bjerhammar, A. (1958). A generalized matrix algebra. *Kungl. Tekniska Högskolans Handlingar (Stockholm)*, 124, 1-32.

Bjerhammar, A. (1973). *Theory of Errors and Generalized Matrix Inverses*. Elsevier Scientific Publishing Company, Amsterdam.

Bodewig, E. (1947). Comparison of some direct methods for computing determinants and inverse matrices. *Koninklijke Nederlandse Akademie van Wetenschappen. Proceedings of the Section of Sciences*, 50, 49-57.

Bodewig, E. (1956). *Matrix Calculus*. North-Holland, Amsterdam. [Second Revised and Enlarged Edition, 1959.]

Boerner, H. (1975). Issai Schur. In *Dictionary of Scientific Biography* (C.C. Gillispie, ed.), Volume XII, Scribner's, New York, p. 237.

Burns, F., Carlson, D., Haynsworth, E., & Markham, T. (1974). Generalized inverse formulas using the Schur complement. *SIAM Journal on Applied Mathematics*, 26, 254-259.

Carlson, D. (1975). Matrix decompositions involving the Schur complement. *SIAM Journal on Applied Mathematics*, 28, 577-587.

Carlson, D., Haynsworth, E., & Markham, T. (1974). A generalization of the Schur complement by means of the Moore-Penrose inverse. *SIAM Journal on Applied Mathematics*, 26, 169-175.

Cline, R.E., & Funderlic, R.E. (1976). The rank of a difference of matrices and associated generalized inverses. Technical Report CS-76-20, Department of Computer Science, Univ. Tennessee.

Cottle, R.W. (1974). Manifestations of the Schur complement. *Linear Algebra and Its Applications*, 8, 189-211.

Crabtree, D.E., & Haynsworth, E.V. (1969). An identity for the Schur complement of a matrix. *Proceedings of the American Mathematical Society*, 22, 364-366.

Dias da Silva, J. (1976). Review of Fiedler (1974). *Mathematical Reviews*, 51, #543.

Duncan, W.J. (1944). Some devices for the solution of large sets of simultaneous linear equations. (With an appendix on the reciprocation of partitioned matrices.) *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Seventh Series*, 35, 660-670.

Edelblute, D.J. (1966). Matrix inversion by rank annihilation. *Mathematics of Computation*, 20, 149-151.

Ellenberg, J.H. (1973). The joint distribution of the standardized least squares residuals from a general linear regression. *Journal of the American Statistical Association*, 68, 941-943.

Faddeev, D.K., & Faddeeva, V.N. (1963). *Computational Methods of Linear Algebra*. Freeman, San Francisco.

Faddeeva, V.N. (1959). *Computational Methods of Linear Algebra*. Dover, New York.

Fiedler, M. (1974). Eigenvalues of nonnegative symmetric matrices. *Linear Algebra and Its Applications*, 9, 119-142.

Frazer, R.A., Duncan, W.J., & Collar, A.R. (1938). *Elementary Matrices, and Some Applications to Dynamics and Differential Equations*. Cambridge University Press.

Frobenius, F.G. (1908). Über Matrizen aus positiven Elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, 471-476. [Reprinted in *Ferdinand Georg Frobenius Gesammelte Abhandlungen*, Springer-Verlag, Berlin, III(1968), 404-409.]

Goddard, L.S., & Schneider, H. (1955). Pairs of matrices with a non-zero commutator. *Proceedings of the Cambridge Philosophical Society*, 51, 551-553.

Guttman, Irwin (1971). A Bayesian analogue of Paulson's lemma and its use in tolerance region construction when sampling from the multi-variate normal. *Annals of the Institute of Statistical Mathematics*, 23, 67-76.

Guttman, Louis (1946). Enlargement methods for computing the inverse matrix. *The Annals of Mathematical Statistics*, 17, 336-343.

Hartfiel, D.J. (1973). An extension of Haynsworth's determinant inequality. *Proceedings of the American Mathematical Society*, 41, 463-465.

Hartwig, R.E. (1976a). Singular value decomposition and the Moore-Penrose inverse of bordered matrices. *SIAM Journal on Applied Mathematics*, 31, 31-41.

Hartwig, R.E. (1976b). Rank factorization and Moore-Penrose inversion. *The Journal of the Industrial Mathematics Society*, 26, 49-63.

Haynsworth, E.V. (1968). Determination of the inertia of a partitioned Hermitian matrix. *Linear Algebra and Its Applications*, 1, 73-81.

Haynsworth, E.V. (1970a). Reduction of a matrix using properties of the Schur complement. *Linear Algebra and Its Applications*, 3, 23-29. [See also: On the Schur complement. *Basle Mathematical Notes*, BMN 20 (1968), 17 pp.]

Haynsworth, E.V. (1970b). Applications of an inequality for the Schur complement. *Proceedings of the American Mathematical Society*, 24, 512-516.

Herstein, I.N., & Small, L.W. (1975). An extension of a theorem of Schur. *Linear and Multilinear Algebra*, 3, 41-43.

Hotelling, H. (1943). Some new methods in matrix calculation. *The Annals of Mathematical Statistics*, 14, 1-34.

Householder, A.S. (1964). *The Theory of Matrices in Numerical Analysis*. Blaisdell, New York.

Jacobi, C.G.J. (1834). De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis, quae solis quadratis variabilium constant; una cum variis theorematis de transformatione et determinatione integralium multiplicium. *Journal für die reine und angewandte Mathematik*, 12, 1-69.

Jossa, F. (1940). Risoluzione progressiva di un sistema di equazioni lineari. Analogia con un problema meccanico. *Rendiconto dell'Accademia delle Scienze Fisiche e Matematiche, Società Reale di Napoli, Serie 4*, 10, 346-352.

Khatri, C.G. (1969). A note on some results on a generalised inverse of a matrix. *Journal of the Indian Statistical Association*, 7, 38-45.

Lanczos, C. (1958). Linear systems in self-adjoint form. *The American Mathematical Monthly*, 65, 665-679.

Marcus, M., & Minc, H. (1964). *A Survey of Matrix Theory and Matrix Inequalities*. Prindle, Weber & Schmidt, Boston.

Marsaglia, G., & Styan, G.P.H. (1974a). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra*, 2, 269-292.

Marsaglia, G., & Styan, G.P.H. (1974b). Rank conditions for generalized inverses of partitioned matrices. *Sankhyā, Series A*, 36, 437-442.

Meyer, C.D., Jr. (1970). Generalized inverses of block triangular matrices. *SIAM Journal on Applied Mathematics*, 19, 741-750.

Meyer, C.D., Jr. (1973). Generalized inverses and ranks of block matrices. *SIAM Journal on Applied Mathematics*, 25, 597-602.

Milliken, G.A., & Akdeniz, F. (1977). A theorem on the difference of the generalized inverses of two nonnegative matrices. *Communications in Statistics - Theory and Methods*, A6, 73-79.

Mirsky, L. (1955). *An Introduction to Linear Algebra*. Oxford University Press.

Mitra, S.K. (1970). A density-free approach to the matrix variate beta distribution. *Sankhyā, Series A*, 32, 81-88.

Newman, M. (1962). Matrix computations. In *Survey of Numerical Analysis* (J. Todd, ed.), McGraw-Hill, New York, pp. 222-254.

Ostrowski, A. (1971). A new proof of Haynsworth's quotient formula for Schur complements. *Linear Algebra and Its Applications*, 4, 389-392.

Pringle, R.M., & Rayner, A.A. (1970). Expressions for generalized inverses of a bordered matrix with application to the theory of constrained linear models. *SIAM Review*, 12, 107-115.

Rao, C.R. (1962). A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *Journal of the Royal Statistical Society, Series B*, 24, 152-158.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*. Second Edition. Wiley, New York.

Rao, C.R., & Mitra, S.K. (1971). *Generalized Inverse of Matrices and Its Applications.* Wiley, New York.

Rohde, C.A. (1965). Generalized inverses of partitioned matrices. *Journal of the Society for Industrial and Applied Mathematics*, 13, 1033-1035.

Ruohonen, M. (1973). Generalization of a matrix inversion lemma. Technical Report No. 44, Institute for Applied Mathematics, University of Turku, Turku, Finland.

Ruohonen, M. (1976). Proof of Wishartness of a residual matrix. Technical Report No. 69, Institute for Applied Mathematics, University of Turku, Turku, Finland.

Sampson, A.R. (1978). Simple BAN estimators of correlations for certain multivariate normal models with known variances. *Journal of the American Statistical Association*, to appear.

Savage, L.J. (1972). *The Foundations of Statistics.* Dover, New York.

Scheffé, H. (1959). *The Analysis of Variance.* Wiley, New York.

Schur, J. [Issai] (1917). Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind. *Journal für die reine und angewandte Mathematik,* 147, 205-232.

Sharpe, G.E., & Styan, G.P.H. (1965). Circuit duality and the general network inverse. *IEEE Transactions on Circuit Theory*, CT-12, 22-27.

Sherman, J., & Morrison, W.J. (1949). Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix. (Abstract.) *The Annals of Mathematical Statistics*, 20, 621.

Sherman, J., & Morrison, W.J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21, 124-127.

Styan, G.P.H., & Pukelsheim, F. (1978). On the nonnegative definiteness of the difference between two matrices. Submitted for publication.

Sverdrup, E. (1967). *Basic Concepts of Statistical Inference. Volume II.* North-Holland, Amsterdam.

Waugh, F.V. (1945). A note concerning Hotelling's method of inverting a partitioned matrix. *The Annals of Mathematical Statistics*, 16, 216-217.

Wedderburn, J.H.M. (1934). *Lectures on Matrices.* American Mathematical Society Colloquium Publications, Volume XVII. [Republished 1964, Dover, New York.]

Westlake, J.R. (1968). *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations.* Wiley, New York.

Wilf, H.S. (1959). Matrix inversion by the annihilation of rank. *Journal of the Society for Industrial and Applied Mathematics,* 7, 149-151.

Woodbury, M.A. (1950). Inverting modified matrices. Memorandum Report 42, Statistical Research Group, Princeton, New Jersey.

Zlobec, S.(1975). The Gauss-bordering method for generalized inversion of matrices. *Zeitschrift für angewandte Mathematik und Mechanik,* 55, 445-449.

Zlobec, S., & Chan, M.W.K. (1974). The Gauss-bordering method for matrix inversion. *Zeitschrift für angewandte Mathematik und Mechanik,* 54, 277-279.

APPENDIX A

CRACOVIANS

Following Banachiewicz (1937, p. 45) we define the *Cracovian product* of an $m \times n$ matrix A and an $m \times p$ matrix B as

$$P = A \circ B = \{p_{rs}\} , \tag{A.1}$$

where p_{rs} is the inner product of the rth column of B with the sth column of A. Hence $p_{rs} = \underset{\sim}{e}_s' A'B \underset{\sim}{e}_r = \underset{\sim}{e}_r' B'A \underset{\sim}{e}_s$, so that

$$A \circ B = B'A \tag{A.2}$$

is a $p \times n$ matrix. It follows at once that

$$I_m \circ B = B' \quad \text{and} \quad A \circ I_m = A . \tag{A.3}$$

Banachiewicz calls the identity matrix I *"Idem"*, remarking that his earlier usage of *"Invers"* "ziehen wir ausdrücklich zurück".

It is found convenient to drop the symbol $\circ$ in (A.3) so that

$$I \circ A = IA = A' , \tag{A.4}$$

since the middle form in (A.4) "nicht vorhanden ist" in ordinary matrix algebra. Thus

$$I \circ \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} IE & IG \\ IF & IH \end{pmatrix}, \tag{A.5}$$

cf. (2.7), *op. cit.*, p. 47. Transposition of a Cracovian product reverses the order, for

$$I(A \circ B) = (A \circ B)' = (B'A)' = A'B = B \circ A . \tag{A.6}$$

When A is nonsingular the Cracovian inverse is the transpose of the usual inverse. To see this, write

$$\text{(A.7)} \qquad A^{-1}A = I = A \circ (A^{-1})' = A \circ IA^{-1} ,$$

using (A.2). The Cracovian inverse of the Cracovian product of two nonsingular matrices is the Cracovian product of their Cracovian inverses in the *same* order, for

$$\begin{aligned} \text{(A.8)} \qquad (A \circ B) \circ (IA^{-1} \circ IB^{-1}) &= (B'A) \circ \{(A^{-1})' \circ (B^{-1})'\} \\ &= (B'A) \circ \{B^{-1} (A^{-1})'\} \\ &= A^{-1}(B^{-1})' B'A = A^{-1}A = I . \end{aligned}$$