Schur Complements and Statistics*

Diane Valérie Ouellette†
Department of Mathematics
McGill University
Montreal, Quebec, Canada

Submitted by Hans Schneider

ABSTRACT

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

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F},$$

which we call the Schur complement of $\mathbf{E}$ in

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

The matrix $\mathbf{E}$ is assumed to be nonsingular. When $\mathbf{E}$ is singular or rectangular we consider the generalized Schur complement $\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-}\mathbf{F}$, where $\mathbf{E}^{-}$ is a generalized inverse of $\mathbf{E}$. 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 for generalized Schur complements. We survey the known results in a historical perspective and obtain several extensions. Numerous applications in numerical ananlysis and statistics are included. The paper ends with an exhaustive bibliography of books and articles related to Schur complements.

I. INTRODUCTION AND NOTATION

1.1. *Introduction*

In recent years, the designation "Schur complement" has been applied to any matrix of the form $\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}$. These objects have undoubtedly been encountered from the time matrices were first used. But today under this new name and with new

*Revised version of a thesis submitted toward the M.Sc. degree, Department of Mathematics, McGill University.

†Now with Postes Canada, Région du Québec.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/81/020187+109$02.50

emphasis on their properties, there is greater awareness of the widespread appearance and utility of Schur complements.

R. W. Cottle [18]

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

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F} \tag{1.1}$$

and the generalized Schur complement

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-}\mathbf{F}, \tag{1.2}$$

where $\mathbf{E}^{-}$ is a generalized inverse of $\mathbf{E}$ satisfying $\mathbf{E}\mathbf{E}^{-}\mathbf{E}=\mathbf{E}$. 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 [66] appears to have been the first author to explicitly consider a matrix of the form (1.1). He used (1.1) to prove that

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

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

Emilie V. Haynsworth [36, p. 74] appears to have been the first author to give the name Schur complement to a matrix of the form (1.1). Following her, we refer to

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F} \tag{1.4}$$

as the *Schur complement of* $\mathbf{E}$ *in* $\mathbf{A}$, where the partitioned matrix

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

The notation

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

is convenient.

We may consider the Schur complement of any nonsingular submatrix in $\mathbf{A}$. 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 $\mathbf{A}$. This is equivalent to premultiplication and/or postmultiplication of $\mathbf{A}$ by a permutation matrix.

In the book by Bodewig [11], the formula (1.3) is said [11, 1st ed., p. 189; 2nd ed., p. 218] to date from Frobenius (1849–1917), who obtained [29]

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

where adj denotes adjugate matrix. In (1.7) $\mathbf{f}$ and $\mathbf{g}$ are column vectors, while h is a scalar. Boerner [12] records that Schur (1875–1941) was a student of Frobenius's. We present (1.7) in Theorem 2.3.

Banachiewicz [5] appears to be the first author to express the inverse of a partitioned matrix in terms of the Schur complement. When the partitioned matrix $\mathbf{A}$ in (1.5) and the submatrix $\mathbf{E}$ are both nonsingular, then the Schur complement of $\mathbf{E}$ in $\mathbf{A}$,

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

is also nonsingular [cf. (1.14) below] and

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-1}\mathbf{G}\mathbf{E}^{-1} & -\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\mathbf{G}\mathbf{E}^{-1} & \mathbf{S}^{-1} \end{pmatrix}; \tag{1.9}$$

cf. Theorem 2.7.

Banachiewicz [5] obtained (1.9) in Cracovian notation, where matrices are multiplied column by column (see Appendix for further details).

The formula (1.9) is often attributed to Schur [66] (see e.g., Marsaglia and Styan [48, 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 $\mathbf{A}$ in (1.5) and the submatrix $\mathbf{H}$ are both nonsingular, then it follows similarly that the Schur complement of $\mathbf{H}$ in $\mathbf{A}$

$$\mathbf{T} = (\mathbf{A}/\mathbf{H}) = \mathbf{E} - \mathbf{F}\mathbf{H}^{-1}\mathbf{G} \tag{1.10}$$

is also nonsingular and

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{T}^{-1} & -\mathbf{T}^{-1}\mathbf{F}\mathbf{H}^{-1} \\ -\mathbf{H}^{-1}\mathbf{G}\mathbf{T}^{-1} & \mathbf{H}^{-1} + \mathbf{H}^{-1}\mathbf{G}\mathbf{T}^{-1}\mathbf{F}\mathbf{H}^{-1} \end{pmatrix}. \tag{1.11}$$

When **A**, **E**, and **H** are all three nonsingular, then

$$(\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G})^{-1}=\mathbf{E}^{-1}+\mathbf{E}^{-1}\mathbf{F}(\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F})^{-1}\mathbf{G}\mathbf{E}^{-1}, \tag{1.12}$$

which was observed by Duncan [21] and reestablished by Woodbury [76]. Equation (1.12) led to formulas like [68, 69, 6]

$$(\mathbf{E}+\mathbf{f}\mathbf{g}')^{-1}=\mathbf{E}^{-1}-\frac{\mathbf{E}^{-1}\mathbf{f}\mathbf{g}'\mathbf{E}^{-1}}{1+\mathbf{g}'\mathbf{E}^{-1}\mathbf{f}}; \tag{1.13}$$

cf. Corollary 2.6.

Bodewig [10] has shown, by establishing a count of the number of operations required, that the usual method of calculating the determinant of the partitioned matrix (1.5) is preferable to Schur's formula (1.3). He claims, however, that the opposite is true when the inverse is calculated; cf. (1.9).

Louis Guttman [32] established that if the matrix **E** in (1.5) is nonsingular, then

$$\begin{aligned} \mathrm{r}(\mathbf{A})=\mathrm{r}\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} &= \mathrm{r}(\mathbf{E})+\mathrm{r}(\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F}) \\ &= \mathrm{r}(\mathbf{E})+\mathrm{r}(\mathbf{A}/\mathbf{E}), \end{aligned} \tag{1.14}$$

where $\mathrm{r}(\cdot)$ denotes rank. We present this as Theorem 2.5. In other words, rank is additive on the Schur complement [47, p. 291]. Wedderburn [73] and Householder [41] gave related results on rank, which turn out to be special cases of (1.14). See Theorems 2.6a and 2.6b.

We conclude Sec. II by showing how Schur complements may be used to prove theorems of Cauchy [16] and Jacobi [42].

In Sec. III, we discuss various properties of the Schur complement of a nonsingular matrix which have appeared more recently. It seems (see the survey paper on Schur complements by Cottle [18]) 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 [36] showed that when the partitioned matrix **A** in (1.5) is Hermitian and **E** is nonsingular, then

$$\mathrm{In}\,\mathbf{A}=\mathrm{In}\,\mathbf{E}+\mathrm{In}(\mathbf{A}/\mathbf{E}), \tag{1.15}$$

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 and Haynsworth [19] and Ostrowski [55] prove that if we partition $\mathbf{E}$ as well, i.e.,

$$\mathbf{A}=\left(\begin{array}{c|c}\mathbf{E} & \mathbf{F}\\ \hline \mathbf{G} & \mathbf{H}\end{array}\right)=\left[\begin{array}{cc|c}\mathbf{K} & \mathbf{L} & \mathbf{F}_1\\ \mathbf{M} & \mathbf{N} & \mathbf{F}_2\\ \hline \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{H}\end{array}\right], \tag{1.16}$$

with $\mathbf{E}$ and $\mathbf{K}$ both nonsingular, then

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

(cf. Theorem 3.3). This result, called the *quotient property*, has led to several determinant inequalities [38, 33].

We conclude Sec. 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 [18].

In Sec. IV, we extend the results in Secs. II and III to generalized Schur complements; cf. (1.2). Let the partitioned matrix $\mathbf{A}$ in (1.5) and the submatrix $\mathbf{E}$ both be square. If either

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

or

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

then

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

for every g-inverse $\mathbf{E}^{-}$; cf. Theorem 4.1.

Following Meyer [50], Marsaglia and Styan [47], Carlson, Haynsworth, and Markham [15], and Carlson [14], we establish several results on rank. Among these, we show (Corollary 4.3) that rank is additive on the Schur complement

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

when (1.18) and (1.19) hold.

Following Rohde [60], Pringle and Rayner [56], Bhimasankaram [7], Marsaglia and Styan [48], and Burns, Carlson, Haynsworth, and Markham [13], we investigate conditions under which the Schur-Banachiewicz inversion formula (1.9) works with generalized inverses replacing regular inverses; see Theorem 4.6.

Following Carlson, Haynsworth, and Markham [15], we find that inertia continues to be additive on the (generalized) Schur complement, that is,

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

where the partitioned matrix is real and symmetric, if

$$\mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{E}, \mathbf{F}) \quad (1.23)$$

[cf. (1.18)], where $\mathbf{E}^{-}$ is any g-inverse of $\mathbf{E}$ (see proof of Theorem 4.7).

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

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

and

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

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

$$|\mathbf{I} - \mathbf{FG}| = |\mathbf{I} - \mathbf{GF}| \quad (1.26a)$$

and

$$\psi(\mathbf{I} - \mathbf{FG}) = \psi(\mathbf{I} - \mathbf{GF}) \quad (1.26b)$$

may be established using Schur complements, where $\psi(\cdot)$ denotes nullity.

Section 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, and Collar [27], the variant given by Jossa [43], and the "second-order enlargement" method due to Louis Guttman [32] are described and are accompanied by numerical examples. A similar method, called "geometric enlargement," due to Louis Guttman [32], is also given.

One of the most useful algorithms, perhaps, is that of partitioned Schur complements, outlined by Louis Guttman [32] and later developed by Zlobec and Chan [78]. Wilf [75] elaborated a method of rank annihilation, and Edelblute [22] considered a special case of the above algorithm which simplifies the calculations performed.

Following Newman [54] and Westlake [74, p. 31], we show in Sec. 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 [77], which computes a g-inverse of a partitioned matrix using partitioned Schur complements. Generalized inversion has also been studied by Ahsanullah and Rahman [1], who extended the method of rank annihilation.

Further details of some of these algorithms are given in the books by Faddeeva [25, pp. 105–111] and Faddeev and Faddeeva [24, pp. 161–167, 173–178].

In Sec. 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 Sec. 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 Sec. 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 Sec. 6.4 we show how a quadratic form which follows a χ^2 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 [53]). We conclude Sec. VI and this paper 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 [4] as the matrix

$$\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F} \end{pmatrix}; \tag{1.27}$$

he refers to

$$\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{G}\mathbf{E}^{-1}\mathbf{F} \end{pmatrix} \tag{1.28}$$

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 boldface capital letters, column vectors by boldface lower case letters, and scalars by lightface lowercase letters. An $n\times n$ matrix $\mathbf{A}$ 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 $\operatorname{diag}(a_{11},a_{22},\ldots,a_{nn})$. In particular, $\mathbf{I}=\{\delta_{ij}\}$ represents the identity matrix with δ_{ij} the Kronecker delta, $\mathbf{e}$ or $\mathbf{e}^{(n)}$ the $n\times 1$ column vector of ones, $\mathbf{e}_i$ or $\mathbf{e}_i^{(n)}$ the $n\times 1$ column vector with all elements zero except for unity in the ith position. The transpose of a matrix $\mathbf{A}$ is denoted $\mathbf{A}'$, with $\mathbf{a}'$ the row vector corresponding to the column vector $\mathbf{a}$. The determinant is denoted by $|\cdot|$, the adjugate (or adjoint) matrix by adj, and the trace by tr. Rank is denoted by $\mathrm{r}(\cdot)$ and nullity by $\psi(\cdot)$. We call $\mathbf{A}^-$ a *generalized inverse* (or g-inverse) of $\mathbf{A}$ if $\mathbf{AA}^-\mathbf{A}=\mathbf{A}$ [57, 59]. If, in addition, $\mathbf{A}^-\mathbf{AA}^-=\mathbf{A}^-$, or $\mathrm{r}(\mathbf{A})=\mathrm{r}(\mathbf{A}^-)$, then $\mathbf{A}^-=\mathbf{A}_r^-$, a reflexive g-inverse. If, in addition, the projectors $\mathbf{AA}_r^-$ and $\mathbf{A}_r^-\mathbf{A}$ are both symmetric, then $\mathbf{A}_r^-=\mathbf{A}^+$, the unique *Moore-Penrose* g-inverse of $\mathbf{A}$.

We denote the characteristic roots of $\mathbf{A}$ by $\mathrm{ch}(\mathbf{A})$, with $\mathrm{ch}_j(\mathbf{A})$ being the jth largest when the roots are real. The inertia $\mathrm{In}\,\mathbf{A}$ of a real symmetric matrix $\mathbf{A}$ is the ordered triple (π,ν,δ), where π is the number of positive, ν the number of negative, and δ the number of zero characteristic roots of $\mathbf{A}$. Thus for a symmetric matrix $\mathbf{A}$ we have that $\pi+\nu=\mathrm{r}(\mathbf{A})$, the rank of $\mathbf{A}$, and $\delta=\psi(\mathbf{A})$, the nullity of $\mathbf{A}$. In this paper, 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 [36]) use positive semidefinite where we use nonnegative definite.

The symbol $\sim$ following a random variable means distributed as. Other symbols used in statistics are: $\mathcal{E}$ for expected value, $\mathcal{V}$ and $\mathcal{C}$ 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 $\mathcal{B}$.

II. EARLY RESULTS ON SCHUR COMPLEMENTS

We are concerned with matrices of the form

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

Emilie V. Haynsworth [36, p. 74] appears to have been the first author tc 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

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

The notation

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

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 [66, Hilfssatz, pp. 216–217]

THEOREM 2.1 (Schur [66]). *Let the matrix* **E** *in* (2.2) *be nonsingular. Then*

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

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

Proof [5, p. 51]. We may write

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

taking determinants, we obtain (2.4). ∎

Similarly, it may be shown that if the matrix **H** in (2.2) is nonsingular, then

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

In the notation (2.3) we thus see that

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

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

COROLLARY 2.1. *Let* **F** *be* $m\times n$ *and* **G** $n\times m$. *Then*

$$|\mathbf{I}_m-\mathbf{F}\mathbf{G}|=|\mathbf{I}_n-\mathbf{G}\mathbf{F}|. \tag{2.8}$$

Proof. Put $\mathbf{E}=\mathbf{I}_m$ and $\mathbf{H}=\mathbf{I}_n$ in (2.2). Then (2.8) follows at once using (2.4) and (2.6). ∎

An alternate proof of (2.8), due to George Tiao, is given in the Appendix to the paper by Irwin Guttman [31].

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

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

Then

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

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

$$\mathbf{B}=\begin{pmatrix}\mathbf{E}+x\mathbf{I} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix}; \tag{2.11}$$

then $(2.9)\Leftrightarrow(\mathbf{E}+x\mathbf{I})\mathbf{G}=\mathbf{G}(\mathbf{E}+x\mathbf{I})$. Thus

$$|\mathbf{B}|=|\mathbf{EH}+x\mathbf{H}-\mathbf{GF}|; \tag{2.12}$$

as $x\to 0$, the matrix $\mathbf{B}\to\mathbf{A}$ and (2.12) becomes (2.10). ∎

It is easily seen that (2.10) need not imply (2.9), since when $\mathbf{F}=\mathbf{0}$ 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 $\mathbf{EH}-\mathbf{GF}$ is nonsingular. In a paper by Herstein and Small (1975) it is shown that, for a fairly wide class $\mathfrak{R}$ of rings, if the matrix (2.2) is over $\mathfrak{R}$, where **E**, **F**, **G**, and **H** are all $n\times n$ over $\mathfrak{R}$ and (2.9) holds, then **A** is invertible if and only if $\mathbf{EH}-\mathbf{GF}$ is invertible. The authors state, as an example, that the result is true when $\mathfrak{R}$ is a (right) artinian ring.

In the book by E. Bodewig [11], the formula (2.4) is said [11, 1st ed., p. 189; 2nd ed., p. 218] to date from Frobenius (1849–1917), who obtained the following theorem [29, p. 405]:

THEOREM 2.3 (Frobenius [29]). *Consider the matrix*

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{f} \\ \mathbf{g}' & h \end{pmatrix}, \tag{2.13}$$

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

$$|\mathbf{A}| = h|\mathbf{E}| - \mathbf{g}'(\operatorname{adj}\mathbf{E})\mathbf{f}, \tag{2.14}$$

where adj *denotes adjugate matrix.*

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

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

follows. Since

$$\mathbf{E}^{-1} = \frac{\operatorname{adj}\mathbf{E}}{|\mathbf{E}|}, \tag{2.16}$$

(2.14) follows. Now suppose that $|\mathbf{E}| = 0$. Then $|\mathbf{E} + x\mathbf{I}| \neq 0$ for all $x \neq -\operatorname{ch}(\mathbf{E})$. Let $\mathbf{B}$ be defined similarly to (2.11):

$$\mathbf{B} = \begin{pmatrix} \mathbf{E} + x\mathbf{I} & \mathbf{f} \\ \mathbf{g}' & h \end{pmatrix}; \tag{2.17}$$

then

$$|\mathbf{B}| = h|\mathbf{E} + x\mathbf{I}| - \mathbf{g}'[\operatorname{adj}(\mathbf{E} + x\mathbf{I})]\mathbf{f}. \tag{2.18}$$

As $x \to 0$, the matrix $\mathbf{B} \to \mathbf{A}$ and (2.18) becomes (2.14). ■

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

$$|\mathbf{A}| = h\left|\mathbf{E} - \frac{\mathbf{f}\mathbf{g}'}{h}\right| \tag{2.19a}$$

$$= \frac{|h\mathbf{E} - \mathbf{f}\mathbf{g}'|}{h^{n-1}}, \tag{2.19b}$$

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

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

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

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

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

THEOREM 2.4 (Bodewig [11, 1st ed., p. 36; 2nd ed., p. 42]). *Let the matrix* $\mathbf{E}$ *be nonsingular and let the matrix* $\mathbf{B}$ *have rank* 1. *Then*

$$|\mathbf{E}+\mathbf{B}|=(1+\operatorname{tr}\mathbf{E}^{-1}\mathbf{B})|\mathbf{E}|, \tag{2.21}$$

where tr *denotes trace.*

Proof. Since $\mathbf{B}$ is of rank 1, we may write

$$\mathbf{B}=\mathbf{f}\mathbf{g}' \tag{2.22}$$

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

$$|\mathbf{E}+\mathbf{B}|=(1+\operatorname{tr}\mathbf{E}^{-1}\mathbf{f}\mathbf{g}')|\mathbf{E}|, \tag{2.23}$$

and using (2.22), (2.21) follows. ∎

When $\mathbf{E}$ is singular Theorem 2.4 reduces to

COROLLARY 2.2 (Bodewig [11, 1st ed., p. 36; 2nd ed., p. 42]). *Let the matrix* $\mathbf{E}$ *be singular, and let the unit-rank matrix* $\mathbf{B}$ *be defined by* (2.22). *Then*

$$|\mathbf{E}+\mathbf{B}|=\operatorname{tr}[(\operatorname{adj}\mathbf{E})\mathbf{B}]=\mathbf{g}'(\operatorname{adj}\mathbf{E})\mathbf{f}. \tag{2.24}$$

We will see later, in Sec. 2.4, how Schur complements are related to Jacobi's theorem on the determinant of a minor (42; cf. [52, p. 25]).

2.2. *Rank*

Schur's determinant formula shows that the partitioned matrix

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix} \tag{2.24a}$$

is singular whenever the Schur complement $\mathbf{S}=(\mathbf{A}/\mathbf{E})=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F}$ is singular ($\mathbf{E}$ is assumed to be nonsingular). This result may be strengthened to show that *rank is additive* on the Schur complement, viz.

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

cf. [47, p. 291].

THEOREM 2.5 (Louis Guttman [32]). *Let the matrix* $\mathbf{E}$ *in* (2.2) *be nonsingular. Then*

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

where $\mathrm{r}(\cdot)$ *denotes rank.*

Proof. Since $\mathbf{E}$ is nonsingular we may write [cf. (2.5)]

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

which yields (2.25). ■

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

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

and when $\mathbf{H}$ is nonsingular,

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

Theorem 2.5 readily yields

COROLLARY 2.3 (Louis Guttman [32]). *If* $\mathbf{A}$ *and* $\mathbf{E}$ *in* (2.2) *are both nonsingular, then the Schur complement* $(\mathbf{A}/\mathbf{E})=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F}$ *is also nonsingular.*

In the book by Wedderburn [73, p. 69] a rank reduction procedure is presented which turns out to be a special case of (2.25). Let the matrix $\mathbf{H}$ be nonnull. Then there clearly exist vectors $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a}'\mathbf{H}\mathbf{b}\neq 0$.

Consider the matrix

$$\mathbf{A}=\begin{pmatrix}\mathbf{a}'\mathbf{Hb} & \mathbf{a}'\mathbf{H}\\ \mathbf{Hb} & \mathbf{H}\end{pmatrix}=\begin{pmatrix}\mathbf{a}'\\ \mathbf{I}\end{pmatrix}\mathbf{H}(\mathbf{b},\mathbf{I}). \tag{2.28}$$

Applying (2.25) yields

$$\mathrm{r}(\mathbf{A})=\mathrm{r}(\mathbf{a}'\mathbf{Hb})+\mathrm{r}(\mathbf{H}-\mathbf{Hba}'\mathbf{H}/\mathbf{a}'\mathbf{Hb})=\mathrm{r}(\mathbf{H}) \tag{2.29}$$

using (2.28), and so we have proved:

THEOREM 2.6a (Wedderburn [73, p. 69]). *If the matrix* $\mathbf{H}\neq 0$, *then there exist vectors* $\mathbf{a},\mathbf{b}$ *such that* $\mathbf{a}'\mathbf{Hb}\neq 0$ *and*

$$\mathrm{r}(\mathbf{H}-\mathbf{Hba}'\mathbf{H}/\mathbf{a}'\mathbf{Hb})=\mathrm{r}(\mathbf{H})-1. \tag{2.30}$$

Theorem 2.6a was extended in the book by Householder [41] as an exercise.

THEOREM 2.6b (Householder, [41, p. 33]). *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be column vectors. Then for* $\lambda\neq 0$

$$\mathrm{r}(\mathbf{H}-\mathbf{uv}'/\lambda)<\mathrm{r}(\mathbf{H}) \tag{2.31}$$

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

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

$$\mathbf{A}=\begin{pmatrix}\lambda & \mathbf{v}'\\ \mathbf{u} & \mathbf{H}\end{pmatrix}. \tag{2.32}$$

Using (2.25), we obtain

$$\mathrm{r}(\mathbf{H})\leqslant \mathrm{r}(\mathbf{A})=1+\mathrm{r}\left(\mathbf{H}-\frac{\mathbf{uv}'}{\lambda}\right)<1+\mathrm{r}(\mathbf{H}) \tag{2.33}$$

when (2.31) holds. Hence $\mathrm{r}(\mathbf{H})=\mathrm{r}(\mathbf{A})$, and so there exist vectors $\mathbf{a}$ and $\mathbf{b}$ such that

$$(\lambda,\mathbf{v}')=\mathbf{a}'(\mathbf{u},\mathbf{H}), \tag{2.34a}$$

$$\begin{pmatrix}\lambda\\ \mathbf{u}\end{pmatrix}=\begin{pmatrix}\mathbf{v}'\\ \mathbf{H}\end{pmatrix}\mathbf{b}, \tag{2.34b}$$

and thus $\mathbf{u}=\mathbf{Hb}$, $\mathbf{v}'=\mathbf{a}'\mathbf{H}$, and $\lambda=\mathbf{a}'\mathbf{u}=\mathbf{v}'\mathbf{b}=\mathbf{a}'\mathbf{Hb}$. ∎

Wedderburn [73, p. 68] derived (2.30) using "the Lagrange method of reducing quadratic forms to a normal form." Rao [58, p. 69] refers to Theorem 2.6a as "Lagrange's Theorem"; for an extension see Sec. 4.6, Theorem 4.11.

2.3. *Matrix inversion*

Banachiewicz [5, p. 54] appears to have been 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 [5] obtained (2.37) in Cracovian notation, where matrices are multiplied column by column (see Appendix for further details); he also rediscovered Theorem 2.1 and proved it using (2.5).

THEOREM 2.7 (Banachiewicz [5]; Frazer, Duncan, and Collar [27, p. 113]). *Suppose that*

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix} \tag{2.35}$$

and $\mathbf{E}$ *are both nonsingular. Then the Schur complement*

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F} \tag{2.36}$$

is also nonsingular and

$$\mathbf{A}^{-1}=\begin{pmatrix}\mathbf{E}^{-1}+\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-1}\mathbf{G}\mathbf{E}^{-1} & -\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-1}\\ -\mathbf{S}^{-1}\mathbf{G}\mathbf{E}^{-1} & \mathbf{S}^{-1}\end{pmatrix}$$

$$=\begin{pmatrix}\mathbf{E}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}+\begin{pmatrix}\mathbf{E}^{-1}\mathbf{F}\\ -\mathbf{I}\end{pmatrix}\mathbf{S}^{-1}(\mathbf{G}\mathbf{E}^{-1},-\mathbf{I}). \tag{2.37}$$

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

$$\mathbf{A}^{-1}=\begin{pmatrix}\mathbf{I} & -\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-1}\\ \mathbf{0} & \mathbf{S}^{-1}\end{pmatrix}\begin{pmatrix}\mathbf{E}^{-1} & \mathbf{0}\\ -\mathbf{G}\mathbf{E}^{-1} & \mathbf{I}\end{pmatrix}, \tag{2.38}$$

which yields (2.37). ∎

COROLLARY 2.4 (Duncan [21]). *Suppose that both* $\mathbf{A}$, *given by* (2.35), *and* $\mathbf{H}$ *are nonsingular. Then the Schur complement*

$$\mathbf{T}=(\mathbf{A}/\mathbf{H})=\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G} \tag{2.39}$$

is nonsingular, and

$$\mathbf{A}^{-1}=\begin{pmatrix} \mathbf{T}^{-1} & -\mathbf{T}^{-1}\mathbf{F}\mathbf{H}^{-1} \\ -\mathbf{H}^{-1}\mathbf{G}\mathbf{T}^{-1} & \mathbf{H}^{-1}+\mathbf{H}^{-1}\mathbf{G}\mathbf{T}^{-1}\mathbf{F}\mathbf{H}^{-1} \end{pmatrix}$$

$$=\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^{-1} \end{pmatrix}+\begin{pmatrix} -\mathbf{I} \\ \mathbf{H}^{-1}\mathbf{G} \end{pmatrix}\mathbf{T}^{-1}(-\mathbf{I},\mathbf{F}\mathbf{H}^{-1}). \tag{2.40}$$

Hotelling [40], moreover, noted that if $\mathbf{A}$, $\mathbf{E}$, and $\mathbf{H}$ are all nonsingular, then

$$\mathbf{A}^{-1}=\begin{pmatrix} \mathbf{T}^{-1} & -\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-1} \\ -\mathbf{H}^{-1}\mathbf{G}\mathbf{T}^{-1} & \mathbf{S}^{-1} \end{pmatrix}, \tag{2.41}$$

which involves four inverses, while (2.37) and (2.40) each require only two (cf. [72]). Duncan [21] observed that (2.37) = (2.40), so that

$$(\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G})^{-1}=\mathbf{E}^{-1}+\mathbf{E}^{-1}\mathbf{F}(\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F})^{-1}\mathbf{G}\mathbf{E}^{-1}, \tag{2.42}$$

which Woodbury [76] reestablished.

THEOREM 2.8 (Woodbury [76]). *Let*

$$\mathbf{A}=\begin{pmatrix} \mathbf{E} & -\mathbf{F}\mathbf{H} \\ \mathbf{H}\mathbf{G} & \mathbf{H} \end{pmatrix}, \tag{2.43}$$

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

$$(\mathbf{E}+\mathbf{F}\mathbf{H}\mathbf{G})^{-1}=\mathbf{E}^{-1}-\mathbf{E}^{-1}\mathbf{F}\mathbf{H}(\mathbf{H}+\mathbf{H}\mathbf{G}\mathbf{E}^{-1}\mathbf{F}\mathbf{H})^{-1}\mathbf{H}\mathbf{G}\mathbf{E}^{-1} \tag{2.44a}$$

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

Proof. Since $\mathbf{E}$ is nonsingular, we may write

$$\mathrm{r}(\mathbf{A})=\mathrm{r}(\mathbf{E})+\mathrm{r}(\mathbf{A}/\mathbf{E})=\mathrm{r}(\mathbf{E})+\mathrm{r}(\mathbf{H}+\mathbf{H}\mathbf{G}\mathbf{E}^{-1}\mathbf{F}\mathbf{H}) \tag{2.45}$$

using (2.25). Assume $\mathbf{H}$ is of rank h. We may write

$$\mathbf{H}=\mathbf{K}\mathbf{L}' \tag{2.46}$$

as a full-rank decomposition, where $\mathbf{K}$ and $\mathbf{L}$ have full column rank h. The matrix $\mathbf{A}$ may now be written as

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

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

Then

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

$$=\mathrm{r}(\mathbf{I}_h)+\mathrm{r}(\mathbf{E}+\mathbf{FKL}'\mathbf{G}) \tag{2.50}$$

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

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

Woodbury [76] implied that (2.44a) might hold if $\mathbf{H}$ is singular. However, this cannot be, for if $\mathbf{E}$ and its Schur complement $\mathbf{H}+\mathbf{HGE}^{-1}\mathbf{FH}=(\mathbf{A}/\mathbf{E})$ are both nonsingular, then by (2.45) $\mathbf{A}$ must be nonsingular. Thus, using (2.51), the fact that both $\mathbf{A}$ and $\mathbf{E}+\mathbf{FHG}$ are nonsingular implies $\mathbf{H}$ nonsingular.

From Theorem 2.8 readily follows:

Corollary 2.5 (Woodbury [76]). *Suppose that*

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

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

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

and the Schur complement

$$(\mathbf{A}/h)=\mathbf{E}+h\mathbf{fg}' \tag{2.54}$$

is nonsingular; and

$$(\mathbf{E}+h\mathbf{fg}')^{-1}=\mathbf{E}^{-1}-\frac{h\mathbf{E}^{-1}\mathbf{fg}'\mathbf{E}^{-1}}{1+h\mathbf{g}'\mathbf{E}^{-1}\mathbf{f}}. \tag{2.55}$$

Woodbury [76] observed that John W. Tukey independently found that

$$(\mathbf{I}+h\mathbf{f}\mathbf{g}')^{-1}=\mathbf{I}-\frac{h\mathbf{f}\mathbf{g}'}{1+h\mathbf{g}'\mathbf{f}}, \tag{2.56}$$

which follows immediately by substituting $\mathbf{E}=\mathbf{I}$ in (2.55).

COROLLARY 2.6 (Bartlett [6]). *Suppose that both* $\mathbf{A}$, *given by* (2.52), *and* $\mathbf{E}$ *are nonsingular. Let* $h=1$. *Then the Schur complement*

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

and the Schur complement

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

is nonsingular; and

$$(\mathbf{E}+\mathbf{f}\mathbf{g}')^{-1}=\mathbf{E}^{-1}-\frac{\mathbf{E}^{-1}\mathbf{f}\mathbf{g}'\mathbf{E}^{-1}}{1+\mathbf{g}'\mathbf{E}^{-1}\mathbf{f}}. \tag{2.59}$$

Sherman and Morrison [68, 69] obtained the following results, which are all special cases of Corollary 2.5;

$$(\mathbf{E}+h\mathbf{e}_i\mathbf{g}')^{-1}=\mathbf{E}^{-1}-\frac{h\mathbf{E}^{-1}\mathbf{e}_i\mathbf{g}'\mathbf{E}^{-1}}{1+h\mathbf{g}'\mathbf{E}^{-1}\mathbf{e}_i}, \qquad h\mathbf{g}'\mathbf{E}^{-1}\mathbf{e}_i\neq -1, \tag{2.60}$$

$$(\mathbf{E}+h\mathbf{f}\mathbf{e}_j')^{-1}=\mathbf{E}^{-1}-\frac{h\mathbf{E}^{-1}\mathbf{f}\mathbf{e}_j'\mathbf{E}^{-1}}{1+h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{f}}, \qquad h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{f}\neq -1, \tag{2.61}$$

$$(\mathbf{E}+h\mathbf{e}_i\mathbf{e}_j')^{-1}=\mathbf{E}^{-1}-\frac{h\mathbf{E}^{-1}\mathbf{e}_i\mathbf{e}_j'\mathbf{E}^{-1}}{1+h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{e}_i}, \qquad h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{e}_i\neq -1, \tag{2.62}$$

where $\mathbf{e}_k$ denotes a column vector with all elements zero except for unity in the kth position.

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

$$(\mathbf{E}+\mathbf{e}_i\mathbf{g}')^{-1}=\mathbf{E}^{-1}-\frac{\mathbf{E}^{-1}\mathbf{e}_i\mathbf{g}'\mathbf{E}^{-1}}{1+\mathbf{g}'\mathbf{E}^{-1}\mathbf{e}_i}. \tag{2.63}$$

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

$$(\mathbf{E}+\mathbf{f}\mathbf{e}_j')^{-1}=\mathbf{E}^{-1}-\frac{\mathbf{E}^{-1}\mathbf{f}\mathbf{e}_j'\mathbf{E}^{-1}}{1+\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{f}}. \tag{2.64}$$

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

It is well known (see, e.g., [2, p. 53]) that for any square matrix $\mathbf{A}$

$$\mathbf{A}(\operatorname{adj}\mathbf{A})=(\operatorname{adj}\mathbf{A})\mathbf{A}=|\mathbf{A}|\mathbf{I}, \tag{2.65}$$

and so if $\mathbf{A}$ is $n\times n$, taking determinants of (2.65) yields

$$|\operatorname{adj}\mathbf{A}|=|\mathbf{A}|^{n-1}, \tag{2.66}$$

which is due to Cauchy (1812) [16]. This result was extended by Jacobi (1834) [42] as follows (see also [2, p. 103]):

THEOREM 2.9 (Jacobi [42]). *Consider the $n\times n$ matrix*

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix}, \tag{2.67}$$

where $\mathbf{E}$ *is* $m\times m$. *Let*

$$\mathbf{A}^*=\operatorname{adj}\mathbf{A}=\begin{pmatrix}\mathbf{E}^* & \mathbf{F}^*\\ \mathbf{G}^* & \mathbf{H}^*\end{pmatrix}, \tag{2.68}$$

where $\mathbf{E}^*$ *is* $m\times m$. *Then*

$$|\mathbf{H}^*|=|\mathbf{A}|^{n-m-1}|\mathbf{E}|, \qquad m=0,1,\dots,n-1. \tag{2.69}$$

Proof. When $m=0$, the matrix $\mathbf{E}$ disappears and (2.69) reduces to (2.66). When $m=n-1$, (2.69) is trivially true. So assume $1\leqslant m\leqslant n-2$. If

$|\mathbf{A}|=0$, then $\mathrm{r}(\mathbf{H}^*)\leqslant \mathrm{r}(\operatorname{adj}\mathbf{A})\leqslant 1$ in view of (2.65), and with $n-m\geqslant 2$ it follows that $|\mathbf{H}^*|=0$ and so (2.69) holds. Now assume $|\mathbf{A}|\neq 0$. Then (2.65) implies that

$$\operatorname{adj}\mathbf{A}=|\mathbf{A}|\mathbf{A}^{-1}. \tag{2.70}$$

Suppose first that $|\mathbf{E}|\neq 0$. We may write

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

using (2.37). Taking determinants, we obtain

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

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

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

$$n=\mathrm{r}(\operatorname{adj}\mathbf{A})=\mathrm{r}(\mathbf{H}^*)+\mathrm{r}(\operatorname{adj}\mathbf{A}/\mathbf{H}^*)=n-m+\mathrm{r}(\operatorname{adj}\mathbf{A}/\mathbf{H}^*), \tag{2.73}$$

and so $(\operatorname{adj}\mathbf{A}/\mathbf{H}^*)$ is nonsingular. The inverse of the Schur complement of $\mathbf{H}^*$ in $\operatorname{adj}\mathbf{A}$ is $\mathbf{E}/|\mathbf{A}|$; cf. (2.40). Hence $|\mathbf{E}|\neq 0$, a contradiction. Thus $|\mathbf{E}|=0$ implies $|\mathbf{H}^*|=0$, and so (2.69) holds. ∎

Similarly, it may be shown that

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

III. RECENT RESULTS ON SCHUR COMPLEMENTS

In Sec. II 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 [18] 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 [36] of the *inertia* of a partitioned real symmetric matrix.

3.1. Inertia

The inertia of a symmetric matrix $\mathbf{A}$ is the ordered triple

$$\operatorname{In}\mathbf{A}=(\pi,\nu,\delta), \tag{3.1}$$

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

$$\operatorname{In}\mathbf{A}=\operatorname{In}\mathbf{CAC}' \tag{3.2}$$

for every nonsingular matrix $\mathbf{C}$ [46, p. 83; 52, p. 377]. The equation (3.2) is called *Sylvester's law of inertia*.

THEOREM 3.1 (Haynsworth [36]). *Consider the $(m+n)\times(m+n)$ symmetric matrix*

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{F}' & \mathbf{H}\end{pmatrix}, \tag{3.3}$$

where $\mathbf{E}$ *is* $m\times m$ *nonsingular. Then*

$$\operatorname{In}\mathbf{A}=\operatorname{In}\mathbf{E}+\operatorname{In}(\mathbf{A}/\mathbf{E}). \tag{3.4}$$

Proof. We may write [cf. (2.26)]

$$\mathbf{B}=\begin{pmatrix}\mathbf{I}_m & \mathbf{0}\\ -\mathbf{F}'\mathbf{E}^{-1} & \mathbf{I}_n\end{pmatrix}\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{F}' & \mathbf{H}\end{pmatrix}\begin{pmatrix}\mathbf{I}_m & -\mathbf{E}^{-1}\mathbf{F}\\ \mathbf{0} & \mathbf{I}_n\end{pmatrix}=\begin{pmatrix}\mathbf{E} & \mathbf{0}\\ \mathbf{0} & (\mathbf{A}/\mathbf{E})\end{pmatrix}. \tag{3.5}$$

Using (3.2),

$$\operatorname{In}\mathbf{B}=\operatorname{In}\mathbf{A}, \tag{3.6}$$

and since the characteristic roots of $\mathbf{B}$ are those of $\mathbf{E}$ and of $(\mathbf{A}/\mathbf{E})$, (3.4) follows. ∎

A matrix of the form $\mathbf{X}'\mathbf{X}$ is said to be *Gramian* or *nonnegative definite* (nnd). If $\mathbf{X}'\mathbf{X}$ is singular, then it will be called *positive semidefinite* (psd). If

$\mathbf{X'X}$ is nonsingular, then it is called *positive definite* (pd). Some authors (e.g., Haynsworth [36]) 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 $\mathbf{A}$ is nonnegative definite $\Leftrightarrow$ $\nu=0$, positive definite $\Leftrightarrow$ $\nu=\delta=0$, positive semidefinite $\Leftrightarrow$ $\{\nu=0 \text{ and } \delta \geqslant 1\}$.

COROLLARY 3.1 (Haynsworth [36]). *Consider the symmetric matrix*

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

and let $\mathbf{E}$ *be positive definite. Then*

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

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

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

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

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

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

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

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

Proof. Consider the matrix

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

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

Haynsworth [36] extended Theorem 3.1 by considering the inertia of partitioned Schur complements. We begin by partitioning the $(m+n)\times(m+n)$ symmetric matrix $\mathbf{A}$ as in (3.3), where $\mathbf{E}$ is $m\times m$ nonsingular. We then compute the Schur complement $\mathbf{S}=(\mathbf{A}/\mathbf{E})$ and obtain (3.4). We partition the Schur complement

$$(\mathbf{A}/\mathbf{E})=\mathbf{S}=\begin{pmatrix}\mathbf{E}_1 & \mathbf{F}_1\\ \mathbf{F}_1' & \mathbf{H}_1\end{pmatrix}, \tag{3.12}$$

where $\mathbf{E}_1$ is $m_1\times m_1$, nonsingular, and compute the Schur complement $\mathbf{S}_1=(\mathbf{S}/\mathbf{E}_1)$. We obtain

$$\operatorname{In}\mathbf{A}=\operatorname{In}\mathbf{E}+\operatorname{In}\mathbf{E}_1+\operatorname{In}\mathbf{S}_1. \tag{3.13}$$

We partition

$$\mathbf{S}_1=\begin{pmatrix}\mathbf{E}_2 & \mathbf{F}_2\\ \mathbf{F}_2' & \mathbf{H}_2\end{pmatrix}, \tag{3.14}$$

where $\mathbf{E}_2$ is $m_2\times m_2$, nonsingular. We compute $\mathbf{S}_2=(\mathbf{S}_1/\mathbf{E}_2)$ and repeat the procedure performed with $\mathbf{S}_1$. We obtain

$$\operatorname{In}\mathbf{A}=\operatorname{In}\mathbf{E}+\operatorname{In}\mathbf{E}_1+\operatorname{In}\mathbf{E}_2+\operatorname{In}\mathbf{S}_2. \tag{3.14a}$$

We may continue this process by defining $\mathbf{E}_{i+1}$ 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 $\mathbf{S}_i=(\mathbf{S}_{i-1}/\mathbf{E}_i)$. The process stops as soon as a Schur complement, $\mathbf{S}_k$ say, is a scalar or has no top left nonsingular submatrix. Then

$$\operatorname{In}\mathbf{A}=\operatorname{In}\mathbf{E}+\sum_{j=1}^{k}\operatorname{In}\mathbf{E}_j+\operatorname{In}\mathbf{S}_k. \tag{3.15}$$

3.2. *The quotient property and related determinant inequalities*

Consider the matrix

$$\mathbf{A}=\left(\begin{array}{c:c}\mathbf{E} & \mathbf{F}\\ \hdashline \mathbf{G} & \mathbf{H}\end{array}\right)=\left[\begin{array}{cc:c}\mathbf{K} & \mathbf{L} & \mathbf{F}_1\\ \mathbf{M} & \mathbf{N} & \mathbf{F}_2\\ \hdashline \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{H}\end{array}\right], \tag{3.16}$$

where $\mathbf{E}$ and $\mathbf{K}$ are nonsingular. Then the Schur complement $(\mathbf{E}/\mathbf{K})$ is a nonsingular leading principal submatrix of the Schur complement $(\mathbf{A}/\mathbf{K})$;

Crabtree and Haynsworth [19] and Ostrowski [55] proved that

$$(\mathbf{A}/\mathbf{E})=((\mathbf{A}/\mathbf{K})/(\mathbf{E}/\mathbf{K})), \tag{3.17}$$

which they called the *quotient property*. We note that the parallel relationship

$$\mathbf{A}\mathbf{E}^{-1}=(\mathbf{A}\mathbf{K}^{-1})(\mathbf{E}\mathbf{K}^{-1})^{-1} \tag{3.18}$$

also holds.

THEOREM 3.3 (Crabtree and Haynsworth [19]). *Consider the matrix* (3.16), *where both* $\mathbf{E}$ *and* $\mathbf{K}$ *are nonsingular. Then the Schur complement* $(\mathbf{E}/\mathbf{K})$ *is a nonsingular leading principal submatrix of the Schur complement* $(\mathbf{A}/\mathbf{K})$. *Moreover*, (3.17) *holds*.

Proof. Since

$$(\mathbf{A}/\mathbf{K})=\begin{pmatrix}\mathbf{N} & \mathbf{F}_2\\ \mathbf{G}_2 & \mathbf{H}\end{pmatrix}-\begin{pmatrix}\mathbf{M}\\ \mathbf{G}_1\end{pmatrix}\mathbf{K}^{-1}(\mathbf{L},\mathbf{F}_1) \tag{3.19}$$

$$=\begin{pmatrix}\mathbf{N}-\mathbf{M}\mathbf{K}^{-1}\mathbf{L} & \mathbf{F}_2-\mathbf{M}\mathbf{K}^{-1}\mathbf{F}_1\\ \mathbf{G}_2-\mathbf{G}_1\mathbf{K}^{-1}\mathbf{L} & \mathbf{H}-\mathbf{G}_1\mathbf{K}^{-1}\mathbf{F}_1\end{pmatrix}, \tag{3.20}$$

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

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

and so $(\mathbf{E}/\mathbf{K})$ is nonsingular. Also,

$$((\mathbf{A}/\mathbf{K})/(\mathbf{E}/\mathbf{K}))$$

$$=\mathbf{H}-\mathbf{G}_1\mathbf{K}^{-1}\mathbf{F}_1-(\mathbf{G}_2-\mathbf{G}_1\mathbf{K}^{-1}\mathbf{L})(\mathbf{E}/\mathbf{K})^{-1}(\mathbf{F}_2-\mathbf{M}\mathbf{K}^{-1}\mathbf{F}_1) \tag{3.22}$$

$$=\mathbf{H}-(\mathbf{G}_1,\mathbf{G}_2)\begin{bmatrix}\mathbf{K}^{-1}+\mathbf{K}^{-1}\mathbf{L}(\mathbf{E}/\mathbf{K})^{-1}\mathbf{M}\mathbf{K}^{-1} & -\mathbf{K}^{-1}\mathbf{L}(\mathbf{E}/\mathbf{K})^{-1}\\ -(\mathbf{E}/\mathbf{K})^{-1}\mathbf{M}\mathbf{K}^{-1} & (\mathbf{E}/\mathbf{K})^{-1}\end{bmatrix}\begin{pmatrix}\mathbf{F}_1\\ \mathbf{F}_2\end{pmatrix} \tag{3.23}$$

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

using (2.37). ■

Haynsworth [38] extended Theorem 3.2 by showing that if

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

[cf. (3.7)] and

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

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

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

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

LEMMA 3.1 (Haynsworth [38]). *Let* $\mathbf{E}$ *and* $\mathbf{K}$ *both be* $m\times m$ *positive definite. Then if* $\mathbf{F}$ *and* $\mathbf{L}$ *are arbitrary* $m\times n$ *matrices,*

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

is nonnegative definite with the same rank as

$$\mathbf{F}-\mathbf{E}\mathbf{K}^{-1}\mathbf{L}. \tag{3.29}$$

Proof. We may rewrite (3.28) as follows:

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

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

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

positive definite, and so (3.30) may be written

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

which is nnd with rank equal to the rank of $\mathbf{F}-\mathbf{E}\mathbf{K}^{-1}\mathbf{L}$. ∎

THEOREM 3.4 (Haynsworth [38]). *Let*

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

and

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

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*

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

is nonnegative definite.

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

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

so that the Schur complement

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

is nnd. Hence

$$\begin{aligned}&((\mathbf{A}+\mathbf{B})/(\mathbf{E}+\mathbf{K}))-(\mathbf{A}/\mathbf{E})-(\mathbf{B}/\mathbf{K})\\ &\qquad=\mathbf{F}'\mathbf{E}^{-1}\mathbf{F}+\mathbf{L}'\mathbf{K}^{-1}\mathbf{L}-(\mathbf{F}+\mathbf{L})'(\mathbf{E}+\mathbf{K})^{-1}(\mathbf{F}+\mathbf{L}),\end{aligned} \tag{3.38}$$

which is nnd from Lemma 3.1. ■

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 [38] proved that if **A** and **B** are nonnegative definite and **E** and **K** are positive definite, then

$$|(\mathbf{A}+\mathbf{B})/(\mathbf{E}+\mathbf{K})|=\frac{|\mathbf{A}+\mathbf{B}|}{|\mathbf{E}+\mathbf{K}|}\geqslant\frac{|\mathbf{A}|}{|\mathbf{E}|}+\frac{|\mathbf{B}|}{|\mathbf{K}|}. \tag{3.39}$$

To prove (3.39), we use the following:

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

$$|\mathbf{A}+\mathbf{B}| \geqslant |\mathbf{A}|+|\mathbf{B}|. \tag{3.40}$$

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

$$|\mathbf{A}+\mathbf{B}| = |\mathbf{A}|\cdot|\mathbf{I}+\mathbf{A}^{-1}\mathbf{B}| \tag{3.41}$$

$$= |\mathbf{A}| \prod_{i=1}^{n} \left(1+\mathrm{ch}_i\mathbf{A}^{-1}\mathbf{B}\right) \tag{3.42}$$

$$\geqslant |\mathbf{A}|\cdot\left[1+\prod_{i=1}^{n}\mathrm{ch}_i\mathbf{A}^{-1}\mathbf{B}\right], \tag{3.43}$$

since the characteristic roots $\mathrm{ch}_i\mathbf{A}^{-1}\mathbf{B}\geqslant 0$. Using the fact that

$$|\mathbf{A}|\cdot\left[1+\prod_{i=1}^{n}\mathrm{ch}_i\mathbf{A}^{-1}\mathbf{B}\right] = |\mathbf{A}|\cdot\left[1+|\mathbf{A}^{-1}\mathbf{B}|\right] = |\mathbf{A}|+|\mathbf{B}|, \tag{3.44}$$

(3.40) follows at once. ■

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

$$|\mathbf{E}_i| \geqslant |\mathbf{K}_i|, \qquad i=1,\dots,n. \tag{3.45}$$

COROLLARY 3.2 (Haynsworth [38]). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *both be* $(m+n)\times(m+n)$ *matrices defined by* (3.33) *and* (3.34), *where* $\mathbf{E}$ *and* $\mathbf{K}$ *are both* $m\times m$. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are nonnegative definite and* $\mathbf{E}$ *and* $\mathbf{K}$ *are positive definite, then* (3.39) *holds.*

Proof. From Theorems 2.1 and 3.4 it follows that

$$\frac{|\mathbf{A}+\mathbf{B}|}{|\mathbf{E}+\mathbf{K}|}=|((\mathbf{A}+\mathbf{B})/(\mathbf{E}+\mathbf{K}))|\geqslant|(\mathbf{A}/\mathbf{E})+(\mathbf{B}/\mathbf{K})|$$

$$\geqslant|(\mathbf{A}/\mathbf{E})|+|(\mathbf{B}/\mathbf{K})|=\frac{|\mathbf{A}|}{|\mathbf{E}|}+\frac{|\mathbf{B}|}{|\mathbf{K}|},\tag{3.46}$$

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

We may extend (3.39) using Lemma 3.3.

THEOREM 3.5 (Haynsworth [38]). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *both be* $n\times n$ *nonnegative definite matrices. Suppose further that* $\mathbf{E}_i$ *and* $\mathbf{K}_i$, $i=1,\ldots,n-1$, *are the* $i\times i$ *principal submatrices in the upper left corners of the matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *respectively. If* $\mathbf{E}_1,\ldots,\mathbf{E}_{n-1},\mathbf{K}_1,\ldots,\mathbf{K}_{n-1}$ *are all positive definite, then*

$$|\mathbf{A}+\mathbf{B}|\geqslant|\mathbf{A}|\left(1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right)+|\mathbf{B}|\left(1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right).\tag{3.47}$$

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

$$|\mathbf{A}+\mathbf{B}|=|\mathbf{E}_1+\mathbf{K}_1|\cdot|((\mathbf{A}+\mathbf{B})/(\mathbf{E}_1+\mathbf{K}_1))|\tag{3.48}$$

using (2.4). But

$$|\mathbf{E}_1+\mathbf{K}_1|\geqslant|\mathbf{E}_1|+|\mathbf{K}_1|\tag{3.49}$$

by (3.40), and

$$|((\mathbf{A}+\mathbf{B})/(\mathbf{E}_1+\mathbf{K}_1))|\geqslant\frac{|\mathbf{A}|}{|\mathbf{E}_1|}+\frac{|\mathbf{B}|}{|\mathbf{K}_1|}\tag{3.50}$$

by (3.39). Hence, (3.47) holds for $n=2$. Now assume that (3.47) holds for $\mathbf{A}$ and $\mathbf{B}$ $n\times n$. If $\mathbf{A}_1$ and $\mathbf{B}_1$ are $(n+1)\times(n+1)$ nonnegative definite matrices, and $\mathbf{A}=\mathbf{E}_n$ and $\mathbf{B}=\mathbf{K}_n$ are $n\times n$ positive definite submatrices of $\mathbf{A}_1$ and $\mathbf{B}_1$, respectively, then

$$|\mathbf{A}_1+\mathbf{B}_1|=|\mathbf{E}_n+\mathbf{K}_n|\cdot|((\mathbf{A}_1+\mathbf{B}_1)/(\mathbf{E}_n+\mathbf{K}_n))|,\tag{3.51}$$

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

$$|\mathbf{E}_n+\mathbf{K}_n| \geqslant |\mathbf{E}_n|\left(1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right)+|\mathbf{K}_n|\left(1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right) \tag{3.52}$$

and by (3.39),

$$|((\mathbf{A}_1+\mathbf{B}_1)/(\mathbf{E}_n+\mathbf{K}_n))| \geqslant \frac{|\mathbf{A}_1|}{|\mathbf{E}_n|}+\frac{|\mathbf{B}_1|}{|\mathbf{K}_n|}. \tag{3.53}$$

Hence

$$\begin{aligned}
|\mathbf{A}_1+\mathbf{B}_1| &\geqslant \left\{|\mathbf{E}_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{K}_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]\right\}\\
&\quad\times\left\{\frac{|\mathbf{A}_1|}{|\mathbf{E}_n|}+\frac{|\mathbf{B}_1|}{|\mathbf{K}_n|}\right\}\\
&=|\mathbf{A}_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+\frac{|\mathbf{E}_n|}{|\mathbf{K}_n|}|\mathbf{B}_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]\\
&\quad+\frac{|\mathbf{K}_n|}{|\mathbf{E}_n|}|\mathbf{A}_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]+|\mathbf{B}_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]\\
&\geqslant |\mathbf{A}_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+\frac{|\mathbf{E}_n|\cdot|\mathbf{B}_1|}{|\mathbf{K}_n|}\\
&\quad+\frac{|\mathbf{K}_n|\cdot|\mathbf{A}_1|}{|\mathbf{E}_n|}+|\mathbf{B}_1|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]\\
&=|\mathbf{A}_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{B}_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right].
\end{aligned} \tag{3.54}$$

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

In the paper by Haynsworth [38], the formula (3.47) was established with both **A** and **B** positive definite.

COROLLARY 3.3 (Haynsworth [38]). *If* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{A}-\mathbf{B}$ *are* $n\times n$ *positive definite matrices, then*

$$|\mathbf{A}+\mathbf{B}|>|\mathbf{A}|+n|\mathbf{B}|. \tag{3.55}$$

Proof. By Theorem 3.5,

$$|\mathbf{A}+\mathbf{B}|\geqslant|\mathbf{A}|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{B}|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]. \tag{3.56}$$

But, since $\mathbf{A}-\mathbf{B}$ is pd, $|\mathbf{E}_i|>|\mathbf{K}_i|$; cf. (3.45). Hence

$$|\mathbf{A}+\mathbf{B}|>|\mathbf{A}|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+n|\mathbf{B}|>|\mathbf{A}|+n|\mathbf{B}|. \quad \blacksquare \tag{3.57}$$

Hartfiel [33] 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 $x=(b/a)^{1/2}$ and so

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

THEOREM 3.6 (Hartfiel [33]). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *both be* $n\times n$ *nonnegative definite matrices. Suppose further that* $\mathbf{E}_i$ *and* $\mathbf{K}_i$, $i=1,\dots,n-1$, *are the* $i\times i$ *principal submatrices in the upper left corners of the matrices* $\mathbf{A}$ *and* $\mathbf{B}$, *respectively. If* $\mathbf{E}_1,\dots,\mathbf{E}_{n-1},\mathbf{K}_1,\dots,\mathbf{K}_{n-1}$ *are all positive definite, then*

$$|\mathbf{A}+\mathbf{B}|\geqslant|\mathbf{A}|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{B}|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]+(2^n-2n)(|\mathbf{A}|\cdot|\mathbf{B}|)^{1/2}. \tag{3.59}$$

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 $\mathbf{A}$ and $\mathbf{B}$ $n\times n$. If $\mathbf{A}_1$ and $\mathbf{B}_1$ are $(n+1)\times(n+1)$ nonnegative definite, and $\mathbf{A}=\mathbf{E}_n$ and $\mathbf{B}=\mathbf{K}_n$ are $n\times n$ positive definite submatrices of $\mathbf{A}$ and $\mathbf{B}$, respectively, then [cf. (3.51)]

$$|\mathbf{A}_1+\mathbf{B}_1|=|\mathbf{E}_n+\mathbf{K}_n|\cdot|((\mathbf{A}_1+\mathbf{B}_1)/(\mathbf{E}_n+\mathbf{K}_n))|. \tag{3.60}$$

But, by the inductive assumption,

$$|\mathbf{E}_n+\mathbf{K}_n| \geqslant |\mathbf{E}_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{K}_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]$$
$$+(2^n-2n)(|\mathbf{E}_n|\cdot|\mathbf{K}_n|)^{1/2}, \tag{3.61}$$

and using (3.53) it follows that

$$|\mathbf{A}_1+\mathbf{B}_1| \geqslant \left\{|\mathbf{E}_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{K}_n|\cdot\left[1+\sum_{i=1}^{n-1}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]\right.$$
$$\left.+(2^n-2n)(|\mathbf{E}_n|\cdot|\mathbf{K}_n|)^{1/2}\right\}\left[\frac{|\mathbf{A}_1|}{|\mathbf{E}_n|}+\frac{|\mathbf{B}_1|}{|\mathbf{K}_n|}\right] \tag{3.62}$$
$$\geqslant |\mathbf{A}_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{B}_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]$$
$$+\sum_{i=1}^{n-1}\left[\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\frac{|\mathbf{E}_n|}{|\mathbf{K}_n|}|\mathbf{B}_1|+\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\frac{|\mathbf{K}_n|}{|\mathbf{E}_n|}|\mathbf{A}_1|\right]$$
$$+(2^n-2n)(|\mathbf{E}_n|\cdot|\mathbf{K}_n|)^{1/2}\left[\frac{|\mathbf{A}_1|}{|\mathbf{E}_n|}+\frac{|\mathbf{B}_1|}{|\mathbf{K}_n|}\right], \tag{3.63}$$

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

$$\sum_{i=1}^{n-1}\left[\frac{|\mathbf{K}_i|\cdot|\mathbf{E}_n|}{|\mathbf{E}_i|\cdot\mathbf{K}_n|}|\mathbf{B}_1|+\frac{|\mathbf{E}_i|\cdot|\mathbf{K}_n|}{|\mathbf{K}_i|\cdot|\mathbf{E}_n|}|\mathbf{A}_1|\right]$$
$$\geqslant 2(n-1)(|\mathbf{A}_1|\cdot|\mathbf{B}_1|)^{1/2}, \tag{3.64}$$

while using the arithmetic-mean–geometric-mean inequality, we have that

$$\frac{|\mathbf{A}_1|}{|\mathbf{E}_n|}+\frac{|\mathbf{B}_1|}{|\mathbf{K}_n|} \geqslant 2\left[\frac{|\mathbf{A}_1|\cdot|\mathbf{B}_1|}{|\mathbf{E}_n|\cdot|\mathbf{K}_n|}\right]^{1/2} \tag{3.65}$$

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

$$|\mathbf{A}_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}\right]+|\mathbf{B}_1|\cdot\left[1+\sum_{i=1}^{n}\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]$$

$$+\left[2^{n+1}-2(n+1)\right](|\mathbf{A}_1|\cdot|\mathbf{B}_1|)^{1/2},$$

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 $\mathbf{A}$ and $\mathbf{B}$, and the induction proof is complete. ∎

In the paper by Hartfiel [33], the formula (3.59) was established when both $\mathbf{A}$ and $\mathbf{B}$ are positive definite.

COROLLARY 3.4 (Hartfiel [33]). *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $n\times n$ *positive definite matrices, then*

$$|\mathbf{A}+\mathbf{B}|\geqslant|\mathbf{A}|+|\mathbf{B}|+(2^n-2)(|\mathbf{A}|\cdot|\mathbf{B}|)^{1/2}. \tag{3.66}$$

Proof. By Theorem 3.6,

$$|\mathbf{A}+\mathbf{B}|\geqslant|\mathbf{A}|+|\mathbf{B}|+\sum_{i=1}^{n-1}\left[|\mathbf{A}|\frac{|\mathbf{K}_i|}{|\mathbf{E}_i|}+|\mathbf{B}|\frac{|\mathbf{E}_i|}{|\mathbf{K}_i|}\right]$$

$$+(2^n-2n)(|\mathbf{A}|\cdot|\mathbf{B}|)^{1/2}$$

$$\geqslant|\mathbf{A}|+|\mathbf{B}|+2(n-1)(|\mathbf{A}|\cdot|\mathbf{B}|)^{1/2}+(2^n-2n)(|\mathbf{A}|\cdot|\mathbf{B}|)^{1/2}, \tag{3.67}$$

using (3.58), and (3.67) equals

$$|\mathbf{A}|+|\mathbf{B}|+(2^n-2)(|\mathbf{A}|\cdot|\mathbf{B}|)^{1/2}.$$

∎

Corollary 3.4 allows us also to extend Corollary 3.3.

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

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

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

3.3. Characteristic roots

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

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

where $\mathbf{D}$ is diagonal $m \times m$ and $\mathbf{X}$ has full column rank m. More generally, consider [37]

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

where $\mathbf{B}$ is an arbitrary $m \times m$ matrix. Goddard and Schneider [30] call such an $\mathbf{X}$ a *commutator*; they showed that m characteristic roots of $\mathbf{A}$ are characteristic roots of $\mathbf{B}$.

THEOREM 3.7 (Haynsworth [37]). *Let the $n \times n$ matrix*

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

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

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

and

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \tag{3.72}$$

where $\mathbf{X}_1$ is an $m \times m$ nonsingular matrix. Then m characteristic roots of $\mathbf{A}$ are characteristic roots of $\mathbf{B}$, and the remaining $n-m$ characteristic roots are characteristic roots of $(\mathbf{C}/\mathbf{X}_1)$, where

$$\mathbf{C} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{F} \\ \mathbf{X}_2 & \mathbf{H} \end{pmatrix}. \tag{3.73}$$

Proof. Since $\mathbf{X}_1$ is nonsingular, the $n \times n$ matrix

$$\mathbf{J} = \begin{pmatrix} \mathbf{X}_1 & 0 \\ \mathbf{X}_2 & \mathbf{I}_{n-m} \end{pmatrix} \tag{3.74}$$

is nonsingular. Hence, using (3.71),

$$\mathbf{J}^{-1}\mathbf{A}\mathbf{J}=\begin{pmatrix}\mathbf{B} & \mathbf{X}_1^{-1}\mathbf{F}\\ \mathbf{0} & \mathbf{H}-\mathbf{X}_2\mathbf{X}_1^{-1}\mathbf{F}\end{pmatrix} \tag{3.75}$$

has the same characteristic roots as $\mathbf{A}$, and the proof is complete. ∎

COROLLARY 3.6 (Haynsworth [37]). *Let the matrix* $\mathbf{A}$ *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* $\mathbf{X}$ *are the characteristic vectors and that* $\mathbf{X}$ *may be partitioned as in* (3.72). *Then the remaining* $n-m$ *characteristic roots of* $\mathbf{A}$ *are characteristic roots of* $(\mathbf{C}/\mathbf{X}_1)$ *where* $\mathbf{C}$ *is defined by* (3.73).

Proof. Since $\mathbf{AX}=\mathbf{XD}$, where $\mathbf{D}=\operatorname{diag}(\lambda_1,\ldots,\lambda_m)$ [cf. (3.69a)], Theorem 3.7 directly implies the result. ∎

3.4. *Quadratic forms*

An alternative 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 [18]). *Consider the quadratic form*

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

where $\mathbf{A}$ *is symmetric and* $\mathbf{E}$ *is nonsingular. Let* q_r *denote* q *constrained by the system of equations*

$$\mathbf{Ex}+\mathbf{Fy}=\mathbf{0}. \tag{3.77}$$

Then

$$q_r=\mathbf{y}'(\mathbf{A}/\mathbf{E})\mathbf{y}. \tag{3.78}$$

Proof. We may write

$$q=\mathbf{x}'\mathbf{E}\mathbf{x}+2\mathbf{x}'\mathbf{F}\mathbf{y}+\mathbf{y}'\mathbf{H}\mathbf{y}. \tag{3.79}$$

Using (3.77) and the fact that $\mathbf{E}$ is nonsingular, we obtain

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

Substituting (3.80) in (3.79) yields

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

■

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

THEOREM 3.9 (Cottle [18]). *Let q_s denote the quadratic form* (3.76) *constrained by the system of equations*

$$\mathbf{Kx} + \mathbf{Ly} = \mathbf{0}, \tag{3.82}$$

where $\mathbf{K}$ *is nonsingular. Let*

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

and

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

Then

$$q_s = \mathbf{y}'(\mathbf{B}/\mathbf{C})\mathbf{y}. \tag{3.85}$$

Proof. Using (3.82) with $\mathbf{K}$ nonsingular, we obtain

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

Substituting (3.86) in (3.79), we have

$$q_s = \mathbf{y}'\left[\mathbf{H} - 2\mathbf{L}'(\mathbf{K}^{-1})'\mathbf{F} + \mathbf{L}'(\mathbf{K}^{-1})'\mathbf{E}\mathbf{K}^{-1}\mathbf{L}\right]\mathbf{y}. \tag{3.87}$$

Since $\mathbf{K}$ is nonsingular, so is $\mathbf{C}$, and the inverse

$$\mathbf{C}^{-1}=\begin{pmatrix} -(\mathbf{K}')^{-1}\mathbf{E}\mathbf{K}^{-1} & (\mathbf{K}')^{-1} \\ \mathbf{K}^{-1} & \mathbf{0} \end{pmatrix} \tag{3.88}$$

is obtained using (2.40). Hence

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

$$=\mathbf{H}+\mathbf{L}'(\mathbf{K}')^{-1}\mathbf{E}\mathbf{K}^{-1}\mathbf{L}-\mathbf{F}'\mathbf{K}^{-1}\mathbf{L}-\mathbf{L}'(\mathbf{K}')^{-1}\mathbf{F}, \tag{3.89}$$

and so (3.87) = (3.85). ∎

We may combine Theorems 3.8 and 3.9 to obtain

THEOREM 3.10 (Cottle [18]). *Let q_t denote the quadratic form*

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

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

$$q_t=\mathbf{y}'(\mathbf{B}/\mathbf{C})\mathbf{y}. \tag{3.91}$$

IV. RESULTS ON GENERALIZED SCHUR COMPLEMENTS

When the submatrix $\mathbf{E}$ in the partitioned matrix

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

is rectangular, or square but singular, then the definition (2.1) of the Schur complement cannot be used. Using generalized inverses, however, we may define [47, 48]

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-}\mathbf{F} \tag{4.2}$$

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

$$\mathbf{A}\mathbf{A}^-\mathbf{A}=\mathbf{A}. \tag{4.3}$$

Following Rao and Mitra [57, 59], we will call $\mathbf{A}^-$ a *generalized inverse* (or g-inverse) of $\mathbf{A}$. If $\mathbf{A}^-\mathbf{A}\mathbf{A}^-=\mathbf{A}^-$ also holds, then we will call $\mathbf{A}^-$ a *reflexive* g-inverse. Hence, a g-inverse $\mathbf{A}^-$ is reflexive if and only if it has the same rank as $\mathbf{A}$ (see proof in Sec. 4.6). A reflexive g-inverse $\mathbf{A}^-$ such that $\mathbf{A}\mathbf{A}^-$ and $\mathbf{A}^-\mathbf{A}$ are both symmetric is unique and is denoted by $\mathbf{A}^+$, the *Moore-Penrose* g-inverse of $\mathbf{A}$. 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 such that $\mathbf{A}'\mathbf{A}$ and $\mathbf{A}\mathbf{A}'$ are defined and have the same rank as both $\mathbf{A}$ and $\mathbf{A}'$, and such that $(\mathbf{A}'\mathbf{B})'=\mathbf{B}'\mathbf{A}$ [48, p. 438].

Carlson, Haynsworth, and Markham [15] considered matrices over the complex field and used the Moore-Penrose g-inverse $\mathbf{E}^+$ in their definition of the generalized Schur complement. Other writers, such as Rohde [60], Khatri [44], Meyer [49], and Pringle and Rayner [56], used (4.2) without giving it a name. (See also [34, 35].)

4.1. *Determinants*

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

$$\begin{vmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{vmatrix} = |\mathbf{E}|\cdot|\mathbf{H}-\mathbf{G}\mathbf{E}^-\mathbf{F}|, \tag{4.4}$$

need not hold; e.g.,

$$\left|\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & 1 \end{array}\right| = -1 \neq 0 = |0|\cdot|1-1\cdot 0^-\cdot 1|. \tag{4.5}$$

Sufficient conditions for (4.4) to hold, however, were obtained by Carlson, Haynsworth, and Markham [15] using Moore-Penrose g-inverses. We extend their results to arbitrary g-inverses using the following:

LEMMA 4.1 (Marsaglia & Styan [47, p. 274, Theorem 5]). *For matrices over an arbitrary field,*

$$r(\mathbf{E},\mathbf{F})=r(\mathbf{E})+r([\mathbf{I}-\mathbf{E}\mathbf{E}^-]\mathbf{F})=r([\mathbf{I}-\mathbf{F}\mathbf{F}^-]\mathbf{E})+r(\mathbf{F}) \tag{4.6}$$

for every $\mathbf{E}^-,\mathbf{F}^-$, *and*

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

for every $\mathbf{E}^-,\mathbf{G}^-$.

Proof. We may write

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

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

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

since the column spaces of $\mathbf{E}$ and $(\mathbf{I}-\mathbf{E}\mathbf{E}^-)\mathbf{F}$ are virtually disjoint: if $\mathbf{a}=\mathbf{E}\mathbf{b}=(\mathbf{I}-\mathbf{E}\mathbf{E}^-)\mathbf{F}\mathbf{c}$, then $(\mathbf{I}-\mathbf{E}\mathbf{E}^-)\mathbf{a}=\mathbf{0}=(\mathbf{I}-\mathbf{E}\mathbf{E}^-)\mathbf{F}\mathbf{c}=\mathbf{a}$, as $\mathbf{I}-\mathbf{E}\mathbf{E}^-$ 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. ∎

THEOREM 4.1. *Let the matrix*

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

have elements over an arbitrary field, and suppose that both $\mathbf{A}$ *and* $\mathbf{E}$ *are square. If either*

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

or

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

then

$$|\mathbf{A}|=|\mathbf{E}|\cdot|\mathbf{H}-\mathbf{G}\mathbf{E}^-\mathbf{F}| \tag{4.14}$$

for every g-inverse $\mathbf{E}^-$.

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

$$\mathbf{E}\mathbf{E}^{-}\mathbf{F}=\mathbf{F} \tag{4.15}$$

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

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

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

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

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

Corollary 4.1. *If* $\mathbf{A}$ *and* $\mathbf{H}$ *in* (4.11) *are both square and if either*

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

or

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

then

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

for every g-inverse $\mathbf{H}^{-}$.

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

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

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

Carlson, Haynsworth, and Markham [15] refer to the result in Theorem 4.2 below as Sylvester's determinant formula. We notice that this result parallels that of Jacobi [42], our Theorem 2.9.

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

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix}, \tag{4.21}$$

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

$$d_{ij}=\begin{vmatrix}\mathbf{E} & \mathbf{f}_j\\ \mathbf{g}_i' & h_{ij}\end{vmatrix}, \qquad i,j=1,2,\dots,n-m, \tag{4.22}$$

and $\mathbf{f}_j,\mathbf{g}_i'$ *denote, respectively, the* j*th column of* $\mathbf{F}$ *and the* i*th row of* $\mathbf{G}$, *and* $\mathbf{H}=\{h_{ij}\}$. *If either*

$$\mathrm{r}(\mathbf{E},\mathbf{F})=\mathrm{r}(\mathbf{E}) \tag{4.23}$$

or

$$\mathrm{r}\begin{pmatrix}\mathbf{E}\\ \mathbf{G}\end{pmatrix}=\mathrm{r}(\mathbf{E}), \tag{4.24}$$

then

$$\mathbf{D}=|\mathbf{E}|\cdot(\mathbf{A}/\mathbf{E}) \tag{4.25}$$

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

$$|\mathbf{D}|=|\mathbf{E}|^{n-m-1}|\mathbf{A}|. \tag{4.26}$$

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

4.2. Rank

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

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

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

We may, however, following Meyer [50] and Marsaglia and Styan [47], establish the following:

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

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

Three different choices of $\mathbf{E}^{-}$ may be made.

Proof. We note that

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

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

$$\mathbf{X}=(\mathbf{I}-\mathbf{E}\mathbf{E}^{\sim})\mathbf{F}, \qquad \mathbf{Y}=\mathbf{G}(\mathbf{I}-\mathbf{E}^{-}\mathbf{E}) \tag{4.29}$$

and

$$\mathbf{K}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-}\mathbf{F}-\mathbf{Y}\mathbf{E}^{\sim}\mathbf{F}. \tag{4.30}$$

Then

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

since the columns (rows) of $\mathbf{E}$ are linearly independent of the columns of $\mathbf{X}$ (rows of $\mathbf{Y}$). Since

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

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

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

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

where $\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^-\mathbf{F}$, and hence

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

is invariant under choice of $\mathbf{E}^-$. This completes the proof. ∎

Marsaglia and Styan [47, (8.5)] obtained Theorem 4.3 but required that the $\mathbf{E}^-$ in the lower right corner of

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

must be the $\mathbf{E}^-$ either in the lower left or in the upper right corner.

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

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

Three different choices of $\mathbf{H}^-$ *may be made.*

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

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

where

$$\mathbf{U} = -(\mathbf{I}-\mathbf{E}\mathbf{E}^{-})\mathbf{F}\mathbf{S}^{-}\mathbf{G}(\mathbf{I}-\mathbf{E}^{-}\mathbf{E}), \tag{4.39a}$$

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

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

We now use

LEMMA 4.2 (Marsaglia and Styan [47, (8.3)]). *For matrices over an arbitrary field,*

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

Any choices of $\mathbf{X}^{-}$ *and* $\mathbf{Y}^{-}$ *may be made.*

Proof. Using Lemma 4.1 yields

$$\mathrm{r}\left(\begin{array}{cc} \mathbf{0} & \mathbf{X} \\ \hdashline \mathbf{Y} & \mathbf{S} \end{array}\right) = \mathrm{r}(\mathbf{X}) + \mathrm{r}(\mathbf{Y},\mathbf{S})\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}-\mathbf{X}^{-}\mathbf{X} \end{pmatrix} \tag{4.41}$$

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

Applying (4.6) gives (4.40). ■

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

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

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

where

$$\mathbf{Z} = (\mathbf{I}-\mathbf{V}\mathbf{V}^{-})\mathbf{U}(\mathbf{I}-\mathbf{W}^{-}\mathbf{W}), \tag{4.44}$$

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

Meyer [50, Corollary 4.1] proved that

$$\mathrm{r}\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \leqslant \mathrm{r}(\mathbf{E}) + \mathrm{r}(\mathbf{S}) + \mathrm{r}(\mathbf{F}) + \mathrm{r}(\mathbf{G}). \tag{4.45}$$

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

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

and

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

$$\leqslant r(U,V) + r(G) \tag{4.48}$$

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

$$\leqslant r(F) + r(G), \tag{4.50}$$

which proves the inequality (4.45).

Meyer [50, Theorem 4.1] also proved:

THEOREM 4.5. *For matrices over an arbitrary field*

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

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). ■

Marsaglia and Styan [47, (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

$$S = H - GE^{\sim}F \tag{4.52}$$

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

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

COROLLARY 4.3 (Marsaglia and Styan [47, p. 291, Corollary 19.1]). *For matrices over an arbitrary field, rank is additive on the Schur complement:*

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

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

$$(\mathbf{I} - \mathbf{E}\mathbf{E}^{-})\mathbf{F}(\mathbf{I} - \mathbf{S}^{-}\mathbf{S}) = \mathbf{0} \tag{4.55a}$$

$$(\mathbf{I} - \mathbf{S}\mathbf{S}^{-})\mathbf{G}(\mathbf{I} - \mathbf{E}^{-}\mathbf{E}) = \mathbf{0} \tag{4.55b}$$

$$(\mathbf{I} - \mathbf{E}\mathbf{E}^{-})\mathbf{F}\mathbf{S}^{-}\mathbf{G}(\mathbf{I} - \mathbf{E}^{-}\mathbf{E}) = \mathbf{0}, \tag{4.55c}$$

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

Proof. Immediate from Theorem 4.4. ■

Corollary 4.3 was proved by Carlson, Haynsworth, and Markham [15] with $\mathbf{E}^{\sim} = \mathbf{E}^{+}$, the Moore-Penrose g-inverse. They assert that their proof can be used to cover the case where $\mathbf{E}^{\sim}$ is a reflexive g-inverse. See also Carlson [14, Theorem A].

We note that if the conditions in Corollary 4.3 hold, then $|\mathbf{E}| = 0$ implies $|\mathbf{A}| = 0$; 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

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \tag{4.56}$$

and

$$\mathbf{B} = \begin{pmatrix} \mathbf{E}^{-} + \mathbf{E}^{-}\mathbf{F}\mathbf{S}^{-}\mathbf{G}\mathbf{E}^{-} & -\mathbf{E}^{-}\mathbf{F}\mathbf{S}^{-} \\ -\mathbf{S}^{-}\mathbf{G}\mathbf{E}^{-} & \mathbf{S}^{-} \end{pmatrix}, \tag{4.57}$$

where

$$\mathbf{S} = \mathbf{H} - \mathbf{G}\mathbf{E}^{-}\mathbf{F}. \tag{4.58}$$

Rohde [60] showed that if $\mathbf{A}$ is real and nonnegative definite, then indeed $\mathbf{B}$ is a g-inverse of $\mathbf{A}$. This result was extended by Pringle and Rayner [56], who assumed that $\mathbf{A}$ has the structure

$$\mathbf{A}=\begin{pmatrix}\mathbf{K}'\mathbf{K} & \mathbf{K}'\mathbf{L}\\ \mathbf{L}'\mathbf{K} & \mathbf{0}\end{pmatrix}, \tag{4.59}$$

and later by Marsaglia and Styan [48, Corollary 1] for

$$\mathbf{A}=\begin{pmatrix}\mathbf{K}'\mathbf{K} & \mathbf{K}'\mathbf{L}\\ \mathbf{M}'\mathbf{K} & \mathbf{N}\end{pmatrix}; \tag{4.60}$$

cf. Corollary 4.6 below. More generally, Bhimasankaram [7] and Burns, Carlson, Haynsworth, and Markham [13] showed that $\mathbf{B}=\mathbf{A}^-$ if and only if the conditions (4.55) hold. Applying Corollary 4.3, we then get

THEOREM 4.6 (Marsaglia and Styan [48 p. 439]). *Suppose that the matrix* $\mathbf{A}$ *defined by* (4.56) *has elements over an arbitrary field, and that* $\mathbf{E}^\sim$ *is a particular g-inverse of* $\mathbf{E}$. *Let the Schur complement* $\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^\sim\mathbf{F}$, *and let*

$$\mathbf{B}=\begin{pmatrix}\mathbf{E}^\sim+\mathbf{E}^\sim\mathbf{F}\mathbf{S}^-\mathbf{G}\mathbf{E}^\sim & -\mathbf{E}^\sim\mathbf{F}\mathbf{S}^-\\ -\mathbf{S}^-\mathbf{G}\mathbf{E}^\sim & \mathbf{S}^-\end{pmatrix}. \tag{4.61}$$

Then:

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

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

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

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

and

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

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

(ii), (iii): These proofs are straightforward but more lengthy: we refer the reader to [48, pp. 438–439] for details. ■

Bhimasankaram [7] and Burns, Carlson, Haynsworth, and Markham [13] proved that the matrix $\mathbf{B}$ defined by (4.61) is a g-inverse of $\mathbf{A}$ if and only if (4.55) holds.

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

$$\mathbf{C}=\begin{pmatrix} \mathbf{T}^{-} & -\mathbf{T}^{-}\mathbf{F}\mathbf{H}^{\sim} \\ -\mathbf{H}^{\sim}\mathbf{G}\mathbf{T}^{-} & \mathbf{H}^{\sim}+\mathbf{H}^{\sim}\mathbf{G}\mathbf{T}^{-}\mathbf{F}\mathbf{H}^{\sim} \end{pmatrix} \tag{4.64}$$

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

$$\mathrm{r}(\mathbf{A})=\mathrm{r}(\mathbf{H})+\mathrm{r}(\mathbf{E}-\mathbf{F}\mathbf{H}^{\sim}\mathbf{G}), \tag{4.65}$$

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

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

and

$$\mathrm{r}\begin{pmatrix} \mathbf{E} \\ \mathbf{G} \end{pmatrix}=\mathrm{r}(\mathbf{E},\mathbf{F})=\mathrm{r}(\mathbf{T})=\mathrm{r}(\mathbf{E}-\mathbf{F}\mathbf{H}^{\sim}\mathbf{G}). \tag{4.67}$$

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

COROLLARY 4.4. *Let the complex matrix*

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

and let

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

Then

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

the Moore-Penrose g-inverse of $\mathbf{A}$, *if*

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

and

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

Burns, Carlson, Haynsworth, and Markham [13] noted that the Moore-Penrose g-inverse of $\mathbf{A}$ is given by (4.70) if (4.61) and (4.64) equal $\mathbf{A}^+$, since the Moore-Penrose g-inverse is unique. Moreover, Theorem 4.6 yields:

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

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

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

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

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

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

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

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

Burns, Carlson, Haynsworth, and Markham [13] proved that, if

$$\mathbf{G}(\mathbf{I}-\mathbf{E}^+\mathbf{E})=\mathbf{0}, \qquad (\mathbf{I}-\mathbf{E}\mathbf{E}^+)\mathbf{F}=\mathbf{0}, \tag{4.77a}$$

$$(\mathbf{I}-\mathbf{S}\mathbf{S}^+)\mathbf{G}=\mathbf{0}, \qquad \mathbf{F}(\mathbf{I}-\mathbf{S}^+\mathbf{S})=\mathbf{0}, \tag{4.77b}$$

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

COROLLARY 4.6 (Marsaglia and Styan [48]). *Suppose that the real matrix* $\mathbf{A}$ *is defined by*

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

Then rank is additive on the Schur complement:

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

where

$$\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-}\mathbf{F}=\mathbf{N}-\mathbf{M}'\mathbf{K}\mathbf{K}^{+}\mathbf{L} \tag{4.80}$$

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

$$\mathbf{B}=\begin{pmatrix}\mathbf{E}^{-}+\mathbf{E}^{-}\mathbf{F}\mathbf{S}^{-}\mathbf{G}\mathbf{E}^{-} & -\mathbf{E}^{-}\mathbf{F}\mathbf{S}^{-}\\ -\mathbf{S}^{-}\mathbf{G}\mathbf{E}^{-} & \mathbf{S}^{-}\end{pmatrix} \tag{4.81}$$

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

$$\mathrm{r}\begin{pmatrix}\mathbf{F}\\ \mathbf{H}\end{pmatrix}=\mathrm{r}(\mathbf{G},\mathbf{H})=\mathrm{r}(\mathbf{H}-\mathbf{G}\mathbf{E}^{+}\mathbf{F})=\mathrm{r}(\mathbf{S}). \tag{4.82}$$

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

Rohde [60] obtained (4.81) for a g-inverse of $\mathbf{A}$, where $\mathbf{A}$ is defined by (4.78) with $\mathbf{M}=\mathbf{L}$ and $\mathbf{N}=\mathbf{L}'\mathbf{L}$. Pringle and Rayner [56] also established that $\mathbf{B}$, given by (4.81), is a g-inverse of $\mathbf{A}$, where $\mathbf{A}$ is defined by (4.78), with $\mathbf{M}=\mathbf{L}$ and $\mathbf{N}=\mathbf{0}$. The following corollary gives a different approach to Rohde's result. (See also [61].)

COROLLARY 4.7 (Marsaglia and Styan [48]). *Suppose that the real nonnegative definite matrix* $\mathbf{A}$ *is defined by*

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{F}' & \mathbf{H}\end{pmatrix}=\begin{pmatrix}\mathbf{K}'\mathbf{K} & \mathbf{K}'\mathbf{L}\\ \mathbf{L}'\mathbf{K} & \mathbf{L}'\mathbf{L}\end{pmatrix}. \tag{4.83}$$

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

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

$$\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}' & \mathbf{H} \end{pmatrix}^+ = \begin{pmatrix} \mathbf{E}^+ + \mathbf{E}^+\mathbf{F}\mathbf{S}^+\mathbf{F}'\mathbf{E}^+ & -\mathbf{E}^+\mathbf{F}\mathbf{S}^+ \\ -\mathbf{S}^+\mathbf{F}'\mathbf{E}^+ & \mathbf{S}^+ \end{pmatrix}, \tag{4.85}$$

$$\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}' & \mathbf{H} \end{pmatrix}^+ = \begin{pmatrix} \mathbf{T}^+ & -\mathbf{T}^+\mathbf{F}\mathbf{H}^+ \\ -\mathbf{H}^+\mathbf{F}'\mathbf{T}^+ & \mathbf{H}^+ + \mathbf{H}^+\mathbf{F}'\mathbf{T}^+\mathbf{F}\mathbf{H}^+ \end{pmatrix}, \tag{4.86}$$

where the Schur complements of $\mathbf{E}$, *and of* $\mathbf{H}$, *in* $\mathbf{A}$,

$$\mathbf{S} = \mathbf{H} - \mathbf{F}'\mathbf{E}^-\mathbf{F} = \mathbf{L}'(\mathbf{I} - \mathbf{K}\mathbf{K}^+)\mathbf{L} \tag{4.87}$$

and

$$\mathbf{T} = \mathbf{E} - \mathbf{F}\mathbf{H}^-\mathbf{F}' = \mathbf{K}(\mathbf{I} - \mathbf{L}\mathbf{L}^+)\mathbf{K}', \tag{4.88}$$

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

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

$$\mathrm{r}\begin{pmatrix} \mathbf{F} \\ \mathbf{H} \end{pmatrix} = \mathrm{r}(\mathbf{H}), \tag{4.89}$$

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

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

and so (4.82) holds, which implies (4.85). By the reverse argument, (4.85) implies (4.84). The alternative arrangement in (4.86) follows from the "symmetry" in (4.84) with respect to $\mathbf{E}$ and $\mathbf{H}$. ■

4.4. *Inertia*

Consider the real symmetric matrix

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

where $\mathbf{E}$ is singular. Then inertia need not be additive on the generalized Schur complement, in contrast to the case where $\mathbf{E}$ 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*:

$$\mathrm{r}(\mathbf{A}+\mathbf{B})=\mathrm{r}(\mathbf{A})+\mathrm{r}(\mathbf{B}). \tag{4.92}$$

Then

$$\mathrm{In}(\mathbf{A}+\mathbf{B})=\mathrm{In}\,\mathbf{A}+\mathrm{In}\,\mathbf{B}, \tag{4.93}$$

where In *denotes inertia.*

Following Carlson, Haynsworth, and Markham [15, p. 172], by (4.93) we mean that $\pi(\mathbf{A}+\mathbf{B})=\pi(\mathbf{A})+\pi(\mathbf{B})$ and $\nu(\mathbf{A}+\mathbf{B})=\nu(\mathbf{A})+\nu(\mathbf{B})$, where $\pi(\cdot)$ and $\nu(\cdot)$ denote, respectively, the number of positive and negative characteristic roots.

Proof. Following (3.1), let

$$\mathrm{In}\,\mathbf{A}=(\pi_a,\nu_a,\delta_a),\qquad \mathrm{r}(\mathbf{A})=\pi_a+\nu_a=r_a, \tag{4.93a}$$

$$\mathrm{In}\,\mathbf{B}=(\pi_b,\nu_b,\delta_b),\qquad \mathrm{r}(\mathbf{B})=\pi_b+\nu_b=r_b. \tag{4.93b}$$

Since **A** and **B** are both $n\times n$ real symmetric matrices, there exist real nonsingular matrices **S** and **T** such that

$$\mathbf{A}=\mathbf{S}\mathbf{D}_a\mathbf{S}'=(\mathbf{S}_1,\mathbf{S}_2,\mathbf{S}_3)\begin{pmatrix}\mathbf{I}_{\pi_a} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{\nu_a} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{\delta_a}\end{pmatrix}\begin{pmatrix}\mathbf{S}_1'\\ \mathbf{S}_2'\\ \mathbf{S}_3'\end{pmatrix} \tag{4.94a}$$

$$=\mathbf{S}_1\mathbf{S}_1'-\mathbf{S}_2\mathbf{S}_2' \tag{4.94b}$$

and

$$\mathbf{B}=\mathbf{T}\mathbf{D}_b\mathbf{T}'=(\mathbf{T}_1,\mathbf{T}_2,\mathbf{T}_3)\begin{pmatrix}\mathbf{I}_{\pi_b} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{\nu_b} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{\delta_b}\end{pmatrix}\begin{pmatrix}\mathbf{T}_1'\\ \mathbf{T}_2'\\ \mathbf{T}_3'\end{pmatrix} \tag{4.95a}$$

$$=\mathbf{T}_1\mathbf{T}_1'-\mathbf{T}_2\mathbf{T}_2'. \tag{4.95b}$$

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

$$\mathbf{A}+\mathbf{B}=\mathbf{S}_1\mathbf{S}_1'-\mathbf{S}_2\mathbf{S}_2'+\mathbf{T}_1\mathbf{T}_1'-\mathbf{T}_2\mathbf{T}_2' \tag{4.96a}$$

$$=(\mathbf{S}_1,\mathbf{S}_2,\mathbf{T}_1,\mathbf{T}_2)\begin{pmatrix}\mathbf{I}_{\pi_a} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{\nu_a} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & -\mathbf{I}_{\pi_b} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I}_{\nu_b}\end{pmatrix}\begin{pmatrix}\mathbf{S}_1'\\ \mathbf{S}_2'\\ \mathbf{T}_1'\\ \mathbf{T}_2'\end{pmatrix}. \tag{4.96b}$$

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

$$\begin{aligned}\mathrm{r}(\mathbf{A}+\mathbf{B})&=r_a+r_b\leqslant \mathrm{r}(\mathbf{S}_1,\mathbf{S}_2,\mathbf{T}_1,\mathbf{T}_2)\\ &\leqslant \pi_a+\nu_a+\pi_b+\nu_b=r_a+r_b.\end{aligned} \tag{4.97}$$

Hence there exists a matrix $\mathbf{U}$, say, $n\times(n-r_a-r_b)$, such that $\mathbf{V}=(\mathbf{S}_1,\mathbf{T}_1,\mathbf{S}_2,\mathbf{T}_2,\mathbf{U})$ is nonsingular, and writing

$$\mathbf{A}+\mathbf{B}=(\mathbf{S}_1,\mathbf{T}_1,\mathbf{S}_2,\mathbf{T}_2,\mathbf{U})\begin{pmatrix}\mathbf{I}_{\pi_a+\pi_b} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{\nu_a+\nu_b} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{S}_1'\\ \mathbf{T}_1'\\ \mathbf{S}_2'\\ \mathbf{T}_2'\\ \mathbf{U}'\end{pmatrix} \tag{4.98}$$

completes the proof. ∎

THEOREM 4.7 (Carlson, Haynsworth, and Markham [15]). *Consider the real symmetric matrix*

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

Let $\mathbf{E}_r^-$ *be a symmetric reflexive g-inverse of* $\mathbf{E}$, *and let the generalized Schur complement*

$$\mathbf{S}=\mathbf{H}-\mathbf{F}'\mathbf{E}_r^-\mathbf{F}=(\mathbf{A}/\mathbf{E}). \tag{4.100}$$

Then

$$\mathrm{In}\,\mathbf{A}=\mathrm{In}\,\mathbf{E}+\mathrm{In}(\mathbf{A}/\mathbf{E})+\mathrm{In}\begin{pmatrix}-\mathbf{X}\mathbf{S}_r^-\mathbf{X}' & \mathbf{V}\\ \mathbf{V}' & \mathbf{0}\end{pmatrix}, \tag{4.101}$$

where

$$\mathbf{X}=(\mathbf{I}-\mathbf{E}\mathbf{E}_r^-)\mathbf{F} \quad \textit{and} \quad \mathbf{V}=\mathbf{X}(\mathbf{I}-\mathbf{S}_r^-\mathbf{S}), \tag{4.102}$$

and $\mathbf{S}_r^-$ *is a symmetric reflexive g-inverse of* $\mathbf{S}$. *Furthermore, if*

$$\mathrm{r}(\mathbf{E})=\mathrm{r}(\mathbf{E},\mathbf{F}), \tag{4.103}$$

then $\mathbf{S}$ *is unique and*

$$\mathrm{In}\,\mathbf{A}=\mathrm{In}\,\mathbf{E}+\mathrm{In}(\mathbf{A}/\mathbf{E}). \tag{4.104}$$

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

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{F}'\mathbf{E}_r^- & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}' & \mathbf{H} \end{pmatrix}\begin{pmatrix} \mathbf{I} & -\mathbf{E}_r^-\mathbf{F} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}=\begin{pmatrix} \mathbf{E} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix}, \tag{4.105}$$

since $\mathbf{E}_r^-$ is a symmetric reflexive g-inverse of $\mathbf{E}$. Then by Sylvester's law of inertia [cf. (3.2)] we obtain

$$\mathrm{In}\,\mathbf{A}=\mathrm{In}\begin{pmatrix} \mathbf{E} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix}. \tag{4.106}$$

Using Theorem 4.3 yields

$$\mathrm{r}\begin{pmatrix} \mathbf{E} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix}=\mathrm{r}\begin{pmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}+\mathrm{r}\begin{pmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix}, \tag{4.107}$$

which, using Lemma 4.3, yields

$$\mathrm{In}\,\mathbf{A}=\mathrm{In}\,\mathbf{E}+\mathrm{In}\begin{pmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix}. \tag{4.108}$$

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

$$\begin{pmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix}=\begin{pmatrix} \mathbf{I} & \mathbf{X}\mathbf{S}_r^- \\ \mathbf{0} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V}' & \mathbf{S} \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{S}_r^-\mathbf{X}' & \mathbf{I} \end{pmatrix}, \tag{4.109}$$

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

$$\operatorname{In}\begin{pmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{X}' & \mathbf{S} \end{pmatrix} = \operatorname{In}\begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V}' & \mathbf{S} \end{pmatrix} = \operatorname{In}\mathbf{S} + \operatorname{In}\begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V}' & \mathbf{0} \end{pmatrix}, \tag{4.110}$$

since

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

using Corollary 4.2 and Lemma 4.3. Thus (4.101) follows at once. If (4.103) holds, then Lemma 4.1 shows $\mathbf{X}=\mathbf{0}$, and so $\mathbf{V}=\mathbf{0}$ and (4.104) follows. ■

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

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

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

THEOREM 4.8 (Carson, Haynsworth, and Markham [15]). *Consider the matrix*

$$\mathbf{A}=\left(\begin{array}{c:c} \mathbf{E} & \mathbf{F} \\ \hdashline \mathbf{G} & \mathbf{H} \end{array}\right)=\left[\begin{array}{cc:c} \mathbf{K} & \mathbf{L} & \mathbf{F}_1 \\ \mathbf{M} & \mathbf{N} & \mathbf{F}_2 \\ \hdashline \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{H} \end{array}\right]. \tag{4.112}$$

If

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

and

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

then

$$r(\mathbf{K},\mathbf{L},\mathbf{F}_1)=r(\mathbf{K})=r\begin{bmatrix}\mathbf{K}\\ \mathbf{M}\\ \mathbf{G}_1\end{bmatrix}\tag{4.115}$$

and the generalized Schur complements $(\mathbf{A}/\mathbf{E})$, $(\mathbf{E}/\mathbf{K})$, *and* $(\mathbf{A}/\mathbf{K})$ *are uniquely determined, and*

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

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

$$\mathbf{F}=\mathbf{E}\mathbf{E}^-\mathbf{F}\quad\text{and}\quad\mathbf{G}=\mathbf{G}\mathbf{E}^-\mathbf{E},\tag{4.117}$$

so that

$$\mathbf{F}_1=(\mathbf{K},\mathbf{L})\mathbf{E}^-\mathbf{F}\quad\text{and}\quad\mathbf{G}_1=\mathbf{G}\mathbf{E}^-\begin{pmatrix}\mathbf{K}\\ \mathbf{M}\end{pmatrix}.\tag{4.118}$$

Thus

$$r(\mathbf{K},\mathbf{L},\mathbf{F}_1)=r(\mathbf{K},\mathbf{L})\quad\text{and}\quad r\begin{bmatrix}\mathbf{K}\\ \mathbf{M}\\ \mathbf{G}_1\end{bmatrix}=r\begin{pmatrix}\mathbf{K}\\ \mathbf{M}\end{pmatrix}.\tag{4.119}$$

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

$$\mathbf{A}=\begin{bmatrix}\mathbf{K} & \mathbf{K}\mathbf{L}_0 & \mathbf{K}\mathbf{F}_0\\ \mathbf{M}_0\mathbf{K} & \mathbf{N} & \mathbf{F}_2\\ \mathbf{G}_0\mathbf{K} & \mathbf{G}_2 & \mathbf{H}\end{bmatrix}\tag{4.120}$$

for some matrices $\mathbf{L}_0$, $\mathbf{F}_0$, $\mathbf{M}_0$, and $\mathbf{G}_0$. To prove (4.116) we notice that

$$(\mathbf{A}/\mathbf{K})=\begin{pmatrix}(\mathbf{E}/\mathbf{K}) & \mathbf{F}_2-\mathbf{M}_0\mathbf{K}\mathbf{F}_0\\ \mathbf{G}_2-\mathbf{G}_0\mathbf{K}\mathbf{L}_0 & \mathbf{H}-\mathbf{G}_0\mathbf{K}\mathbf{F}_0\end{pmatrix},\tag{4.121}$$

so that

$$((\mathbf{A}/\mathbf{K})/(\mathbf{E}/\mathbf{K}))=\mathbf{H}-\mathbf{G}_0\mathbf{K}\mathbf{F}_0-(\mathbf{G}_2-\mathbf{G}_0\mathbf{K}\mathbf{L}_0)(\mathbf{E}/\mathbf{K})^-(\mathbf{F}_2-\mathbf{M}_0\mathbf{K}\mathbf{F}_0),\tag{4.122}$$

while

$$(\mathbf{A}/\mathbf{E})=\mathbf{H}-(\mathbf{G}_0\mathbf{K},\mathbf{G}_2)\begin{pmatrix}\mathbf{K} & \mathbf{K}\mathbf{L}_0\\ \mathbf{M}_0\mathbf{K} & \mathbf{N}\end{pmatrix}^{-}\begin{pmatrix}\mathbf{K}\mathbf{F}_0\\ \mathbf{F}_2\end{pmatrix}. \tag{4.123}$$

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

$$\begin{pmatrix}\mathbf{K} & \mathbf{K}\mathbf{L}_0\\ \mathbf{M}_0\mathbf{K} & \mathbf{N}\end{pmatrix}^{-}=\begin{pmatrix}\mathbf{K}^{-} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}+\begin{pmatrix}\mathbf{K}^{-}\mathbf{K}\mathbf{L}_0\\ -\mathbf{I}\end{pmatrix}(\mathbf{E}/\mathbf{K})^{-}(\mathbf{M}_0\mathbf{K}\mathbf{K}^{-},-\mathbf{I}). \tag{4.124}$$

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

Carlson, Haynsworth, and Markham [15] also extended Theorem 3.4 using generalized Schur complements. Let

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{F}' & \mathbf{H}\end{pmatrix} \tag{4.125}$$

and

$$\mathbf{B}=\begin{pmatrix}\mathbf{K} & \mathbf{L}\\ \mathbf{L}' & \mathbf{N}\end{pmatrix} \tag{4.126}$$

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

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

The generalized Schur complements

$$(\mathbf{A}/\mathbf{E})=\mathbf{H}-\mathbf{F}'\mathbf{E}^{-}\mathbf{F}, \tag{4.128a}$$

$$(\mathbf{B}/\mathbf{K})=\mathbf{N}-\mathbf{L}'\mathbf{K}^{-}\mathbf{L} \tag{4.128b}$$

are, therefore, uniquely defined, as is

$$((\mathbf{A}+\mathbf{B})/(\mathbf{E}+\mathbf{K}))=\mathbf{H}+\mathbf{N}-(\mathbf{F}+\mathbf{L})'(\mathbf{E}+\mathbf{K})^{-}(\mathbf{F}+\mathbf{L}), \tag{4.129}$$

since $\mathbf{A}+\mathbf{B}$ is nonnegative definite (cf. remarks before Corollary 4.1).

THEOREM 4.9 (Carlson, Haynsworth, and Markham [15]). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be defined as in* (4.125) *and* (4.126). *Then*

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

is nonnegative definite.

Proof. Consider the matrices

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

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

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

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

Carlson, Haynsworth, and Markham [15] also extended Theorem 3.5 by allowing the principal submatrices $\mathbf{E}_i$ and $\mathbf{K}_i$ to be nonnegative definite. The inequality (3.47), however, is only meaningful when the $\mathbf{E}_i$ and $\mathbf{K}_i$ are all nonsingular. If we substitute $|\mathbf{A}|/|\mathbf{E}_i|=|(\mathbf{A}/\mathbf{E}_i)|$ and $|\mathbf{B}|/|\mathbf{K}_i|=|(\mathbf{B}/\mathbf{K}_i)|$, then we obtain:

THEOREM 4.10 (Carlson, Haynsworth, and Markham [15]). *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *both be* $n\times n$ *nonnegative definite matrices. Suppose further that* $\mathbf{E}_i$ *and* $\mathbf{K}_i$, $i=1,\dots,n$, *are the* $i\times i$ *principal submatrices in the upper left corners of the matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *respectively. Then*

$$|\mathbf{A}+\mathbf{B}|\geqslant|\mathbf{A}|+|\mathbf{B}|+\sum_{i=1}^{n-1}\left[|(\mathbf{A}/\mathbf{E}_i)|\cdot|\mathbf{K}_i|+|(\mathbf{B}/\mathbf{K}_i)|\cdot|\mathbf{E}_i|\right]. \tag{4.133}$$

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

$$|\mathbf{A}+\mathbf{B}|=|\mathbf{E}_1+\mathbf{K}_1|\cdot|((\mathbf{A}+\mathbf{B})/(\mathbf{E}_1+\mathbf{K}_1))|; \tag{4.134}$$

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

$$|\mathbf{E}_1+\mathbf{K}_1| \geqslant |\mathbf{E}_1|+|\mathbf{K}_1|, \tag{4.135}$$

while

$$|((\mathbf{A}+\mathbf{B})/(\mathbf{E}_1+\mathbf{K}_1))| \geqslant |(\mathbf{A}/\mathbf{E}_1)|+|(\mathbf{B}/\mathbf{K}_1)| \tag{4.136}$$

follows from (4.130). Hence

$$\begin{aligned}|\mathbf{A}+\mathbf{B}| &\geqslant (|\mathbf{E}_1|+|\mathbf{K}_1|)(|(\mathbf{A}/\mathbf{E}_1)|+|(\mathbf{B}/\mathbf{K}_1)|)\\ &= |\mathbf{E}_1|\cdot|(\mathbf{A}/\mathbf{E}_1)|+|\mathbf{K}_1|\cdot|(\mathbf{B}/\mathbf{K}_1)|\\ &\quad +|(\mathbf{A}/\mathbf{E}_1)|\cdot|\mathbf{K}_1|+|(\mathbf{B}/\mathbf{K}_1)|\cdot|\mathbf{E}_1|\\ &= |\mathbf{A}|+|\mathbf{B}|+|(\mathbf{A}/\mathbf{E}_1)|\cdot|\mathbf{K}_1|+|(\mathbf{B}/\mathbf{K}_1)|\cdot|\mathbf{E}_1|. \end{aligned} \tag{4.137}$$

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

Now assume that (4.133) holds for $\mathbf{A}$ and $\mathbf{B}$ $n\times n$. If $\mathbf{A}_1$ and $\mathbf{B}_1$ are $(n+1)\times(n+1)$ nonnegative definite matrices, and $\mathbf{A}=\mathbf{E}_n$ and $\mathbf{B}=\mathbf{K}_n$ are $n\times n$ submatrices of $\mathbf{A}_1$ and $\mathbf{B}_1$, respectively, then

$$|\mathbf{A}_1+\mathbf{B}_1| = |\mathbf{E}_n+\mathbf{K}_n|\cdot|((\mathbf{A}_1+\mathbf{B}_1)/(\mathbf{E}_n+\mathbf{K}_n))|. \tag{4.138}$$

But, by the inductive assumption,

$$|\mathbf{E}_n+\mathbf{K}_n| \geqslant \sum_{i=1}^{n}\left[|(\mathbf{E}_n/\mathbf{E}_i)|\cdot|\mathbf{K}_i|+|(\mathbf{K}_n/\mathbf{K}_i)|\cdot|\mathbf{E}_i|\right] \tag{4.139}$$

and by (4.130),

$$|((\mathbf{A}_1+\mathbf{B}_1)/(\mathbf{E}_n+\mathbf{K}_n))| \geqslant |(\mathbf{A}_1/\mathbf{E}_n)|+|(\mathbf{B}_1/\mathbf{K}_n)|. \tag{4.140}$$

Hence,

$$
\begin{aligned}
|\mathbf{A}_1+\mathbf{B}_1| \geqslant & \left\{ \sum_{i=1}^{n} \left[|(\mathbf{E}_n/\mathbf{E}_i)|\cdot|\mathbf{K}_i| + |(\mathbf{K}_n/\mathbf{K}_i)|\cdot|\mathbf{E}_i| \right] \right\} \\
& \times \left\{ |(\mathbf{A}_1/\mathbf{E}_n)| + |(\mathbf{B}_1/\mathbf{K}_n)| \right\} \\
= & \sum_{i=1}^{n} |(\mathbf{A}_1/\mathbf{E}_n)|\cdot|(\mathbf{E}_n/\mathbf{E}_i)|\cdot|\mathbf{K}_i| \\
& + \sum_{i=1}^{n} |(\mathbf{A}_1/\mathbf{E}_n)|\cdot|(\mathbf{K}_n/\mathbf{K}_i)|\cdot|\mathbf{E}_i| \\
& + \sum_{i=1}^{n} |(\mathbf{B}_1/\mathbf{K}_n)|\cdot|(\mathbf{E}_n/\mathbf{E}_i)|\cdot|\mathbf{K}_i| + \sum_{i=1}^{n} |(\mathbf{B}_1/\mathbf{K}_n)|\cdot|(\mathbf{K}_n/\mathbf{K}_i)|\cdot|\mathbf{E}_i| \\
\geqslant & \sum_{i=1}^{n} |(\mathbf{A}_1/\mathbf{E}_n)|\cdot|(\mathbf{E}_n/\mathbf{E}_i)|\cdot|\mathbf{K}_i| + |\mathbf{A}_1| \\
& + |\mathbf{B}_1| + \sum_{i=1}^{n} |(\mathbf{B}_1/\mathbf{K}_n)|\cdot|(\mathbf{K}_n/\mathbf{K}_i)|\cdot|\mathbf{E}_i| \\
= & \sum_{i=1}^{n+1} \left[|(\mathbf{A}_1/\mathbf{E}_i)|\cdot|\mathbf{K}_i| + |(\mathbf{B}_1/\mathbf{K}_i)|\cdot|\mathbf{E}_i| \right]. \qquad (4.141)
\end{aligned}
$$

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

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)]

$$
\begin{vmatrix} \mathbf{I}_m & \mathbf{F} \\ \mathbf{G} & \mathbf{I}_n \end{vmatrix} = |\mathbf{I}_m - \mathbf{FG}| = |\mathbf{I}_n - \mathbf{GF}|. \qquad (4.142)
$$

Using (2.4) and (2.6), we similarly obtain

$$\begin{vmatrix} \lambda\mathbf{I}_m & \mathbf{F} \\ \mathbf{G} & \mathbf{I}_n \end{vmatrix} = |\lambda\mathbf{I}_m - \mathbf{FG}| = |\lambda\mathbf{I}_m|\cdot\left|\mathbf{I}_n - \frac{1}{\lambda}\mathbf{GF}\right|, \tag{4.143}$$

and so

$$\lambda^n|\lambda\mathbf{I}_m - \mathbf{FG}| = \lambda^m|\lambda\mathbf{I}_n - \mathbf{GF}|, \tag{4.144}$$

which shows that **FG** and **GF** have the same nonzero characteristic roots; see, e.g., [52, p. 200].

Furthermore if we replace $\mathbf{I}_n$ in the lower right corner of (4.143) by $\lambda\mathbf{I}_n$, $\mathbf{F}$ by $-\mathbf{F}$, and $\mathbf{G}$ by $-\mathbf{F}'$, then

$$\begin{vmatrix} \lambda\mathbf{I}_m & -\mathbf{F} \\ -\mathbf{F}' & \lambda\mathbf{I}_n \end{vmatrix} = |\lambda\mathbf{I}_m|\cdot|\lambda\mathbf{I}_n - \mathbf{F}'\mathbf{F}/\lambda| = \lambda^{m-n}|\lambda^2\mathbf{I}_n - \mathbf{F}'\mathbf{F}|, \tag{4.145}$$

and so the nonzero characteristic roots of

$$\begin{pmatrix} \mathbf{0} & \mathbf{F} \\ \mathbf{F}' & \mathbf{0} \end{pmatrix} \tag{4.146}$$

are the pairs of positive and negative singular values of **F**; see [45].

A similar result to (4.142) is [17]

$$\psi(\mathbf{I}_m - \mathbf{FG}) = \psi(\mathbf{I}_n - \mathbf{GF}), \tag{4.147}$$

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

To prove (4.147) we use Theorem 2.5 to write

$$\mathrm{r}\begin{pmatrix} \mathbf{I}_m & \mathbf{F} \\ \mathbf{G} & \mathbf{I}_n \end{pmatrix} = n + \mathrm{r}(\mathbf{I}_m - \mathbf{FG}) = m + \mathrm{r}(\mathbf{I}_n - \mathbf{GF}), \tag{4.148}$$

from which (4.147) follows at once.

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

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

and

$$\mathrm{r}\begin{pmatrix} \mathbf{I}_m & \mathbf{F} \\ \mathbf{FG} & \mathbf{F} \end{pmatrix} = m + \mathrm{r}(\mathbf{F} - \mathbf{FGF}) = \mathrm{r}(\mathbf{F}) + \mathrm{r}(\mathbf{I}_m - \mathbf{FG}). \tag{4.150}$$

Hence

$$\psi(\mathbf{I} - \mathbf{FG}) = \mathrm{r}(\mathbf{F}) - \mathrm{r}(\mathbf{F} - \mathbf{FGF}) \tag{4.151a}$$

$$= \mathrm{r}(\mathbf{G}) - \mathrm{r}(\mathbf{G} - \mathbf{GFG}) \tag{4.151b}$$

$$= \psi(\mathbf{I} - \mathbf{GF}), \tag{4.151c}$$

and

$$\mathrm{r}(\mathbf{F}) - \mathrm{r}(\mathbf{G}) = \mathrm{r}(\mathbf{F} - \mathbf{FGF}) - \mathrm{r}(\mathbf{G} - \mathbf{GFG}). \tag{4.152}$$

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

$$\mathrm{r}(\mathbf{A}^-) = \mathrm{r}(\mathbf{A}) + \mathrm{r}(\mathbf{A}^- - \mathbf{A}^-\mathbf{A}\mathbf{A}^-) \geqslant \mathrm{r}(\mathbf{A}), \tag{4.153}$$

and so

$$\mathrm{r}(\mathbf{A}^-) = \mathrm{r}(\mathbf{A}) \quad \Leftrightarrow \quad \mathbf{A}^- = \mathbf{A}^-\mathbf{A}\mathbf{A}^-. \tag{4.154}$$

That is, a g-inverse $\mathbf{A}^-$ of a matrix $\mathbf{A}$ is reflexive if and only if the ranks of $\mathbf{A}^-$ and $\mathbf{A}$ are the same [8; 9, p. 383].

We may extend Theorem 2.6a [73], which showed that if $\mathbf{H} \neq \mathbf{0}$, then there exist column vectors $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a}'\mathbf{H}\mathbf{b} \neq 0$ and

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

THEOREM 4.11. *Let the matrices* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{H}$ *satisfy*

$$\mathrm{r}(\mathbf{A}'\mathbf{H}\mathbf{B}) = \mathrm{r}(\mathbf{A}'\mathbf{H}) = \mathrm{r}(\mathbf{H}\mathbf{B}). \tag{4.156}$$

Then

$$\mathrm{r}\big(\mathbf{H} - \mathbf{H}\mathbf{B}(\mathbf{A}'\mathbf{H}\mathbf{B})^-\mathbf{A}'\mathbf{H}\big) = \mathrm{r}(\mathbf{H}) - \mathrm{r}(\mathbf{A}'\mathbf{H}\mathbf{B}) \tag{4.157}$$

for any choice of generalized inverse.

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

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

where

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

Using the rank-cancellation rules of Marsaglia and Styan [47, Theorem 2], we see that

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

and

$$\mathrm{r}(\mathbf{A'HB}) = \mathrm{r}(\mathbf{A'H}) \quad \Rightarrow \quad \left[\mathbf{I} - \mathbf{A'HB}(\mathbf{A'HB})^{-}\right]\mathbf{A'H} = \mathbf{0}, \tag{4.161}$$

since $\mathbf{A'HB} = \mathbf{A'HB}(\mathbf{A'HB})^{-}\mathbf{A'HB}$. Hence (4.156)$\Rightarrow$(4.157), and the proof is complete. ∎

Rao [57, p. 69] presents (4.157) as an exercise when $\mathbf{A'HB}$ is square and nonsingular; this condition clearly implies (4.156).

In a statistical study of the residuals from a linear model, Ellenberg [23] 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*

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

If

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

then the Schur complement

$$(\mathbf{A}/\mathbf{E}) = \mathbf{H} - \mathbf{G}\mathbf{E}^{-}\mathbf{F} = (\mathbf{A}/\mathbf{E})^2 \tag{4.164}$$

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

$$r(\mathbf{A}/\mathbf{E}) = r(\mathbf{A}) - r(\mathbf{E}). \tag{4.165}$$

Proof. From (4.162) we obtain

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

while (4.163) yields, using Lemma 4.1,

$$\mathbf{EE}^-\mathbf{F} = \mathbf{F} \quad \text{and} \quad \mathbf{GE}^-\mathbf{E} = \mathbf{G} \tag{4.167}$$

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

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

$$= (\mathbf{H} - \mathbf{GF}) + \mathbf{GE}^-(\mathbf{E} - \mathbf{E}^2)\mathbf{E}^-\mathbf{F} - (\mathbf{G} - \mathbf{GE})\mathbf{E}^-\mathbf{F} - \mathbf{GE}^-(\mathbf{F} - \mathbf{EF}) \tag{4.168b}$$

$$\begin{aligned} &= \mathbf{H} - \mathbf{GF} + (\mathbf{GE}^-\mathbf{E})\mathbf{E}^-\mathbf{F} - (\mathbf{GE}^-\mathbf{E})(\mathbf{EE}^-\mathbf{F}) \\ &\quad - \mathbf{GE}^-\mathbf{F} + \mathbf{G}(\mathbf{EE}^-\mathbf{F}) - \mathbf{GE}^-\mathbf{F} + (\mathbf{GE}^-\mathbf{E})\mathbf{F} \end{aligned} \tag{4.168c}$$

$$= \mathbf{H} - \mathbf{GE}^-\mathbf{F} = (\mathbf{A}/\mathbf{E}), \tag{4.168d}$$

and (4.164) is proved; (4.165) follows using Corollary 4.3. ■

The special case of Theorem 4.12 considered by Ellenberg [23] supposed that $\mathbf{A}$ is symmetric and $\mathbf{E}$ nonsingular. It is clear that when the idempotent matrix $\mathbf{A}$ is symmetric, then it is nonnegative definite and so (4.163) always holds. Moreover, $\mathbf{E}$ nonsingular implies (4.163) even if $\mathbf{A} \neq \mathbf{A}'$. When $\mathbf{E}$ is symmetric idempotent but $\mathbf{A}$ is idempotent and not symmetric, then (4.164) need not hold, for let

$$\mathbf{A} = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right]. \tag{4.169}$$

Then $\mathbf{A}=\mathbf{A}^2$ and

$$\mathbf{E}^- = \begin{bmatrix} 1 & a & b \\ c & d & e \\ f & g & h \end{bmatrix} \tag{4.170}$$

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

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

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

A theorem by Milliken and Akdeniz [51] showed that if $\mathbf{H}$ and $\mathbf{E}-\mathbf{H}$ are both symmetric nonnegative definite matrices, then for the difference between the Moore-Penrose g-inverses we have

$$\mathbf{H}^+ - \mathbf{E}^+ \text{ is nnd} \quad \Leftrightarrow \quad \mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{H}). \tag{4.172}$$

This has been extended by Styan and Pukelsheim [70], who use symmetric reflexive g-inverses rather than Moore-Penrose g-inverses. See also Theorem 3.2.

THEOREM 4.13 (Styan and Pukelsheim [70]). *Let* $\mathbf{H}$ *and* $\mathbf{E}-\mathbf{H}$ *be symmetric nonnegative definite matrices. Then* $\mathbf{E}$ *is nonnegative definite and* $\mathrm{r}(\mathbf{E},\mathbf{H}) = \mathrm{r}(\mathbf{E})$. *Let* $\mathbf{E}_r^-$ *and* $\mathbf{H}_r^-$ *be symmetric reflexive g-inverses. Then*

$$\mathbf{H}_r^- - \mathbf{E}_r^- \textit{ is nnd} \quad \Leftrightarrow \quad \mathbf{E}\mathbf{E}_r^- = \mathbf{H}\mathbf{H}_r^-, \tag{4.173}$$

and then $\mathrm{r}(\mathbf{E}-\mathbf{H}) = \mathrm{r}(\mathbf{H}_r^- - \mathbf{E}_r^-)$.

Proof. Using Theorem 4.7, it follows that

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

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

$$\mathrm{In}\begin{pmatrix} \mathbf{E} & \mathbf{E}\mathbf{E}_r^- \\ \mathbf{E}_r^-\mathbf{E} & \mathbf{H}_r^- \end{pmatrix} = \mathrm{In}\,\mathbf{E} + \mathrm{In}(\mathbf{H}_r^- - \mathbf{E}_r^-), \tag{4.175}$$

using Theorem 4.7 again and the fact that $\mathrm{r}(\mathbf{E}, \mathbf{E}\mathbf{E}_r^-) = \mathrm{r}(\mathbf{E})$.

Let $\mathbf{E}\mathbf{E}_r^- = \mathbf{H}\mathbf{H}_r^-$. Then

$$\operatorname{In}\begin{pmatrix} \mathbf{E} & \mathbf{E}\mathbf{E}_r^- \\ \mathbf{E}_r^-\mathbf{E} & \mathbf{H}_r^- \end{pmatrix} = \operatorname{In}\begin{pmatrix} \mathbf{E} & \mathbf{H}\mathbf{H}_r^- \\ \mathbf{H}_r^-\mathbf{H} & \mathbf{H}_r^- \end{pmatrix}$$

$$= \operatorname{In}\mathbf{H}_r^- + \operatorname{In}(\mathbf{E}-\mathbf{H}), \tag{4.176}$$

using Theorem 4.7 and choosing $(\mathbf{H}_r^-)_r^- = \mathbf{H}$. Thus (4.175) = (4.176), and so $\mathbf{H}_r^- - \mathbf{E}_r^-$ is nnd and $\mathrm{r}(\mathbf{E}-\mathbf{H}) = \mathrm{r}(\mathbf{H}_r^- - \mathbf{E}_r^-)$.

Now, let $\mathbf{H}_r^- - \mathbf{E}_r^-$ be nonnegative definite. Then the partitioned matrix in (4.175) is nnd. Thus

$$\mathrm{r}(\mathbf{E}_r^-\mathbf{E}, \mathbf{H}_r^-) = r(\mathbf{H}_r^-) \tag{4.177}$$

which, in turn, implies

$$\left[\mathbf{I} - \mathbf{H}_r^-(\mathbf{H}_r^-)^-\right]\mathbf{E}_r^-\mathbf{E} = \mathbf{0} \tag{4.178}$$

so that, choosing $(\mathbf{H}_r^-)^- = \mathbf{H}$,

$$\mathbf{E}_r^-\mathbf{E} = \mathbf{H}_r^-\mathbf{H}\mathbf{E}_r^-\mathbf{E} = \mathbf{H}_r^-\mathbf{H}, \tag{4.179}$$

since $\mathrm{r}(\mathbf{E}, \mathbf{H}) = \mathrm{r}(\mathbf{E})$. Transposing yields $\mathbf{E}\mathbf{E}_r^- = \mathbf{H}\mathbf{H}_r^-$. ∎

COROLLARY 4.8 (Milliken and Akdeniz [51]). *Let* $\mathbf{H}$ *and* $\mathbf{E}-\mathbf{H}$ *be symmetric nonnegative definite matrices. Then*

$$\mathbf{H}^+ - \mathbf{E}^+ \textit{ is nnd} \quad \Leftrightarrow \quad \mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{H}). \tag{4.180}$$

Proof. If $\mathbf{H}^+ - \mathbf{E}^+$ is nonnegative definite, then

$$\mathbf{E}\mathbf{E}^+ = \mathbf{H}\mathbf{H}^+ \tag{4.181}$$

follows from (4.173), and

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

Now, suppose $\mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{H})$. Then Theorem 4.13 implies that $\mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{H}) = \mathrm{r}(\mathbf{E}, \mathbf{H})$. Then

$$\mathbf{E} = \mathbf{H}\mathbf{H}^+\mathbf{E} \quad \text{and} \quad \mathbf{H} = \mathbf{E}\mathbf{E}^+\mathbf{H}. \tag{4.183}$$

So postmultiplying the first equation in (4.183) by $\mathbf{E}^+$ yields

$$\mathbf{EE}^+ = \mathbf{HH}^+\mathbf{EE}^+ = (\mathbf{HH}^+\mathbf{EE}^+)' \tag{4.184a}$$

$$= (\mathbf{EE}^+)'(\mathbf{HH}^+)' \tag{4.184b}$$

$$= \mathbf{EE}^+\mathbf{HH}^+ = \mathbf{HH}^+, \tag{4.184c}$$

and the proof is complete. ■

Note that $\mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{H})$ does not always imply that $\mathbf{H}_r^- - \mathbf{E}_r^-$ is nonnegative definite. For example, let

$$\mathbf{E} = \mathbf{H} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.185}$$

which is nonnegative definite, so that

$$\mathbf{E} - \mathbf{H} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.186}$$

Then

$$\mathbf{E}_r^- = \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} \tag{4.187}$$

for some scalar x, and

$$\mathbf{H}_r^- = \begin{pmatrix} 1 & y \\ y & y^2 \end{pmatrix} \tag{4.188}$$

for some scalar y. Hence

$$\mathbf{H}_r^- - \mathbf{E}_r^- = \begin{pmatrix} 0 & y-x \\ y-x & y^2-x^2 \end{pmatrix} \tag{4.189}$$

is nonnegative definite if and only if $x = y$. But

$$\mathbf{EE}_r^- = \mathbf{HH}_r^- \quad \Leftrightarrow \quad \begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & y \\ 0 & 0 \end{pmatrix} \tag{4.190}$$

if and only if $x = y$. From this example, we conclude that, although $\mathrm{r}(\mathbf{E}) = \mathrm{r}(\mathbf{H})$

does not always imply that $\mathbf{H}_r^- - \mathbf{E}_r^-$ is nonnegative definite, the condition $\mathbf{E}\mathbf{E}_r^- = \mathbf{H}\mathbf{H}_r^-$ always does.

In a study of the existence of a nonnegative definite matrix with prescribed characteristic roots, Fiedler [26] based his proofs on a lemma, which Dias da Silva [20] 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 [26]). *Let* $\mathbf{A}$ *be a symmetric* $m\times m$ *matrix with characteristic roots* $\alpha_1, \alpha_2, \ldots, \alpha_m$, *and let* $\mathbf{u}$ *be a normalized characteristic vector corresponding to* α_1. *Let* $\mathbf{B}$ *be a symmetric* $n\times n$ *matrix with characteristic roots* $\beta_1, \beta_2, \ldots, \beta_n$, *and let* $\mathbf{v}$ *be a normalized characteristic vector corresponding to* β_1. *Then, for any* γ, *the matrix*

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \gamma\mathbf{u}\mathbf{v}' \\ \gamma\mathbf{v}\mathbf{u}' & \mathbf{B} \end{pmatrix} \tag{4.191}$$

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

$$p = \begin{vmatrix} \mathbf{A}-\lambda\mathbf{I} & \gamma\mathbf{u}\mathbf{v}' \\ \gamma\mathbf{v}\mathbf{u}' & \mathbf{B}-\lambda\mathbf{I} \end{vmatrix} \tag{4.193a}$$

$$= |\mathbf{A}-\lambda\mathbf{I}|\cdot|\mathbf{B}-\lambda\mathbf{I}-\gamma\mathbf{v}\mathbf{u}'(\mathbf{A}-\lambda\mathbf{I})^{-1}\gamma\mathbf{u}\mathbf{v}'|, \tag{4.193b}$$

using (2.4), and so

$$p = |\mathbf{A}-\lambda\mathbf{I}|\cdot|\mathbf{B}-\lambda\mathbf{I}-\gamma^2\mathbf{v}\mathbf{u}'(\alpha_1-\lambda)^{-1}\mathbf{u}\mathbf{v}'|, \tag{4.194}$$

since $(\mathbf{A}-\lambda\mathbf{I})\mathbf{u} = (\alpha_1-\lambda)\mathbf{u}$. Hence

$$p = |\mathbf{A}-\lambda\mathbf{I}|\cdot\left|\mathbf{B}-\lambda\mathbf{I}-\frac{\gamma^2\mathbf{v}\mathbf{v}'}{\alpha_1-\lambda}\right|, \tag{4.195}$$

since $\mathbf{u}'\mathbf{u}=1$, and so

$$p=|\mathbf{A}-\lambda\mathbf{I}|\cdot|\mathbf{B}-\lambda\mathbf{I}|\cdot\left|\mathbf{I}-\frac{\gamma^2(\mathbf{B}-\lambda\mathbf{I})^{-1}\mathbf{v}\mathbf{v}'}{\alpha_1-\lambda}\right| \tag{4.196a}$$

$$=|\mathbf{A}-\lambda\mathbf{I}|\cdot|\mathbf{B}-\lambda\mathbf{I}|\cdot\left|\mathbf{I}-\frac{\gamma^2\mathbf{v}\mathbf{v}'}{(\alpha_1-\lambda)(\beta_1-\lambda)}\right|, \tag{4.196b}$$

since $(\mathbf{B}-\lambda\mathbf{I})\mathbf{v}=(\beta_1-\lambda)\mathbf{v}$. Hence

$$p=\begin{vmatrix}\mathbf{A}-\lambda\mathbf{I} & \gamma\mathbf{u}\mathbf{v}' \\ \gamma\mathbf{v}\mathbf{u}' & \mathbf{B}-\lambda\mathbf{I}\end{vmatrix} \tag{4.197a}$$

$$=\left(\prod_{i=2}^{m}(\alpha_i-\lambda)\right)\cdot\left(\prod_{j=2}^{n}(\beta_j-\lambda)\right)\left[(\alpha_1-\lambda)(\beta_1-\lambda)-\gamma^2\right]. \quad \blacksquare \tag{4.197b}$$

Theorem 4.14 may be used to find the characteristic roots of a special correlation matrix structure; cf. the remarks after (6.70).

Emilie Haynsworth (1978, personal communication) has noted that Theorem 4.14 may be generalized, since our proof does not require that $\mathbf{A}$ and $\mathbf{B}$ be symmetric.

THEOREM 4.15. *Let* $\mathbf{A}$ *be an* $m\times m$ *complex matrix with characteristic roots* $\alpha_1,\alpha_2,\ldots,\alpha_m$, *and let* $\mathbf{a}$ *be a characteristic vector corresponding to* α_1. *Let* $\mathbf{B}$ *be an* $n\times n$ *complex matrix with characteristic roots* $\beta_1,\beta_2,\ldots,\beta_n$, *and let* $\mathbf{b}$ *be a characteristic vector corresponding to* β_1. *Let the complex row vectors* $\mathbf{c}^*$ *and* $\mathbf{d}^*$ *be* $1\times m$ *and* $1\times n$, *respectively. Then the matrix*

$$\begin{pmatrix}\mathbf{A} & \mathbf{a}\mathbf{d}^* \\ \mathbf{b}\mathbf{c}^* & \mathbf{B}\end{pmatrix} \tag{4.198}$$

has characteristic roots $\alpha_2,\ldots,\alpha_m,\beta_2,\ldots,\beta_n$, *and the characteristic roots of*

$$\begin{pmatrix}\alpha_1 & \mathbf{d}^*\mathbf{b} \\ \mathbf{c}^*\mathbf{a} & \beta_1\end{pmatrix}. \tag{4.199}$$

Proof. Following the proof of Theorem 4.14, the characteristic polynomial is

$$\begin{vmatrix}\mathbf{A}-\lambda\mathbf{I} & \mathbf{a}\mathbf{d}^* \\ \mathbf{b}\mathbf{c}^* & \mathbf{B}-\lambda\mathbf{I}\end{vmatrix}=|\mathbf{A}-\lambda\mathbf{I}|\cdot|\mathbf{B}-\lambda\mathbf{I}|\cdot h, \tag{4.200}$$

where

$$\begin{aligned} h &= |\mathbf{I}-(\mathbf{B}-\lambda\mathbf{I})^{-1}\mathbf{bc}^*(\mathbf{A}-\lambda\mathbf{I})^{-1}\mathbf{ad}^*| \\ &= |\mathbf{I}-(\beta_1-\lambda)^{-1}\mathbf{bc}^*(\alpha_1-\lambda)^{-1}\mathbf{ad}^*| \\ &= 1-\frac{\mathbf{c}^*\mathbf{a}\cdot\mathbf{d}^*\mathbf{b}}{(\alpha_1-\lambda)(\beta_1-\lambda)}, \end{aligned} \tag{4.201}$$

from which the result follows. ∎

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 [27].

5.1. *The bordering method*

Consider the matrix

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \tag{5.1}$$

The method proposed by Frazer, Duncan, and Collar [27, p. 114] considers the principal leading submatrices

$$\mathbf{E}_i=\begin{pmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix}, \qquad i=1,2,\ldots,n, \tag{5.2}$$

$$=\left[\begin{array}{c|c} \mathbf{E}_{i-1} & \begin{matrix} a_{1i} \\ \vdots \\ a_{i-1,i} \end{matrix} \\ \hline a_{i1} \quad \cdots \quad a_{i,i-1} & a_{ii} \end{array}\right], \qquad i=2,3,\ldots,n, \tag{5.3}$$

$$=\begin{pmatrix} \mathbf{E}_{i-1} & \mathbf{f}_{i-1} \\ \mathbf{g}'_{i-1} & a_{ii} \end{pmatrix}, \qquad i=2,3,\ldots,n, \tag{5.4}$$

where $\mathbf{f}_{i-1}=(a_{1i},\ldots,a_{i-1,i})'$ and $\mathbf{g}'_{i-1}=(a_{i1},\ldots,a_{i,i-1})$; $\mathbf{E}_1\equiv a_{11}$. Assuming $\mathbf{E}_{i-1}^{-1}$ known, we can apply (2.37). Then

$$\mathbf{E}_i^{-1}=\begin{bmatrix} \mathbf{E}_{i-1}^{-1}+\dfrac{\mathbf{E}_{i-1}^{-1}\mathbf{f}_{i-1}\mathbf{g}'_{i-1}\mathbf{E}_{i-1}^{-1}}{s_{i-1}} & -\dfrac{\mathbf{E}_{i-1}^{-1}\mathbf{f}_{i-1}}{s_{i-1}} \\ -\dfrac{\mathbf{g}'_{i-1}\mathbf{E}_{i-1}^{-1}}{s_{i-1}} & \dfrac{1}{s_{i-1}} \end{bmatrix}, \tag{5.5}$$

where $s_{i-1}=a_{ii}-\mathbf{g}'_{i-1}\mathbf{E}_{i-1}^{-1}\mathbf{f}_{i-1}$, $i=2,3,\ldots,n$. Define $s_0\equiv a_{11}$. This method requires that all the $\mathbf{E}_i$'s $(i=1,2,\ldots,n)$ be nonsingular. Hence, if one or more of the $\mathbf{E}_i$'s is singular, $i=1,2,\ldots,n$ (i.e., if at least one Schur complement $s_{i-1}=0$), then we can find a permutation matrix

$$\mathbf{\Pi}=(\mathbf{e}_{i_1},\mathbf{e}_{i_2},\ldots,\mathbf{e}_{i_n}), \tag{5.6}$$

where $\{i_1,i_2,\ldots,i_n\}$ is a permutation of $\{1,2,\ldots,n\}$, such that all the principal submatrices in $\mathbf{\Pi A}$ are nonsingular. Having obtained $(\mathbf{\Pi A})^{-1}$, we postmultiply by $\mathbf{\Pi}$ to obtain $\mathbf{A}^{-1}$.

Example 5.1. To find the inverse of

$$\mathbf{A}=\begin{bmatrix} 5 & 3 & -3 \\ 2 & -4 & 4 \\ 3 & 2 & 1 \end{bmatrix} \tag{5.7}$$

we write

$$\mathbf{A}=\begin{pmatrix} \mathbf{E}_2 & \mathbf{f}_2 \\ \mathbf{g}'_2 & a_{33} \end{pmatrix}, \tag{5.8}$$

where

$$\mathbf{E}_2=\begin{pmatrix} 5 & 3 \\ 2 & -4 \end{pmatrix}, \tag{5.9}$$

$\mathbf{f}_2=(-3,4)'$, $\mathbf{g}_2'=(3,2)$, and $a_{33}=1$. To apply (5.5) we compute

$$\mathbf{E}_2^{-1}=\frac{1}{26}\begin{pmatrix}4 & 3\\ 2 & -5\end{pmatrix}=\begin{bmatrix}\frac{2}{13} & \frac{3}{26}\\ \frac{1}{13} & -\frac{5}{26}\end{bmatrix}, \tag{5.10}$$

$$\mathbf{E}_2^{-1}\mathbf{f}_2=\frac{1}{26}\begin{pmatrix}4 & 3\\ 2 & -5\end{pmatrix}\begin{pmatrix}-3\\ 4\end{pmatrix}=\begin{pmatrix}0\\ -1\end{pmatrix}, \tag{5.11}$$

$$\mathbf{g}_2'\mathbf{E}_2^{-1}=\tfrac{1}{26}(3\quad 2)\begin{pmatrix}4 & 3\\ 2 & -5\end{pmatrix}=\tfrac{1}{26}(16\quad -1)$$

$$=\left(\tfrac{8}{13}\quad -\tfrac{1}{26}\right), \tag{5.12}$$

$$s_2=a_{33}-\mathbf{g}_2'\mathbf{E}_2^{-1}\mathbf{f}_2=1-(3\quad 2)\begin{pmatrix}0\\ -1\end{pmatrix}=3, \tag{5.13}$$

$$\mathbf{E}_2^{-1}+\frac{\mathbf{E}_2^{-1}\mathbf{f}_2\mathbf{g}_2'\mathbf{E}_2^{-1}}{s_2}=\frac{1}{26}\begin{pmatrix}4 & 3\\ 2 & -5\end{pmatrix}+\frac{1}{26}\begin{pmatrix}0\\ -1\end{pmatrix}(16\quad -1)/3$$

$$=\frac{1}{78}\begin{pmatrix}12 & 9\\ -10 & -14\end{pmatrix}=\begin{bmatrix}\frac{2}{13} & \frac{3}{26}\\ -\frac{5}{39} & -\frac{7}{39}\end{bmatrix}. \tag{5.14}$$

Hence

$$\mathbf{A}^{-1}=\frac{1}{78}\begin{bmatrix}12 & 9 & 0\\ -10 & -14 & 26\\ -16 & 1 & 26\end{bmatrix}=\begin{bmatrix}\frac{2}{13} & \frac{3}{26} & 0\\ -\frac{5}{39} & -\frac{7}{39} & \frac{1}{3}\\ -\frac{8}{39} & \frac{1}{78} & \frac{1}{3}\end{bmatrix}. \tag{5.15}$$

A variant of the above method was given by Jossa [43], who showed that, when $\mathbf{E}_{i-1}^{-1}$ is known, then the following operations yield $\mathbf{E}_i^{-1}=\{a_i^{k,l}\}$,

$i=2,3,\ldots,n$. Define $a_1^{1,1}=1/a_{11}$. For $i=2,3,\ldots,n$, set

$$r_{ki}=-\sum_{h=1}^{i-1} a_{i-1}^{k,h}a_{hi}, \qquad k=1,2,\ldots,i-1, \tag{5.16}$$

$$a_i^{i,i}=\left(a_{ii}+\sum_{h=1}^{i-1} a_{ih}r_{hi}\right)^{-1}, \tag{5.17}$$

$$a_i^{k,i}=a_i^{i,i}r_{ki}, \qquad k=1,2,\ldots,i-1, \tag{5.18}$$

$$a_i^{i,l}=-a_i^{i,i}\left(\sum_{h=1}^{i-1} a_{ih}a_{i-1}^{h,l}\right), \qquad l=1,2,\ldots,i-1, \tag{5.19}$$

$$a_i^{k,l}=a_{i-1}^{k,l}+r_{ki}a_i^{i,l}, \qquad k,l=1,2,\ldots,i-1. \tag{5.20}$$

The above equations may be obtained simply by rewriting (5.5) in scalar notation.

EXAMPLE 5.1 (reprise). Using Jossa's method, we obtain

$$a_1^{1,1}=\tfrac{1}{5},$$

$$\text{(5.16)} \qquad r_{12}=-a_1^{1,1}a_{12}=-\tfrac{3}{5} \qquad \begin{cases} r_{13}=-\left(a_2^{1,1}a_{13}+a_2^{1,2}a_{23}\right)=0 \\ r_{23}=-\left(a_2^{2,1}a_{13}+a_2^{2,2}a_{23}\right)=1 \end{cases}$$

$$\text{(5.17)} \quad a_2^{2,2}=\frac{1}{a_{22}+a_{21}r_{12}}=-\tfrac{5}{26} \quad a_3^{3,3}=\frac{1}{a_{33}+a_{31}r_{13}+a_{32}r_{23}}=\tfrac{1}{3}$$

$$\text{(5.18)} \qquad a_2^{1,2}=a_2^{2,2}r_{12}=\tfrac{3}{26} \qquad \begin{cases} a_3^{1,3}=a_3^{3,3}r_{13}=0 \\ a_3^{2,3}=a_3^{3,3}r_{23}=\tfrac{1}{3} \end{cases}$$

$$\text{(5.19)} \quad a_2^{2,1}=-a_2^{2,2}a_{21}a_1^{1,1}=\tfrac{2}{26} \quad \begin{cases} a_3^{3,1}=-a_3^{3,3}\left(a_{31}a_2^{1,1}+a_{32}a_2^{2,1}\right)=-\tfrac{16}{78} \\ a_3^{3,2}=-a_3^{3,3}\left(a_{31}a_2^{1,2}+a_{32}a_2^{2,2}\right)=\tfrac{1}{78} \end{cases}$$

$$\text{(5.20)} \quad a_2^{1,1}=a_1^{1,1}+r_{12}a_2^{2,1}=\tfrac{4}{26} \quad \begin{cases} a_3^{1,1}=a_2^{1,1}+r_{13}a_3^{3,1}=\tfrac{4}{26}=\tfrac{12}{78} \\ a_3^{1,2}=a_2^{1,2}+r_{13}a_3^{3,2}=\tfrac{3}{26}=\tfrac{9}{78} \\ a_3^{2,1}=a_2^{2,1}+r_{23}a_3^{3,1}=-\tfrac{10}{78} \\ a_3^{2,2}=a_2^{2,2}+r_{23}a_3^{3,2}=-\tfrac{14}{78}. \end{cases}$$

Hence, we obtain (5.15).

Louis Guttman [32] called this bordering method "first order enlargement" in view of the partitioning (5.3) adding a single row and column to $\mathbf{E}_{i-1}$. We now consider the partitioning with $\mathbf{E}_{i-2}$ bordered by 2 rows and 2 columns:

$$\mathbf{E}_i=\left[\begin{array}{ccc|cc} & \mathbf{E}_{i-2} & & \begin{matrix} a_{1,i-1} \\ \cdots \\ a_{i-2,i-1}\end{matrix} & \begin{matrix} a_{1,i} \\ \cdots \\ a_{i-2,i}\end{matrix} \\ \hline a_{i-1,1} & \cdots & a_{i-1,i-2} & a_{i-1,i-1} & a_{i-1,i} \\ a_{i,1} & \cdots & a_{i,i-2} & a_{i,i-1} & a_{i,i} \end{array}\right]$$

$$=\begin{pmatrix} \mathbf{E}_{i-2} & \mathbf{F}_{i-2} \\ \mathbf{G}_{i-2} & \mathbf{H}_{i-2} \end{pmatrix}, \qquad i=3,4,\ldots,n, \tag{5.21}$$

where

$$\mathbf{F}_{i-2}=\begin{bmatrix} a_{1,i-1} & a_{1,i} \\ \vdots & \vdots \\ a_{i-2,i-1} & a_{i-2,i} \end{bmatrix} \qquad [(i-2)\times 2], \tag{5.22}$$

$$\mathbf{G}_{i-2}=\begin{pmatrix} a_{i-1,1} & \cdots & a_{i-1,i-2} \\ a_{i,1} & \cdots & a_{i,i-2} \end{pmatrix} \qquad [2\times(i-2)], \tag{5.23}$$

$$\mathbf{H}_{i-2}=\begin{pmatrix} a_{i-1,i-1} & a_{i-1,i} \\ a_{i,i-1} & a_{ii} \end{pmatrix} \qquad [2\times 2]. \tag{5.24}$$

Assuming $\mathbf{E}_{i-2}^{-1}$ known, we can apply (2.37).

Example 5.2. To find the inverse of the matrix $\mathbf{A}$ defined by (5.7) using the method described above, we write

$$\mathbf{A}=\begin{pmatrix} e_1 & \mathbf{f}_1' \\ \mathbf{g}_1 & \mathbf{H}_1 \end{pmatrix}, \tag{5.25}$$

where $e_1=5$, $\mathbf{f}_1'=(3,-3)$, $\mathbf{g}_1=(2,3)'$ and

$$\mathbf{H}_1=\begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix}. \tag{5.26}$$

The Schur complement $\mathbf{S}=(\mathbf{A}/e_1)$ is given by

$$\mathbf{S}=\begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix}-\tfrac{1}{5}\begin{pmatrix} 2 \\ 3 \end{pmatrix}(3 \quad -3)=\begin{bmatrix} -\frac{26}{5} & \frac{26}{5} \\ \frac{1}{5} & \frac{14}{5} \end{bmatrix}. \tag{5.27}$$

Applying (2.37), we obtain

$$\mathbf{A}^{-1}=\begin{bmatrix} \dfrac{1}{e_1}+\dfrac{1}{e_1^2}\mathbf{f}_1'\mathbf{S}^{-1}\mathbf{g}_1 & -\dfrac{\mathbf{f}_1'\mathbf{S}^{-1}}{e_1} \\ -\dfrac{\mathbf{S}^{-1}\mathbf{g}_1}{e_1} & \mathbf{S}^{-1} \end{bmatrix}. \tag{5.28}$$

To apply (5.28) we compute:

$$\mathbf{S}^{-1}=-\tfrac{5}{78}\begin{bmatrix} \frac{14}{5} & -\frac{26}{5} \\ -\frac{1}{5} & -\frac{26}{5} \end{bmatrix}=\begin{bmatrix} -\frac{7}{39} & \frac{1}{3} \\ \frac{1}{78} & \frac{1}{3} \end{bmatrix}, \tag{5.29}$$

$$-\frac{\mathbf{f}_1'\mathbf{S}^{-1}}{e_1}=\tfrac{1}{78}\begin{pmatrix} 3 & -3 \end{pmatrix}\begin{bmatrix} \frac{14}{5} & -\frac{26}{5} \\ -\frac{1}{5} & -\frac{26}{5} \end{bmatrix}$$

$$=\begin{pmatrix} \frac{9}{78} & 0 \end{pmatrix}=\begin{pmatrix} \frac{3}{26} & 0 \end{pmatrix}, \tag{5.30}$$

$$-\frac{\mathbf{S}^{-1}\mathbf{g}_1}{e_1}=\tfrac{1}{78}\begin{bmatrix} \frac{14}{5} & -\frac{26}{5} \\ -\frac{1}{5} & -\frac{26}{5} \end{bmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix}=\begin{bmatrix} -\frac{10}{78} \\ -\frac{16}{78} \end{bmatrix}=\begin{bmatrix} -\frac{5}{39} \\ -\frac{8}{39} \end{bmatrix}, \tag{5.31}$$

$$\frac{1}{e_1}+\frac{1}{e_1^2}\mathbf{f}_1'\mathbf{S}^{-1}\mathbf{g}_1=\tfrac{1}{5}-\tfrac{1}{5}\begin{pmatrix} \frac{9}{78} & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix}=\tfrac{12}{78}=\tfrac{2}{13}. \tag{5.32}$$

Hence, we obtain (5.15).

5.2. *Geometric enlargement*

The method of "geometric enlargement" due to Louis Guttman [32] allows the inverse of the matrix $\mathbf{A}$ in (5.1) to be obtained by successively constructing the inverses of the principal submatrices

$$\mathbf{E}_{2i}=\begin{pmatrix} a_{11} & \cdots & a_{1,2i} \\ \vdots & \ddots & \vdots \\ a_{2i,1} & \cdots & a_{2i,2i} \end{pmatrix} \qquad (2i\times 2i) \tag{5.33}$$

$$=\begin{pmatrix} \mathbf{E}_i & \mathbf{F}_i \\ \mathbf{G}_i & \mathbf{H}_i \end{pmatrix}, \qquad i=1,2,\ldots,\left[\tfrac{n}{2}\right]. \tag{5.34}$$

Assuming $\mathbf{E}_i^{-1}$ is known, we can apply (2.37). If $i=1$, then

$$\mathbf{E}_2=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}; \tag{5.35}$$

setting $\mathbf{E}_1 = a_{11}$, we see that "geometric enlargement" reduces to "first-order enlargement." Also, if $i=2$, we have

$$\mathbf{E}_4=\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \tag{5.36}$$

and $\mathbf{E}_2$ 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

$$\mathbf{A}=\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}, \tag{5.37}$$

where $\mathbf{E}$ is $n_0\times n_0$, nonsingular, and readily invertible (e.g., $n_0=1$ or 2, $\mathbf{E}$ diagonal). We then compute the Schur complement $\mathbf{S}=(\mathbf{A}/\mathbf{E})$, and if it is easily invertible, then we compute $\mathbf{A}^{-1}$ using the Schur-Banachiewicz formula (2.37). Otherwise we partition the Schur complement

$$\mathbf{S}=\begin{pmatrix} \mathbf{E}_1 & \mathbf{F}_1 \\ \mathbf{G}_1 & \mathbf{H}_1 \end{pmatrix}, \tag{5.38}$$

where $\mathbf{E}_1$ is $n_1\times n_1$, nonsingular, and readily invertible. We now compute the Schur complement $\mathbf{S}_1=(\mathbf{S}/\mathbf{E}_1)$, and if it is easily invertible, then we compute $\mathbf{S}^{-1}$ using (2.37), from which $\mathbf{A}^{-1}$ follows using (2.37) again. Otherwise we partition

$$\mathbf{S}_1=\begin{pmatrix} \mathbf{E}_2 & \mathbf{F}_2 \\ \mathbf{G}_2 & \mathbf{H}_2 \end{pmatrix}, \tag{5.39}$$

where $\mathbf{E}_2$ is $n_2 \times n_2$, nonsingular, and readily invertible. We compute $\mathbf{S}_2 = (\mathbf{S}_1/\mathbf{E}_2)$ and repeat the procedure performed with $\mathbf{S}_1$. And so on. Writing

$$\mathbf{S}_k = \begin{pmatrix} \mathbf{E}_{k+1} & \mathbf{F}_{k+1} \\ \mathbf{G}_{k+1} & \mathbf{H}_{k+1} \end{pmatrix}, \qquad k = 1, 2, \ldots, m-1, \tag{5.40}$$

the forward part of this algorithm stops at $k = m$ when $\mathbf{S}_m = (\mathbf{S}_{m-1}/\mathbf{E}_m)$ is easily invertible. At most $m = n-2$. Some $m+1$ **E**-matrices will have been inverted. We now invert $\mathbf{S}_m$ and proceed backwards, computing in turn each inverse $\mathbf{S}_k^{-1}$, $k = m-1, m-2, \ldots, 3, 2, 1, 0$, using $\mathbf{S}_{k+1}^{-1}$ and (2.37), with $\mathbf{S} = \mathbf{S}_0$. $\mathbf{A}^{-1}$ follows using (2.37) again. Louis Guttman sketched the above algorithm with $n_k = 1$ or 2; $k = 0, 1, \ldots, m$. Zlobec and Chan [32] 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

$$\mathbf{A} = \begin{bmatrix} 5 & 3 & -3 \\ 2 & -4 & 4 \\ 3 & 2 & 1 \end{bmatrix} \tag{5.41}$$

using partitioned Schur complements.

The forward part. We may write

$$\mathbf{A} = \begin{pmatrix} e & \mathbf{f}' \\ \mathbf{g} & \mathbf{H} \end{pmatrix} \tag{5.42}$$

where $e = 5 (\neq 0)$, $\mathbf{f}' = (3, -3), \mathbf{g}' = (2, 3)$, and

$$\mathbf{H} = \begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix}, \tag{5.43}$$

$$\mathbf{S} = \begin{pmatrix} -4 & 4 \\ 2 & 1 \end{pmatrix} - \tfrac{1}{5} \begin{pmatrix} 2 \\ 3 \end{pmatrix} (3, -3) \tag{5.44}$$

$$= \begin{bmatrix} -\frac{26}{5} & \frac{26}{5} \\ \frac{1}{5} & \frac{14}{5} \end{bmatrix}. \tag{5.45}$$

Partition **S** as follows:

$$\mathbf{S} = \begin{pmatrix} e_1 & f_1 \\ g_1 & h_1 \end{pmatrix}, \tag{5.46}$$

where $e_1 = -\frac{28}{5}$ $(\neq 0)$, $f_1 = \frac{28}{5}$, $g_1 = \frac{1}{5}$, $h_1 = \frac{14}{5}$. Hence,

$$s_1 = \tfrac{14}{5} - \left(\tfrac{1}{5}\right)\left(-\tfrac{28}{5}\right)^{-1}\left(\tfrac{28}{5}\right) = 3. \tag{5.47}$$

The backward part. $1/s_1 = \frac{1}{3}$. Applying (2.37), we obtain

$$\mathbf{S}^{-1} = \begin{bmatrix} -\frac{14}{78} & \frac{1}{3} \\ \frac{1}{78} & \frac{1}{3} \end{bmatrix}. \tag{5.48}$$

Applying (2.37) again, we obtain (5.15).

5.4. *Rank annihilation*

We express the $n \times n$ nonsingular matrix $\mathbf{A}$ as the sum of a nonsingular matrix $\mathbf{D}$ and the sum of h matrices each of rank one (cf. [75]),

$$\mathbf{A} = \mathbf{D} + \sum_{i=1}^{h} \mathbf{f}_i \mathbf{g}_i'. \tag{5.49}$$

The matrix $\mathbf{D}$ is easily invertible, e.g., diagonal. Clearly $h \leqslant n$. Let us write

$$\begin{aligned} &\mathbf{E}_0 = \mathbf{D} \\ &\mathbf{E}_1 = \mathbf{D} + \mathbf{f}_1 \mathbf{g}_1', \\ &\quad\vdots \\ &\mathbf{E}_j = \mathbf{E}_{j-1} + \mathbf{f}_j \mathbf{g}_j', \qquad j = 1, 2, \ldots, h-1, \\ &\quad\vdots \\ &\mathbf{E}_h = \mathbf{A}. \end{aligned} \tag{5.50}$$

Then we compute, in turn, $\mathbf{E}_0^{-1}, \mathbf{E}_1^{-1}, \ldots, \mathbf{E}_h^{-1} = \mathbf{A}^{-1}$ using (2.59),

$$\mathbf{E}_j^{-1} = \mathbf{E}_{j-1}^{-1} - \frac{\mathbf{E}_{j-1}^{-1} \mathbf{f}_j \mathbf{g}_j' \mathbf{E}_{j-1}^{-1}}{1 + \mathbf{g}_j' \mathbf{E}_{j-1}^{-1} \mathbf{f}_j}, \tag{5.51}$$

where

$$1 + \mathbf{g}_j' \mathbf{E}_{j-1}^{-1} \mathbf{f}_j \neq 0, \qquad j = 1, \ldots, h. \tag{5.52}$$

This method requires that all the $\mathbf{E}_j$'s $(j=0,1,\ldots,h)$ be nonsingular; $\mathbf{A}^{-1}=\mathbf{E}_h^{-1}$.

Edelblute [22] considered the special case of (5.49) with $\mathbf{D}=\mathbf{I}$, $\mathbf{f}_j=(\mathbf{A}-\mathbf{I})\mathbf{e}_j$, $\mathbf{g}_i=\mathbf{e}_i$, and $h=n$. Then

$$\begin{aligned}
&\mathbf{A}=\mathbf{I}+\sum_{j=1}^{n}(\mathbf{A}-\mathbf{I})\mathbf{e}_j\mathbf{e}_j',\\
&\mathbf{E}_0=\mathbf{I},\\
&\mathbf{E}_1=\mathbf{I}+(\mathbf{A}-\mathbf{I})\mathbf{e}_1\mathbf{e}_1',\\
&\quad\vdots\\
&\mathbf{E}_j=\mathbf{E}_{j-1}+(\mathbf{A}-\mathbf{I})\mathbf{e}_j\mathbf{e}_j', \qquad j=1,\ldots,n,\\
&\quad\vdots\\
&\mathbf{E}_n=\mathbf{A}.
\end{aligned}\tag{5.53}$$

Hence

$$\mathbf{E}_1^{-1}=\mathbf{I}-\frac{(\mathbf{A}-\mathbf{I})\mathbf{e}_1\mathbf{e}_1'}{a_{11}},\tag{5.54a}$$

$$\mathbf{E}_j^{-1}=\mathbf{E}_{j-1}^{-1}-\frac{\mathbf{E}_{j-1}^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_j\mathbf{e}_j'\mathbf{E}_{j-1}^{-1}}{1+\mathbf{e}_j'\mathbf{E}_{j-1}^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_j}, \qquad j=2,\ldots,n.\tag{5.54b}$$

Example 5.4. Find the inverse of

$$\mathbf{A}=\begin{bmatrix}5 & 3 & -3\\ 2 & -4 & 4\\ 3 & 2 & 1\end{bmatrix}\tag{5.55}$$

by rank annihilation.

We write

$$\mathbf{A}=\mathbf{I}+\sum_{i=1}^{3}\mathbf{f}_i\mathbf{e}_i',\tag{5.56}$$

where $\mathbf{f}_1=(4,2,3)'$, $\mathbf{f}_2=(3,-5,2)'$, and $\mathbf{f}_3=(-3,4,0)'$. Then

$$\mathbf{E}_1^{-1}=\mathbf{I}-\begin{bmatrix}4\\2\\3\end{bmatrix}(1\quad 0\quad 0)/5 \tag{5.57}$$

$$=\begin{bmatrix}\frac{1}{5} & 0 & 0\\ -\frac{2}{5} & 1 & 0\\ -\frac{3}{5} & 0 & 1\end{bmatrix}=\frac{1}{5}\begin{bmatrix}1 & 0 & 0\\ -2 & 5 & 0\\ -3 & 0 & 5\end{bmatrix}. \tag{5.58}$$

Thus, using (5.51),

$$\mathbf{E}_2^{-1}=\mathbf{E}_1^{-1}-\mathbf{E}_1^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_2\mathbf{e}_2'\mathbf{E}_1^{-1}/a, \tag{5.59}$$

where

$$a=1+\left(-\tfrac{2}{5},1,0\right)\begin{bmatrix}3\\-5\\2\end{bmatrix}=-\frac{26}{5}, \tag{5.60}$$

so that

$$\mathbf{E}_2^{-1}=\mathbf{E}_1^{-1}+\tfrac{5}{26}\mathbf{E}_1^{-1}\begin{bmatrix}3\\-5\\2\end{bmatrix}\left(-\tfrac{2}{5},1,0\right) \tag{5.61}$$

$$=\mathbf{E}_1^{-1}+\frac{1}{26}\begin{bmatrix}3\\-31\\1\end{bmatrix}\left(-\tfrac{2}{5},1,0\right) \tag{5.62}$$

$$=\frac{1}{130}\begin{bmatrix}26 & 0 & 0\\ -52 & 130 & 0\\ -78 & 0 & 130\end{bmatrix}+\frac{1}{130}\begin{bmatrix}-6 & 15 & 0\\ 62 & -155 & 0\\ -2 & 5 & 0\end{bmatrix} \tag{5.63}$$

$$=\frac{1}{26}\begin{bmatrix}4 & 3 & 0\\ 2 & -5 & 0\\ -16 & 1 & 26\end{bmatrix}. \tag{5.64}$$

Hence

$$\mathbf{A}^{-1}=\mathbf{E}_3^{-1}=\mathbf{E}_2^{-1}-\mathbf{E}_2^{-1}\begin{bmatrix}-3\\4\\0\end{bmatrix}(0,0,1)\mathbf{E}_2^{-1}/b, \tag{5.65}$$

where

$$b = 1 + (0,0,1)\mathbf{E}_2^{-1}\begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix} = 3. \tag{5.66}$$

Thus

$$\mathbf{A}^{-1} = \mathbf{E}_2^{-1} - \frac{1}{78}\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}(-16,1,26) \tag{5.67}$$

$$= \frac{1}{78}\begin{bmatrix} 12 & 9 & 0 \\ -10 & -14 & 26 \\ -16 & 1 & 26 \end{bmatrix} \tag{5.68}$$

[cf. (5.15)].

5.5. Complex matrices

Let the $n \times n$ complex matrix $\mathbf{E} + i\mathbf{F}$, where $\mathbf{E}, \mathbf{F}$ are both real matrices, be nonsingular, and let us write its inverse as $\mathbf{K} + i\mathbf{L}$, where $\mathbf{K}$ and $\mathbf{L}$ are real. Then [54; 74, p. 31]

$$\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ -\mathbf{F} & \mathbf{E} \end{pmatrix}\begin{pmatrix} \mathbf{K} & \mathbf{L} \\ -\mathbf{L} & \mathbf{K} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{5.69}$$

Thus, we note that $\mathbf{E} + i\mathbf{F}$ is nonsingular if and only if

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ -\mathbf{F} & \mathbf{E} \end{pmatrix} \tag{5.70}$$

is nonsingular. If $\mathbf{E}$ is nonsingular, then $\mathbf{E} + i\mathbf{F}$ is nonsingular if and only if the Schur complement $(\mathbf{A}/\mathbf{E}) = \mathbf{E} + \mathbf{F}\mathbf{E}^{-1}\mathbf{F}$ is nonsingular [cf. (2.4)], and then

$$\mathbf{K} = (\mathbf{E} + \mathbf{F}\mathbf{E}^{-1}\mathbf{F})^{-1}, \qquad \mathbf{L} = -\mathbf{E}^{-1}\mathbf{F}(\mathbf{E} + \mathbf{F}\mathbf{E}^{-1}\mathbf{F})^{-1} \tag{5.71}$$

[cf. (2.37)].

If $\mathbf{E}$ is singular we may rearrange the columns of $\mathbf{A}$ in order to obtain a submatrix in the top left-hand corner which is nonsingular. This is possible because (5.69) implies that $r(\mathbf{E}, \mathbf{F}) = n$. 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 Sec. 5.3 by partitioning the rectangular or singular matrix

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix}, \tag{5.72}$$

where $\mathbf{E}$ is $n_0 \times n_0$, nonsingular, and readily invertible (e.g., $n_0=1$ or 2, $\mathbf{E}$ diagonal). We then compute the Schur complement $\mathbf{S}=(\mathbf{A}/\mathbf{E})$, and if it is easy to find a g-inverse of $\mathbf{S}$, then we compute (cf. [77]) a g-inverse $\mathbf{A}^-$ using [cf. (4.28) and (4.61)]

THEOREM 5.1. *If*

$$\mathbf{A}=\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix} \tag{5.73}$$

and $\mathbf{E}$ *is nonsingular, then*

$$\begin{pmatrix}\mathbf{E}^{-1}+\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-}\mathbf{G}\mathbf{E}^{-1} & -\mathbf{E}^{-1}\mathbf{F}\mathbf{S}^{-}\\ -\mathbf{S}^{-}\mathbf{G}\mathbf{E}^{-1} & \mathbf{S}^{-}\end{pmatrix}=\mathbf{A}^{-}, \tag{5.74}$$

where $\mathbf{S}=\mathbf{H}-\mathbf{G}\mathbf{E}^{-1}\mathbf{F}$.

Otherwise we partition the Schur complement

$$\mathbf{S}=\begin{pmatrix}\mathbf{E}_1 & \mathbf{F}_1\\ \mathbf{G}_1 & \mathbf{H}_1\end{pmatrix}, \tag{5.75}$$

where $\mathbf{E}_1$ is $n_1 \times n_1$, nonsingular, and readily invertible. We now compute the Schur complement $\mathbf{S}_1=(\mathbf{S}/\mathbf{E}_1)$, and if it is easy to find an $\mathbf{S}_1^-$, then we compute $\mathbf{S}^-$ using Theorem 5.1, from which $\mathbf{A}^-$ follows using Theorem 5.1 again. Otherwise we partition

$$\mathbf{S}_1=\begin{pmatrix}\mathbf{E}_2 & \mathbf{F}_2\\ \mathbf{G}_2 & \mathbf{H}_2\end{pmatrix}, \tag{5.76}$$

where $\mathbf{E}_2$ is $n_2 \times n_2$, nonsingular, and readily invertible. We compute $\mathbf{S}_2=(\mathbf{S}_1/\mathbf{E}_2)$ and repeat the procedure performed with $\mathbf{S}_1$. And so on. Writing

$$\mathbf{S}_k=\begin{pmatrix}\mathbf{E}_{k+1} & \mathbf{F}_{k+1}\\ \mathbf{G}_{k+1} & \mathbf{H}_{k+1}\end{pmatrix}, \qquad k=1,2,\ldots,r-1, \tag{5.77}$$

the forward part of this algorithm stops at $k=r$ when $\mathbf{S}_r=(\mathbf{S}_{r-1}/\mathbf{E}_r)$ has a g-inverse $\mathbf{S}_r^-$ which is easy to find. At most $r=\mathrm{r}(\mathbf{A})$. We now proceed backwards computing in turn each g-inverse $\mathbf{S}_k^-$, $k=r-1,r-2,\ldots,3,2,1,0$, using $\mathbf{S}_{k+1}^-$ and Theorem 5.1, with $\mathbf{S}=\mathbf{S}_0$. And so $\mathbf{A}^-$ follows using Theorem 5.1 again.

EXAMPLE 5.5. Find a g-inverse of

$$\mathbf{A}=\begin{bmatrix}2&1&2\\2&0&2\\0&1&0\end{bmatrix}\tag{5.78}$$

using partitioned Schur complements.

The matrix is clearly singular and has rank 2. We may partition it as

$$\mathbf{A}=\begin{pmatrix}e&\mathbf{f}'\\\mathbf{g}&\mathbf{H}\end{pmatrix},\tag{5.79}$$

where $e=2$, $\mathbf{f}'=(1,2)$, $\mathbf{g}=(2,0)'$, and

$$\mathbf{H}=\begin{pmatrix}0&2\\1&0\end{pmatrix}.\tag{5.80}$$

Then

$$\mathbf{S}=(\mathbf{A}/\mathbf{E})=\begin{pmatrix}0&2\\1&0\end{pmatrix}-\frac{1}{2}\begin{pmatrix}2\\0\end{pmatrix}(1,2)=\begin{pmatrix}-1&0\\1&0\end{pmatrix}.\tag{5.81}$$

Noting that

$$\mathbf{S}=\begin{pmatrix}-1\\1\end{pmatrix}(1\quad 0)\tag{5.82}$$

we see at once that

$$\tfrac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}(-1\quad 1)=\mathbf{S}^+,\tag{5.83}$$

the Moore-Penrose g-inverse of $\mathbf{S}$. Hence we use Theorem 5.1 to compute

$$\mathbf{A}^-=\begin{bmatrix}\frac{1}{2}+\frac{1}{4}(1,2)\mathbf{S}^+(2,0)' & -\frac{1}{2}(1,2)\mathbf{S}^+\\ -\frac{1}{2}\mathbf{S}^+(2,0)' & \mathbf{S}^+\end{bmatrix}$$

$$=\begin{bmatrix}\frac{1}{4}&\frac{1}{4}&-\frac{1}{4}\\\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\0&0&0\end{bmatrix}.\tag{5.84}$$

If we had not noticed the factorization (5.82), we could have partitioned (5.81) as

$$\mathbf{S}=\left(\begin{array}{c|c} -1 & 0 \\ \hline 1 & 0 \end{array}\right) \tag{5.85}$$

and computed, using Theorem 5.1, the g-inverse

$$\mathbf{S}^{\sim}=\left(\begin{array}{c|c} -1 & 0 \\ \hline 1 & 1 \end{array}\right). \tag{5.86}$$

Hence, again using Theorem 5.1, we obtain the alternative g-inverse

$$\mathbf{A}^{\sim}=\begin{bmatrix} \frac{1}{2}+\frac{1}{4}(1,2)\mathbf{S}^{\sim}(2,0)' & -\frac{1}{2}(1,2)\mathbf{S}^{\sim} \\ -\frac{1}{2}\mathbf{S}^{\sim}(2,0)' & \mathbf{S}^{\sim} \end{bmatrix}$$

$$=\begin{bmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix}. \tag{5.87}$$

A third g-inverse of $\mathbf{A}$ may be found by noting that any g-inverse of $\mathbf{S}$ must have the form

$$\mathbf{S}^{-}=\begin{pmatrix} a & 1+a \\ b & c \end{pmatrix}, \tag{5.88}$$

where a, b, and c are arbitrary scalars. Thus $a=-\frac{1}{2}$ and $b=c=0$ yield $\mathbf{S}^{+}$, while $a=-1$ and $b=c=1$ yield $\mathbf{S}^{\sim}$. Letting $a=b=c=0$, we obtain

$$\mathbf{S}^{-}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{5.89}$$

and using Theorem 5.1 again yields

$$\mathbf{A}^{-}=\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \tag{5.90}$$

5.7. *Generalized inversion by rank annihilation*

When $1+\mathbf{g}'\mathbf{E}^{-1}\mathbf{f}\neq 0$, the expression

$$(\mathbf{E}+\mathbf{f}\mathbf{g}')^{-1}=\mathbf{E}^{-1}-\frac{\mathbf{E}^{-1}\mathbf{f}\mathbf{g}'\mathbf{E}^{-1}}{1+\mathbf{g}'\mathbf{E}^{-1}\mathbf{f}} \tag{5.91}$$

[cf. (2.59) and (5.51)] was used repeatedly in Sec. 5.4 to find the inverse by rank annihilation. When $1+\mathbf{g}'\mathbf{E}^{-1}\mathbf{f}=0$, it follows [1] that $\mathbf{E}^{-1}$ is a g-inverse of $\mathbf{E}+\mathbf{f}\mathbf{g}'$, as is easily verified.

We express the $n\times n$ matrix $\mathbf{A}$ as in (5.53), and we write

$$\mathbf{A}=\mathbf{I}+(\mathbf{A}-\mathbf{I})\begin{pmatrix}\mathbf{I}_a & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}+(\mathbf{A}-\mathbf{I})\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{n-a}\end{pmatrix} \tag{5.92}$$

with

$$\mathbf{E}_a=\mathbf{I}+(\mathbf{A}-\mathbf{I})\begin{pmatrix}\mathbf{I}_a & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \tag{5.93a}$$

$$=\mathbf{I}+\begin{pmatrix}\mathbf{E}-\mathbf{I}_a & \mathbf{F}\\ \mathbf{G} & \mathbf{H}-\mathbf{I}_{n-a}\end{pmatrix}\begin{pmatrix}\mathbf{I}_a & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \tag{5.93b}$$

$$=\begin{pmatrix}\mathbf{E} & \mathbf{0}\\ \mathbf{G} & \mathbf{I}_{n-a}\end{pmatrix}, \tag{5.93c}$$

where $\mathbf{A}$ is partitioned as usual [see e.g. (2.2)], with $\mathbf{E}$ $a\times a$. Then [1, p. 3]

$$|\mathbf{E}_a|=0, \qquad \mathrm{r}(\mathbf{A})<a<n-1. \tag{5.94}$$

To prove (5.94) we note that

$$\mathrm{r}(\mathbf{E}_a)=\mathrm{r}(\mathbf{E})+n-a \tag{5.95a}$$

$$\leqslant \mathrm{r}(\mathbf{A})+n-a \tag{5.95b}$$

$$<a+n-a=n. \tag{5.95c}$$

If $\mathrm{r}(\mathbf{A})=n-1$ and $\mathbf{E}_j$ is nonsingular for all j such that $1\leqslant j\leqslant n-1$, then $\mathbf{E}_{n-1}^{-1}=\mathbf{A}^-$. More generally, let

$$\mathbf{A}=\mathbf{D}+\mathbf{F}_1\mathbf{G}_1'+\mathbf{F}_2\mathbf{G}_2', \tag{5.96}$$

where $\mathbf{F}_1$ and $\mathbf{G}_1$ are $n\times r$, $\mathbf{F}_2$ and $\mathbf{G}_2$ are $n\times(n-r)$, and $\mathrm{r}(\mathbf{A})=r$. If $\mathbf{D}+\mathbf{F}_1\mathbf{G}_1'$ $(=\mathbf{E}_r$ say) is nonsingular, then

$$(\mathbf{D}+\mathbf{F}_1\mathbf{G}_1')^{-1}=\mathbf{A}^{-}. \tag{5.97}$$

To prove this, consider

$$\mathbf{M}=\begin{pmatrix} \mathbf{D}+\mathbf{F}_1\mathbf{G}_1' & \mathbf{F}_2 \\ \mathbf{G}_2' & -\mathbf{I}_{n-r} \end{pmatrix}. \tag{5.98}$$

Then

$$\mathrm{r}(\mathbf{M})=n-r+\mathrm{r}(\mathbf{D}+\mathbf{F}_1\mathbf{G}_1'+\mathbf{F}_2\mathbf{G}_2') \tag{5.99a}$$

$$=n-r+\mathrm{r}(\mathbf{A})=n \tag{5.99b}$$

$$=\mathrm{r}(\mathbf{D}+\mathbf{F}_1\mathbf{G}_1')+\mathrm{r}\big(-\mathbf{I}_{n-r}-\mathbf{G}_2'(\mathbf{D}+\mathbf{F}_1\mathbf{G}_1')^{-1}\mathbf{F}_2\big) \tag{5.99c}$$

$$=n+\mathrm{r}\big(\mathbf{I}_{n-r}+\mathbf{G}_2'(\mathbf{D}+\mathbf{F}_1\mathbf{G}_1')^{-1}\mathbf{F}_2\big). \tag{5.99d}$$

Thus

$$\mathbf{I}_{n-r}+\mathbf{G}_2'(\mathbf{D}+\mathbf{F}_1\mathbf{G}_1')^{-1}\mathbf{F}_2=\mathbf{0}. \tag{5.100}$$

Now let $\mathbf{E}_r=\mathbf{D}+\mathbf{F}_1\mathbf{G}_1'$. Then

$$\mathbf{A}\mathbf{E}_r^{-1}\mathbf{A}=(\mathbf{E}_r+\mathbf{F}_2\mathbf{G}_2')\mathbf{E}_r^{-1}(\mathbf{E}_r+\mathbf{F}_2\mathbf{G}_2') \tag{5.101a}$$

$$=\big(\mathbf{I}_r+\mathbf{F}_2\mathbf{G}_2'\mathbf{E}_r^{-1}\big)(\mathbf{E}_r+\mathbf{F}_2\mathbf{G}_2') \tag{5.101b}$$

$$=\mathbf{E}_r+\mathbf{F}_2\mathbf{G}_2'+\mathbf{F}_2\mathbf{G}_2'+\mathbf{F}_2\mathbf{G}_2'\mathbf{E}_r^{-1}\mathbf{F}_2\mathbf{G}_2'. \tag{5.101c}$$

It follows that

$$\mathbf{A}\mathbf{E}_r^{-1}\mathbf{A}=\mathbf{A} \quad\Leftrightarrow\quad \mathbf{F}_2\big(\mathbf{I}+\mathbf{G}_2'\mathbf{E}_r^{-1}\mathbf{F}_2\big)\mathbf{G}_2'=\mathbf{0}, \tag{5.102}$$

which is implied by (5.100). Hence the proof is complete.

Thus if

$$\mathbf{A}=\mathbf{D}+\sum_{i=1}^{r}\mathbf{f}_i\mathbf{g}_i'+\sum_{i=r+1}^{n}\mathbf{f}_i\mathbf{g}_i', \tag{5.103}$$

so that $\mathbf{D}+\Sigma_{i=1}^{r}\mathbf{f}_i\mathbf{g}_i'$ is nonsingular and $r=\mathrm{r}(\mathbf{A})$, then

$$\left(\mathbf{D}+\sum_{i=1}^{r}\mathbf{f}_i\mathbf{g}_i'\right)^{-1}=\mathbf{A}^{-}. \tag{5.104}$$

If $\mathbf{D}=\mathbf{I}$, $\mathbf{f}_i=(\mathbf{A}-\mathbf{I})\mathbf{e}_i$, $\mathbf{g}_i=\mathbf{e}_i$, then (5.103) may not be possible. For example, suppose

$$\mathbf{A}=\begin{bmatrix}2&1&2\\2&1&2\\3&0&3\end{bmatrix}\qquad(\text{rank}=2) \tag{5.105}$$

$$=\mathbf{I}+\begin{bmatrix}1\\2\\3\end{bmatrix}\mathbf{e}_1'+\begin{bmatrix}1\\0\\0\end{bmatrix}\mathbf{e}_2'+\begin{bmatrix}2\\2\\2\end{bmatrix}\mathbf{e}_3'. \tag{5.106}$$

We may write

$$\mathbf{E}_1=\begin{bmatrix}2&0&0\\2&1&0\\3&0&1\end{bmatrix} \tag{5.107}$$

and its inverse

$$\mathbf{E}_1^{-1}=\mathbf{I}-\begin{bmatrix}1\\2\\3\end{bmatrix}\mathbf{e}_1'/2 \tag{5.108}$$

$$=\begin{bmatrix}\frac{1}{2}&0&0\\-1&1&0\\-\frac{3}{2}&0&1\end{bmatrix} \tag{5.109}$$

using (5.91). Then $\mathbf{g}_i'\mathbf{E}_1^{-1}\mathbf{f}_i=-1$, $i=2,3$, since

$$\mathbf{e}_2'\mathbf{E}_1^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_2=(-1,1,0)\begin{bmatrix}1\\0\\0\end{bmatrix}=-1 \tag{5.110}$$

and

$$\mathbf{e}_3'\mathbf{E}_1^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_3=\left(-\tfrac{3}{2},0,1\right)\begin{bmatrix}2\\2\\2\end{bmatrix}=-1. \tag{5.111}$$

Moreover,

$$\mathbf{e}_2'\mathbf{E}_1^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_3=(-1,1,0)\begin{bmatrix}2\\2\\2\end{bmatrix}=0 \tag{5.112}$$

and

$$\mathbf{e}_3'\mathbf{E}_1^{-1}(\mathbf{A}-\mathbf{I})\mathbf{e}_2=\left(-\tfrac{3}{2},0,1\right)\begin{bmatrix}1\\0\\0\end{bmatrix}=-\tfrac{3}{2}\neq 0. \tag{5.113}$$

Hence $\mathbf{E}_1^{-1}\neq\mathbf{A}^{-}$. In fact,

$$\mathbf{A}\mathbf{E}_1^{-1}\mathbf{A}=\begin{bmatrix}2&1&2\\2&1&2\\3&0&3\end{bmatrix}\begin{bmatrix}\frac{1}{2}&0&0\\-1&1&0\\-\frac{3}{2}&0&1\end{bmatrix}\mathbf{A} \tag{5.114}$$

and so

$$\mathbf{A}\mathbf{E}_1^{-1}\mathbf{A}=\begin{bmatrix}-3&1&2\\-3&1&2\\-3&0&3\end{bmatrix}\begin{bmatrix}2&1&2\\2&1&2\\3&0&3\end{bmatrix}$$

$$=\begin{bmatrix}2&-2&2\\2&-2&2\\3&-3&3\end{bmatrix}\neq\mathbf{A}=\begin{bmatrix}2&1&2\\2&1&2\\3&0&3\end{bmatrix} \tag{5.115}$$

[cf. (5.105)].

Since the matrix $\mathbf{A}$ given by (5.105) is 3×3 and has rank 2, it would follow that $\mathbf{E}_2^{-1}$ is a g-inverse of $\mathbf{A}$ if $\mathbf{E}_2$ were nonsingular: cf. remarks just above (5.96). In this example, however,

$$\mathbf{E}_2=\begin{bmatrix}2&1&0\\2&1&0\\3&0&1\end{bmatrix} \tag{5.116}$$

has rank 2.

VI. STATISTICAL APPLICATIONS OF SCHUR COMPLEMENTS

The Schur complement arises in a number of different areas of mathematical statistics. As observed by Cottle [18, 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

$$\mathbf{x}=\begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix} \tag{6.1}$$

follow a p-variate normal distribution with mean vector

$$\boldsymbol{\mu}=\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix} \tag{6.2}$$

and covariance matrix

$$\boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}, \tag{6.3}$$

where $\boldsymbol{\Sigma}_{22}$ is positive definite. Then the conditional distribution of $\mathbf{x}_1$ given $\mathbf{x}_2$ is multivariate normal with mean vector

$$\boldsymbol{\nu}_1=\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2) \tag{6.4}$$

and covariance matrix the Schur complement of $\boldsymbol{\Sigma}_{22}$ in $\boldsymbol{\Sigma}$,

$$(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22})=\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}. \tag{6.5}$$

To prove this result we note first that the joint distribution of

$$\begin{pmatrix}\mathbf{x}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{x}_2\\ \mathbf{x}_2\end{pmatrix} \tag{6.6}$$

is multivariate normal with mean vector

$$\begin{pmatrix}\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2\\ \boldsymbol{\mu}_2\end{pmatrix} \tag{6.7}$$

and covariance matrix

$$\begin{pmatrix} (\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \tag{6.8}$$

[cf. (2.26)]. Hence $\mathbf{x}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{x}_2$ is distributed independently of $\mathbf{x}_2$, and so its conditional distribution given $\mathbf{x}_2$ is the same as its unconditional distribution. Thus $\mathbf{x}_1$ given $\mathbf{x}_2$ is multivariate normal with mean vector (6.4) and covariance matrix (6.5).

Consider now the density function of the multivariate normal distribution

$$\phi(\mathbf{x}) = (2\pi)^{-\frac{1}{2}p}|\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}, \tag{6.9}$$

[3, p. 17]. Then the above result concerning the conditional distribution of $\mathbf{x}_1$ given $\mathbf{x}_2$ yields

$$\phi(\mathbf{x}) = \phi(\mathbf{x}_1|\mathbf{x}_2)\phi(\mathbf{x}_2); \tag{6.10}$$

thus

$$|\boldsymbol{\Sigma}| = |(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22})|\cdot|\boldsymbol{\Sigma}_{22}| \tag{6.11}$$

[cf. (2.6)], and

$$(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = (\mathbf{x}_1-\boldsymbol{\nu}_1)'(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22})^{-1}(\mathbf{x}_1-\boldsymbol{\nu}_1) + (\mathbf{x}_2-\boldsymbol{\mu}_2)'\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2). \tag{6.12}$$

To verify (6.12) directly we use (2.40) to write

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \mathbf{I} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} \end{pmatrix}(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22})^{-1}\left(\mathbf{I}, -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\right) + \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}. \tag{6.13}$$

Substituting (6.13) into the left-hand side of (6.12) yields the right-hand side directly, since

$$\left(\mathbf{I}, -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\right)(\mathbf{x}-\boldsymbol{\mu}) = \mathbf{x}_1 - \boldsymbol{\nu}_1. \tag{6.14}$$

When $\boldsymbol{\Sigma}_{22}$ is positive semidefinite and singular, the covariance matrix of $\boldsymbol{\Sigma}$ is also singular (Corollary 4.5), and so $\mathbf{x}$ does not have a density function. Using generalized inverses, however, we may evaluate [58, pp. 522–523] the joint distribution of

$$\begin{pmatrix} \mathbf{x}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-}\mathbf{x}_2 \\ \mathbf{x}_2 \end{pmatrix} \tag{6.15}$$

[cf. (6.6)] as multivariate normal with mean vector

$$\begin{pmatrix} \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-}\boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_2 \end{pmatrix} \tag{6.16}$$

and covariance matrix (6.8), where

$$(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22}) = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-}\boldsymbol{\Sigma}_{21} \tag{6.17}$$

is the generalized Schur complement of $\boldsymbol{\Sigma}_{22}$ in $\boldsymbol{\Sigma}$ [cf. (4.2)]. From (4.52) we see that (6.17) is unique for all choices of g-inverse $\boldsymbol{\Sigma}_{22}^{-}$. It should be noted that

$$\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{22}^{-}\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{12} \tag{6.18}$$

is needed to establish that the off-diagonal blocks in (6.8) are still $\mathbf{0}$. The equation (6.18) is equivalent to

$$\mathrm{r}(\boldsymbol{\Sigma}_{12}, \boldsymbol{\Sigma}_{22}) = \mathrm{r}(\boldsymbol{\Sigma}_{22}), \tag{6.19}$$

in view of Lemma 4.1, and (6.19) holds because of the nonnegative definiteness of $\boldsymbol{\Sigma}$. It is interesting to note that (6.19) is just the condition for consistency of

$$\mathbf{A}\boldsymbol{\Sigma}_{22} = \boldsymbol{\Sigma}_{12}, \tag{6.20}$$

which is analogous to the "normal equations" in regression analysis.

It follows at once that the conditional distribution of $\mathbf{x}_1$ given $\mathbf{x}_2$ is multivariate normal with mean vector

$$\boldsymbol{\nu}_1 = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \tag{6.21}$$

[cf. (6.4)] and covariance matrix (6.17). The mean vector (6.21) is unique provided $\mathbf{x}_2 - \boldsymbol{\mu}_2$ lies in the column space of $\boldsymbol{\Sigma}_{22}$ (with probability 1), in view of (6.18). This is assured by the distribution of $\mathbf{x}_2 - \boldsymbol{\mu}_2$ being multivariate normal with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{22}$.

Cottle [18, p. 195] gives an interesting interpretation of the "quotient property" for the multivariate normal distribution. (See also [3, p. 33].) Let the random vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}. \tag{6.22}$$

Suppose that we have the conditional distribution of $\mathbf{x}_1$ and $\mathbf{x}_2$ given $\mathbf{x}_3$. How do we find the conditional distribution of $\mathbf{x}_1$ given $\mathbf{x}_2$ and $\mathbf{x}_3$? Let us partition the covariance matrix of $\mathbf{x}$ as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{13} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{23} \\ \mathbf{\Sigma}_{31} & \mathbf{\Sigma}_{32} & \mathbf{\Sigma}_{33} \end{bmatrix}, \tag{6.23}$$

and write

$$\mathbf{\Sigma}_{2\&3} = \begin{pmatrix} \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{23} \\ \mathbf{\Sigma}_{32} & \mathbf{\Sigma}_{33} \end{pmatrix}. \tag{6.24}$$

Then (3.17) yields

$$(\mathbf{\Sigma}/\mathbf{\Sigma}_{2\&3}) = ((\mathbf{\Sigma}/\mathbf{\Sigma}_{33})/(\mathbf{\Sigma}_{2\&3}/\mathbf{\Sigma}_{33})). \tag{6.25}$$

Thus the conditional distribution of $\mathbf{x}_1$ given $\mathbf{x}_2$ and $\mathbf{x}_3$ is the conditional distribution of

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \Big| \mathbf{x}_3 \qquad \text{given} \qquad (\mathbf{x}_2 | \mathbf{x}_3).$$

In other words, we may condition sequentially.

6.2. *Partial correlation coefficients*

In Sec. 6.1 we saw that $(\mathbf{\Sigma}/\mathbf{\Sigma}_{22})$, the Schur complement of $\mathbf{\Sigma}_{22}$ in the covariance matrix

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}, \tag{6.26}$$

is also a covariance matrix. Anderson [3, p. 29] defines the elements of $(\mathbf{\Sigma}/\mathbf{\Sigma}_{22})$ to be *partial covariances.* Writing

$$(\mathbf{\Sigma}/\mathbf{\Sigma}_{22})=\{\sigma_{ij}^{(2)}\}, \tag{6.27}$$

we may define the *partial correlation coefficient* as

$$\rho_{ij}^{(2)}=\frac{\sigma_{ij}^{(2)}}{\left(\sigma_{ii}^{(2)}\sigma_{jj}^{(2)}\right)^{1/2}}, \tag{6.28}$$

provided $\sigma_{ii}^{(2)}>0$ for all i (which is assured when $\mathbf{\Sigma}$ is positive definite).

The diagonal elements of the covariance matrix $\mathbf{\Sigma}$ are the variances of the components of the underlying random vector $\mathbf{x}$. When these variances are all positive we may form the correlation matrix of $\mathbf{x}$ as

$$\mathbf{R}=\mathbf{\Delta}^{-1}\mathbf{\Sigma}\mathbf{\Delta}^{-1}=\begin{pmatrix}\mathbf{R}_{11} & \mathbf{R}_{12}\\ \mathbf{R}_{21} & \mathbf{R}_{22}\end{pmatrix}, \tag{6.29}$$

where

$$\mathbf{\Delta}=\operatorname{diag}\left(\sigma_{ii}^{1/2}\right)=\begin{pmatrix}\mathbf{\Delta}_1 & 0\\ 0 & \mathbf{\Delta}_2\end{pmatrix}, \tag{6.30}$$

say, is the diagonal matrix of standard deviations. If the Schur complement

$$(\mathbf{R}/\mathbf{R}_{22})=\{r_{ij}^{(2)}\}, \tag{6.31}$$

then

$$\rho_{ij}^{(2)}=\frac{r_{ij}^{(2)}}{\left(r_{ii}^{(2)}r_{jj}^{(2)}\right)^{1/2}}, \tag{6.32}$$

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

$$\begin{aligned}\mathbf{R}_{11}-\mathbf{R}_{12}\mathbf{R}_{22}^{-}\mathbf{R}_{21}&=\mathbf{\Delta}_1^{-1}\left(\mathbf{\Sigma}_{11}-\mathbf{\Sigma}_{12}\mathbf{\Delta}_2^{-1}\left(\mathbf{\Delta}_2^{-1}\mathbf{\Sigma}_{22}\mathbf{\Delta}_2^{-1}\right)^{-}\mathbf{\Delta}_2^{-1}\mathbf{\Sigma}_{21}\right)\mathbf{\Delta}_1^{-1}\\ &=\mathbf{\Delta}_1^{-1}(\mathbf{\Sigma}/\mathbf{\Sigma}_{22})\mathbf{\Delta}_1^{-1}.\end{aligned} \tag{6.33}$$

We may exploit the quotient property (6.25) to obtain a recursion formula for partial correlation coefficients [3, p. 34]. Partition the random vector

$$\mathbf{x}=\begin{bmatrix}\mathbf{x}_1\\ x_2\\ \mathbf{x}_3\end{bmatrix}, \tag{6.34}$$

where x_2 is a scalar, and write

$$\mathbf{R}=\begin{bmatrix}\mathbf{R}_{11} & \boldsymbol{\rho}_{12} & \mathbf{R}_{13}\\ \boldsymbol{\rho}_{12}' & 1 & \boldsymbol{\rho}_{23}'\\ \mathbf{R}_{13}' & \boldsymbol{\rho}_{23} & \mathbf{R}_{33}\end{bmatrix}. \tag{6.35}$$

Then, using (6.25), we obtain

$$\left\{\rho_{ij}^{(2\&3)}\right\}=\mathbf{R}_{11}-\mathbf{R}_{13}\mathbf{R}_{33}^{-}\mathbf{R}_{13}' - \frac{(\boldsymbol{\rho}_{12}-\mathbf{R}_{13}\mathbf{R}_{33}^{-}\boldsymbol{\rho}_{23})(\boldsymbol{\rho}_{12}-\mathbf{R}_{13}\mathbf{R}_{33}^{-}\boldsymbol{\rho}_{23})'}{1-\boldsymbol{\rho}_{23}'\mathbf{R}_{33}^{-}\boldsymbol{\rho}_{23}} \tag{6.36}$$

$$=\left\{r_{ij}^{(3)}-\frac{r_{i2}^{(3)}r_{j2}^{(3)}}{r_{22}^{(3)}}\right\}. \tag{6.37}$$

Hence

$$\rho_{ij}^{(2\&3)}=\frac{\rho_{ij}^{(3)}-\rho_{i2}^{(3)}\rho_{j2}^{(3)}}{\left(1-\left[\rho_{i2}^{(3)}\right]^2\right)^{1/2}\left(1-\left[\rho_{j2}^{(3)}\right]^2\right)^{1/2}}; \tag{6.38}$$

cf. [3, (34), p. 34].

Now suppose that $\mathbf{x}_1$ in (6.34) is also a scalar, and partition

$$\mathbf{R}=\left[\begin{array}{cc:c}1 & \rho_{12} & \boldsymbol{\rho}_{13}'\\ \rho_{12} & 1 & \boldsymbol{\rho}_{23}'\\ \hdashline \boldsymbol{\rho}_{13} & \boldsymbol{\rho}_{23} & \mathbf{R}_3\end{array}\right]. \tag{6.39}$$

Then the Schur complement

$$(\mathbf{R}/\mathbf{R}_3)=\begin{pmatrix}1-\boldsymbol{\rho}_{13}'\mathbf{R}_3^{-}\boldsymbol{\rho}_{13} & \rho_{12}-\boldsymbol{\rho}_{13}'\mathbf{R}_3^{-}\boldsymbol{\rho}_{23}\\ \rho_{12}-\boldsymbol{\rho}_{13}'\mathbf{R}_3^{-}\boldsymbol{\rho}_{23} & 1-\boldsymbol{\rho}_{23}'\mathbf{R}_3^{-}\boldsymbol{\rho}_{23}\end{pmatrix}, \tag{6.40}$$

and so

$$\rho_{12}^{(3)} = \frac{\rho_{12} - \rho_{13}' \mathbf{R}_3^- \rho_{23}}{\left(1 - \rho_{13}' \mathbf{R}_3^- \rho_{13}\right)^{1/2} \left(1 - \rho_{23}' \mathbf{R}_3^- \rho_{23}\right)^{1/2}}. \tag{6.41}$$

When $\mathbf{R}$ is nonsingular, we may obtain an alternate formula for $\rho_{12}^{(3)}$ using $\mathbf{R}^{-1}$. From (2.40) we may write

$$\{\rho^{ij}\} = \mathbf{R}^{-1} = \begin{pmatrix} (\mathbf{R}/\mathbf{R}_3)^{-1} & \cdot \\ \cdot & \cdot \end{pmatrix}, \tag{6.42}$$

and

$$(\mathbf{R}/\mathbf{R}_3)^{-1} = \frac{1}{|(\mathbf{R}/\mathbf{R}_3)|} \begin{pmatrix} 1 - \rho_{23}' \mathbf{R}_3^{-1} \rho_{23} & -\rho_{12} + \rho_{13}' \mathbf{R}_3^{-1} \rho_{23} \\ -\rho_{12} + \rho_{13}' \mathbf{R}_3^{-1} \rho_{23} & 1 - \rho_{13}' \mathbf{R}_3^{-1} \rho_{13} \end{pmatrix}. \tag{6.43}$$

Hence

$$\rho_{12}^{(3)} = \frac{-\rho^{12}}{\left(\rho^{11}\rho^{22}\right)^{1/2}}, \tag{6.44}$$

the negative of the corresponding correlation coefficient in $\mathbf{R}^{-1}$ (note that the minus sign has been dropped in (4g.2.8) in [58, 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:

$$\mathbf{R} = (1-\rho)\mathbf{I}_n + \rho \mathbf{e}\mathbf{e}', \tag{6.45}$$

which arises, for example, in the one-way random-effects analysis of variance (see e.g. [65, p. 225]). Consider the model

$$y_{ij} = \mu + a_i + u_{ij}, \qquad j = 1, \ldots, n_i, \quad i = 1, \ldots, k, \tag{6.46}$$

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

$$\mathcal{V}(a_i)=\sigma_a^2, \qquad i=1,\ldots,k \tag{6.48a}$$

$$\mathcal{V}(u_{ij})=\sigma^2, \qquad j=1,\ldots,n_i, \quad i=1,\ldots,k. \tag{6.48b}$$

Let $\mathbf{y}_i=\{y_{ij}\}_{j=1,\ldots,n_i}$ and $\mathbf{y}=\{\mathbf{y}_i\}_{i=1,\ldots,k}$. Then the covariance matrix of $\mathbf{y}$ is

$$\operatorname{diag}\left(\sigma^2\mathbf{I}_{n_i}+\sigma_a^2\mathbf{e}^{(n_i)}\mathbf{e}^{(n_i)'}\right)_{i=1,\ldots,k}, \tag{6.49}$$

where $\mathbf{e}^{(n_i)}$ is the $n_i\times 1$ vector of ones. The correlation matrix of $\mathbf{y}_i$ is, therefore, of the type (6.45), with

$$\rho=\frac{\sigma_a^2}{\sigma^2+\sigma_a^2}; \tag{6.50}$$

this is called the "intraclass" correlation between y_{ij} and $y_{ij'}$, where $j\neq j'$. If $n_1=n_2=\cdots=n_k=m$, then (6.49) becomes

$$\mathbf{I}_k\otimes\left(\sigma^2\mathbf{I}_m+\sigma_a^2\mathbf{e}^{(m)}\mathbf{e}^{(m)'}\right), \tag{6.51}$$

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, (6.9).

The determinant of the $n\times n$ matrix $\mathbf{R}$ given by (6.45) is

$$|\mathbf{R}|=|(1-\rho)\mathbf{I}_n+\rho\mathbf{e}\mathbf{e}'|=(1-\rho)^n\left|\mathbf{I}_n+\frac{\rho\mathbf{e}\mathbf{e}'}{1-\rho}\right| \tag{6.52a}$$

$$=(1-\rho)^n\left[1+\frac{\rho n}{1-\rho}\right] \tag{6.52b}$$

$$=(1-\rho)^{n-1}[1+\rho(n-1)], \tag{6.52c}$$

using (2.8). Thus $\mathbf{R}$ is nonsingular provided $\rho\neq 1$ or $-1/(n-1)$, and then we may compute the inverse $\mathbf{R}^{-1}$ using the formula (2.59), i.e.,

$$\mathbf{R}^{-1}=\left[(1-\rho)\mathbf{I}_n+\rho\mathbf{e}\mathbf{e}'\right]^{-1} \tag{6.53a}$$

$$=\frac{1}{1-\rho}\mathbf{I}_n-\frac{\rho}{(1-\rho)^2[1+\rho n/(1-\rho)]}\mathbf{e}\mathbf{e}' \tag{6.53b}$$

$$=\frac{1}{1-\rho}\left\{\mathbf{I}_n-\frac{\rho\mathbf{e}\mathbf{e}'}{1+\rho(n-1)}\right\}. \tag{6.53c}$$

We may find the characteristic roots by solving

$$|\mathbf{R}-\lambda\mathbf{I}_n|=|(1-\rho-\lambda)\mathbf{I}_n+\rho\mathbf{e}\mathbf{e}'|=0. \tag{6.54}$$

Using (2.8), we obtain

$$|\mathbf{R}-\lambda\mathbf{I}_n|=(1-\rho-\lambda)^{n-1}(1-\rho-\lambda+n\rho), \tag{6.55}$$

and so the characteristic roots are $1-\rho$ with multiplicity $n-1$ and $1+\rho(n-1)$ with multiplicity 1.

The matrix $\mathbf{R}$ 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 [63], is

$$\mathbf{R}=\begin{pmatrix}\mathbf{I}_m & \rho\mathbf{e}^{(m)}\mathbf{e}^{(n)'}\\ \rho\mathbf{e}^{(n)}\mathbf{e}^{(m)'} & \mathbf{I}_n\end{pmatrix}, \tag{6.57}$$

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

$$\mathbf{X}=\left[\begin{array}{c:c}\mathbf{e}^{(n)}\mathbf{e}_1^{(m)'} & \mathbf{I}_n\\ \vdots & \vdots\\ \mathbf{e}^{(n)}\mathbf{e}_m^{(m)'} & \mathbf{I}_n\end{array}\right], \tag{6.58}$$

(see e.g., [65, p. 100]), where $\mathbf{e}_i^{(k)}$ is the $k\times 1$ vector with 1 in the ith cell and 0 elsewhere. The matrix $\mathbf{X}$ is $mn\times(m+n)$, where m is the number of rows and n the number of columns in the experimental design. Hence

$$\mathbf{X}'\mathbf{X}=\begin{pmatrix} n\mathbf{I}_m & \mathbf{e}^{(m)}\mathbf{e}^{(n)\prime} \\ \mathbf{e}^{(n)}\mathbf{e}^{(m)\prime} & m\mathbf{I}_n \end{pmatrix}. \tag{6.59}$$

When the vector $\mathbf{y}$ of observations on the dependent variable has covariance matrix $\sigma^2\mathbf{I}$, then $\mathbf{X}'\mathbf{y}$ has covariance matrix $\sigma^2\mathbf{X}'\mathbf{X}$. The corresponding correlation matrix has the structure (6.57) with $\rho=(mn)^{-1/2}$, which is the maximum value of ρ such that (6.57) is nonnegative definite: cf. (6.70) below.

The determinant of (6.57) is

$$\begin{vmatrix} \mathbf{I}_m & \rho\mathbf{e}^{(m)}\mathbf{e}^{(n)\prime} \\ \rho\mathbf{e}^{(n)}\mathbf{e}^{(m)\prime} & \mathbf{I}_n \end{vmatrix}=|\mathbf{I}_n-m\rho^2\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}| \tag{6.60a}$$

$$=1-mn\rho^2, \tag{6.60b}$$

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

$$\mathrm{r}(\mathbf{X})=\mathrm{r}(\mathbf{X}'\mathbf{X})=m+\mathrm{r}\big(m\mathbf{I}_n-m\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}/n\big) \tag{6.61a}$$

$$=m+\mathrm{r}\big(\mathbf{I}_n-\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}/n\big). \tag{6.61b}$$

The matrix $\mathbf{C}_n=\mathbf{I}_n-\mathbf{e}\mathbf{e}'/n$ may be called the "centering matrix" [67]. The corresponding correlation matrix is the intraclass correlation matrix (6.45) with $\rho=-1/(n-1)$; this value of ρ is the lowest such 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

$$\mathrm{r}(\mathbf{X})=m+n-1. \tag{6.62}$$

To compute the inverse of

$$\mathbf{R}=\begin{pmatrix} \mathbf{I}_m & \rho\mathbf{e}^{(m)}\mathbf{e}^{(n)\prime} \\ \rho\mathbf{e}^{(n)}\mathbf{e}^{(m)\prime} & \mathbf{I}_n \end{pmatrix}, \qquad \rho^2\neq\frac{1}{mn}, \tag{6.63}$$

we use (2.41) and the Schur complements

$$\mathbf{S}=(\mathbf{R}/\mathbf{I}_m)=\mathbf{I}_n-m\rho^2\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}, \tag{6.64a}$$

$$\mathbf{T}=(\mathbf{R}/\mathbf{I}_n)=\mathbf{I}_m-n\rho^2\mathbf{e}^{(m)}\mathbf{e}^{(m)\prime}, \tag{6.64b}$$

and their inverses

$$\mathbf{S}^{-1}=\mathbf{I}_n+\frac{m\rho^2\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}}{1-mn\rho^2}, \tag{6.65a}$$

$$\mathbf{T}^{-1}=\mathbf{I}_m+\frac{n\rho^2\mathbf{e}^{(m)}\mathbf{e}^{(m)\prime}}{1-mn\rho^2}, \tag{6.65b}$$

which may be found using (2.59). Hence

$$\mathbf{R}^{-1}=\begin{bmatrix} \mathbf{I}_m+\dfrac{n\rho^2\mathbf{e}^{(m)}\mathbf{e}^{(m)\prime}}{1-mn\rho^2} & -\dfrac{\rho\mathbf{e}^{(m)}\mathbf{e}^{(n)\prime}}{1-mn\rho^2} \\ -\dfrac{\rho\mathbf{e}^{(n)}\mathbf{e}^{(m)\prime}}{1-mn\rho^2} & \mathbf{I}_n+\dfrac{m\rho^2\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}}{1-mn\rho^2} \end{bmatrix}. \tag{6.66}$$

To compute a generalized inverse of (6.59),

$$\mathbf{X}'\mathbf{X}=\begin{pmatrix} n\mathbf{I}_m & \mathbf{e}^{(m)}\mathbf{e}^{(n)\prime} \\ \mathbf{e}^{(n)}\mathbf{e}^{(m)\prime} & m\mathbf{I}_n \end{pmatrix}, \tag{6.67}$$

we may use (4.61), since (6.67) is nonnegative definite. The Schur complement

$$(\mathbf{X}'\mathbf{X}/n\mathbf{I}_m)=m\left(\mathbf{I}_n-\frac{\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}}{n}\right)=m\mathbf{C}_n, \tag{6.68a}$$

where $\mathbf{C}_n$ is the centering matrix [cf. (6.61)]. Since $\mathbf{C}_n$ is idempotent, it follows that $\mathbf{C}_n=\mathbf{C}_n^-$ and so $\mathbf{C}_n/m=(\mathbf{X}'\mathbf{X}/n\mathbf{I}_m)^-$. Hence

$$\begin{pmatrix} \mathbf{I}_m/n & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_n/m \end{pmatrix}=(\mathbf{X}'\mathbf{X})^-. \tag{6.68b}$$

The characteristic roots of (6.57) may be obtained from

$$|\mathbf{R}-\lambda\mathbf{I}_{m+n}|=\begin{vmatrix}(1-\lambda)\mathbf{I}_m & \rho\mathbf{e}^{(m)}\mathbf{e}^{(n)\prime}\\ \rho\mathbf{e}^{(n)}\mathbf{e}^{(m)\prime} & (1-\lambda)\mathbf{I}_n\end{vmatrix} \tag{6.69a}$$

$$=(1-\lambda)^m\left|(1-\lambda)\mathbf{I}_n-\frac{\rho^2 m\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}}{1-\lambda}\right| \tag{6.69b}$$

$$=(1-\lambda)^{m+n}\left|\mathbf{I}_n-\frac{\rho^2 m\mathbf{e}^{(n)}\mathbf{e}^{(n)\prime}}{(1-\lambda)^2}\right| \tag{6.69c}$$

$$=(1-\lambda)^{m+n-2}\left[(1-\lambda)^2-\rho^2 mn\right], \tag{6.69d}$$

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

$$-(mn)^{-1/2}<\rho<(mn)^{-1/2}. \tag{6.70}$$

We note that the correlation structure (6.57) is a special case of that considered in Theorem 4.14. In (4.191) set $\mathbf{A}=\mathbf{I}_m$, $\mathbf{B}=\mathbf{I}_n$; then $\mathbf{u}=m^{-1/2}\mathbf{e}^{(m)}$ is a normalized characteristic vector of $\mathbf{A}$ corresponding to a unit root. Similarly $\mathbf{v}=n^{-1/2}\mathbf{e}^{(n)}$ for $\mathbf{B}$. Hence put $\gamma=\rho(mn)^{1/2}$. Then the characteristic roots of (6.57) are 1 with multiplicity $m+n-2$ and the two roots of

$$\begin{Bmatrix}1 & \rho(mn)^{1/2}\\ \rho(mn)^{1/2} & 1\end{Bmatrix};$$

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 [58, p. 189] states are "fundamental to the theory of least squares."

Consider the general linear model with normality

$$\mathbf{y}\sim\mathcal{N}(\mathbf{X}\gamma,\sigma^2\mathbf{I}_n), \tag{6.71}$$

where $\mathbf{X}$ has rank r. The residual sum of squares

$$S_e = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y} = \mathbf{y}'\left[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\right]\mathbf{y} \tag{6.72}$$

is the Schur complement of $\mathbf{X}'\mathbf{X}$ in the matrix

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{y} \\ \mathbf{y}'\mathbf{X} & \mathbf{y}'\mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{X}' \\ \mathbf{y}' \end{pmatrix} (\mathbf{X}, \mathbf{y}). \tag{6.73}$$

Hence,

$$S_e \sim \sigma^2 \chi^2_{n-r}, \tag{6.74}$$

central chi-squared with $n-r$ degrees of freedom; cf [58, p. 189].

Now consider the multivariate general linear model with normality,

$$\mathbf{Y} = \mathbf{X}\mathbf{\Gamma} + \mathbf{U}, \tag{6.75}$$

where $\mathbf{Y}$ and $\mathbf{U}$ are $n \times p$ with rows following independent p-variate normal distributions with covariance matrix $\mathbf{\Sigma}$. The residual matrix of sums of squares and cross products

$$S_e = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y} = \mathbf{Y}'\left[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\right]\mathbf{Y} \tag{6.76}$$

is the Schur complement of $\mathbf{X}'\mathbf{X}$ in the matrix

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Y} \\ \mathbf{Y}'\mathbf{X} & \mathbf{Y}'\mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{X}' \\ \mathbf{Y}' \end{pmatrix} (\mathbf{X}, \mathbf{Y}), \tag{6.77}$$

and

$$S_e \sim \mathcal{W}_p(n-r, \mathbf{\Sigma}), \tag{6.78}$$

the p-variate central Wishart distribution with $n-r$ degrees of freedom and scale parameter $\mathbf{\Sigma}$ [58, p. 534]. When $p=1$, then $\mathbf{\Sigma} = \sigma^2$ and (6.78) reduces to (6.74).

To prove (6.78) we may use the following result [58, p. 536]. Let the random $n \times p$ matrix $\mathbf{Z}$ have independent rows, each normally distributed with covariance matrix $\mathbf{\Sigma}$. Suppose $\mathcal{E}(\mathbf{Z}) = \mathbf{\Omega}$. If $\mathbf{A}$ is a nonrandom symmetric

$n \times n$ matrix, then

$$\mathbf{W} = \mathbf{Z}'\mathbf{A}\mathbf{Z} \sim \mathcal{W}_p(f, \boldsymbol{\Sigma}) \tag{6.79}$$

if and only if $\mathbf{A} = \mathbf{A}^2$ and $\mathbf{A}\boldsymbol{\Omega} = \mathbf{0}$, and then $f = \mathrm{r}(\mathbf{A})$. Clearly $\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}' = \mathbf{M} = \mathbf{M}^2$ and $\mathrm{r}(\mathbf{M}) = n - \mathrm{r}(\mathbf{X}) = n - r$. Since $\boldsymbol{\Omega} = \mathbf{X}\boldsymbol{\Gamma}$, then $\mathbf{A}\boldsymbol{\Omega} = \mathbf{M}\mathbf{X}\boldsymbol{\Gamma} = \mathbf{0}$.

A somewhat different result concerning the Wishart distribution of a Schur complement may be obtained from (6.79) by setting $\mathbf{A} = \mathbf{I}$ and partitioning

$$\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2), \tag{6.80}$$

where $\mathbf{Z}_1$ is $n \times p_1$ and $\mathbf{Z}_2$ is $n \times p_2$, and $n > p = p_1 + p_2$. Then

$$\mathbf{W} = \begin{pmatrix} \mathbf{Z}_1'\mathbf{Z}_1 & \mathbf{Z}_1'\mathbf{Z}_2 \\ \mathbf{Z}_2'\mathbf{Z}_1 & \mathbf{Z}_2'\mathbf{Z}_2 \end{pmatrix}. \tag{6.81}$$

Partition $\boldsymbol{\Sigma}$ similarly so that

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \tag{6.82}$$

and suppose $\boldsymbol{\Omega} = \mathcal{E}(\mathbf{Z}) = \mathbf{0}$. Then

$$(\mathbf{W}/\mathbf{Z}_2'\mathbf{Z}_2) \sim \mathcal{W}_{p_1}(n - r_2, (\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22})), \tag{6.83}$$

where $r_2 = \mathrm{r}(\boldsymbol{\Sigma}_{22})$.

To prove (6.83) we consider

$$(\mathbf{W}/\mathbf{Z}_2'\mathbf{Z}_2) = \mathbf{Z}_1'[\mathbf{I} - \mathbf{Z}_2(\mathbf{Z}_2'\mathbf{Z}_2)^{-}\mathbf{Z}_2']\mathbf{Z}_1. \tag{6.84}$$

Moreover, given $\mathbf{Z}_2$ the rows of $\mathbf{Z}_1$ are independently normally distributed with covariance matrix $(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{22})$, while

$$\mathcal{E}(\mathbf{Z}_1|\mathbf{Z}_2) = \mathbf{Z}_2\boldsymbol{\Sigma}_{22}^{-}\boldsymbol{\Sigma}_{21}; \tag{6.85}$$

cf. (6.17) and (6.21). Then (6.83) follows at once, since $\mathbf{Z}_2$ has rank $r(\mathbf{\Sigma}_{22})$ with probability 1. Rao [58, p. 539] proves (6.83) when $\mathbf{\Sigma}_{22}$ is positive definite, and Ruohonen [62] establishes (6.83) using Moore-Penrose g-inverses.

Mitra [53] derives a result analogous to (6.83) for the matrix-variate beta distribution. Let $\mathbf{W}_1$ and $\mathbf{W}_2$ be independent $p\times p$ random matrices such that

$$\mathbf{W}_i \sim \mathcal{W}_p(k_i, \mathbf{\Sigma}), \qquad i=1,2, \tag{6.86}$$

and $k_1+k_2 \geqslant p$. Then $\mathbf{W}=\mathbf{W}_1+\mathbf{W}_2$ is positive definite with probability 1, and we may define

$$\mathbf{B}=\mathbf{W}^{-1/2}\mathbf{W}_1\mathbf{W}^{-1/2} \sim \mathcal{B}_p(k_1, k_2), \tag{6.87}$$

the p-variate beta distribution with k_1, k_2 degrees of freedom. If we partition

$$\mathbf{B}=\begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}, \tag{6.88}$$

where $\mathbf{B}_{11}$ is $p_1\times p_1$ and $\mathbf{B}_{22}$ is $p_2\times p_2$, then the Schur complement

$$(\mathbf{B}/\mathbf{B}_{22}) \sim \mathcal{B}_{p_1}(k_1, k_2-r_2), \tag{6.89}$$

the p_1-variate beta distribution with k_1, k_2-r_2 degrees of freedom.

6.5. *The Cramér-Rao inequality*

Let $\mathbf{x}_1,\ldots,\mathbf{x}_n$ be independently and identically distributed as the random vector $\mathbf{x}$, whose distribution depends on the unknown parameter vector $\boldsymbol{\theta}$. Then the score vector is defined as

$$\mathbf{s}=\frac{\partial \log l}{\partial \boldsymbol{\theta}}, \tag{6.90}$$

where l denotes the likelihood function of $\mathbf{x}_1,\ldots,\mathbf{x}_n$. Let $\mathbf{t}$ be an unbiased estimator for $\boldsymbol{\theta}$, i.e.,

$$\mathcal{E}(\mathbf{t})=\boldsymbol{\theta}. \tag{6.91}$$

Then the random vector

$$\mathbf{u}=\begin{pmatrix} \mathbf{s} \\ \mathbf{t} \end{pmatrix} \tag{6.92}$$

has, under certain regularity conditions, mean vector

$$\boldsymbol{\mu}=\begin{pmatrix} \mathbf{0} \\ \boldsymbol{\theta} \end{pmatrix} \tag{6.93}$$

and covariance matrix structure

$$\boldsymbol{\Sigma}=\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \mathbf{I}\\ \mathbf{I} & \boldsymbol{\Sigma}_{22}\end{pmatrix}. \tag{6.94}$$

If $\boldsymbol{\Sigma}_{11}$ is positive definite, then it follows from Theorem 3.2 that the Schur complement

$$(\boldsymbol{\Sigma}/\boldsymbol{\Sigma}_{11})=\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{11}^{-1} \tag{6.95}$$

is nonnegative definite. If, then, an unbiased statistic $\mathbf{t}_0$, say, can be found with covariance matrix $\boldsymbol{\Sigma}_{11}^{-1}=[\mathcal{V}(\partial\log l/\partial\boldsymbol{\theta})]^{-1}$, then $\mathbf{t}_0$ is the *minimum-variance unbiased* or *Markov* estimator of $\boldsymbol{\theta}$. This result is usually called the Cramér-Rao inequality, though Sverdrup [71, p. 72] and Savage [64, p. 238] claim that it is due to Fréchet [28].

To prove $\mathcal{E}(\mathbf{s})=\mathbf{0}$ and $\mathcal{C}(\mathbf{s},\mathbf{t})=\mathbf{I}$, we note first that

$$\int l\,d\mathbf{x}_1\cdots d\mathbf{x}_n=1, \tag{6.96}$$

which implies that, under appropriate regularity conditions,

$$\mathcal{E}(\mathbf{s})=\int\frac{\partial\log l}{\partial\boldsymbol{\theta}}\cdot l\,d\mathbf{z}=\int\frac{\partial l}{\partial\boldsymbol{\theta}}d\mathbf{z} \tag{6.97a}$$

$$=\frac{\partial}{\partial\boldsymbol{\theta}}\int l\,d\mathbf{z}=\mathbf{0}, \tag{6.97b}$$

where $d\mathbf{z}=d\mathbf{x}_1\,d\mathbf{x}_2\cdots d\mathbf{x}_n$. Moreover,

$$\mathcal{C}(\mathbf{s},\mathbf{t})=\mathcal{E}(\mathbf{s}\mathbf{t}')=\{\mathcal{E}(s_i t_j)\} \tag{6.98a}$$

$$=\left\{\int\frac{\partial\log l}{\partial\theta_i}\cdot t_j l\,d\mathbf{z}\right\} \tag{6.98b}$$

$$=\left\{\int\frac{\partial l}{\partial\theta_i}\cdot t_j\,d\mathbf{z}\right\} \tag{6.98c}$$

$$=\left\{\frac{\partial}{\partial\theta_i}\int t_j l\,d\mathbf{z}\right\} \tag{6.98d}$$

$$=\left\{\frac{\partial}{\partial\theta_i}\mathcal{E}(t_j)\right\}=\left\{\frac{\partial\theta_j}{\partial\theta_i}\right\}=\{\delta_{ij}\}=\mathbf{I}. \tag{6.98e}$$

In particular, if

$$\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\gamma}, \sigma^2\mathbf{I}), \tag{6.99}$$

where $\mathbf{X}$ has full column rank, then

$$l = (2\pi\sigma^2)^{-\frac{1}{2}n} \exp\{-\tfrac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\gamma})'(\mathbf{y}-\mathbf{X}\boldsymbol{\gamma})/\sigma^2\}, \tag{6.100}$$

$$\log l = -\tfrac{1}{2}n\log 2\pi - n\log\sigma - \tfrac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\gamma})'(\mathbf{y}-\mathbf{X}\boldsymbol{\gamma})/\sigma^2, \tag{6.101}$$

$$\mathbf{s} = \frac{\partial \log l}{\partial \boldsymbol{\gamma}} = \mathbf{X}'(\mathbf{y}-\mathbf{X}\boldsymbol{\gamma}) = -(\mathbf{X}'\mathbf{X}\boldsymbol{\gamma} - \mathbf{X}'\mathbf{y}), \tag{6.102}$$

$$\mathcal{V}(\mathbf{s}) = \mathbf{X}'\mathbf{X}/\sigma^2, \tag{6.103}$$

$$\boldsymbol{\Sigma}_{11}^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \tag{6.104}$$

The maximum-likelihood estimator of $\boldsymbol{\gamma}$ is

$$\hat{\boldsymbol{\gamma}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \tag{6.105}$$

and this has covariance matrix (6.104). Hence $\hat{\boldsymbol{\gamma}}$ is the minimum-variance unbiased or Markov estimator of $\boldsymbol{\gamma}$.

APPENDIX. CRACOVIANS

Following Banachiewicz [5, p. 45], we define the *Cracovian product* of an $m \times n$ matrix $\mathbf{A}$ and an $m \times p$ matrix $\mathbf{B}$ as

$$\mathbf{P} = \mathbf{A} \circ \mathbf{B} = \{p_{rs}\}, \tag{A.1}$$

where p_{rs} is the inner product of the rth column of $\mathbf{B}$ with the sth column of $\mathbf{A}$. Hence $p_{rs} = \mathbf{e}_s'\mathbf{A}'\mathbf{B}\mathbf{e}_r = \mathbf{e}_r'\mathbf{B}'\mathbf{A}\mathbf{e}_s$, so that

$$\mathbf{A} \circ \mathbf{B} = \mathbf{B}'\mathbf{A} \tag{A.2}$$

is a $p \times n$ matrix. It follows at once that

$$\mathbf{I}_m \circ \mathbf{B} = \mathbf{B}' \quad \text{and} \quad \mathbf{A} \circ \mathbf{I}_m = \mathbf{A}. \tag{A.3}$$

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):

$$\mathbf{I}\circ\mathbf{A}=\mathbf{IA}=\mathbf{A}', \tag{A.4}$$

since the middle form in (A.4) "*nicht vorhanden ist*" in ordinary matrix algebra. Thus

$$\mathbf{I}\circ\begin{pmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{pmatrix}=\begin{pmatrix}\mathbf{IE} & \mathbf{IG}\\ \mathbf{IF} & \mathbf{IH}\end{pmatrix}; \tag{A.5}$$

cf. [5, (2.7), p. 47]. Transposition of a Cracovian product reverses the order, for

$$\mathbf{I}(\mathbf{A}\circ\mathbf{B})=(\mathbf{A}\circ\mathbf{B})'=(\mathbf{B}'\mathbf{A})'=\mathbf{A}'\mathbf{B}=\mathbf{B}\circ\mathbf{A}. \tag{A.6}$$

When **A** is nonsingular the Cracovian inverse is the transpose of the usual inverse. To see this, write

$$\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}=\mathbf{A}\circ(\mathbf{A}^{-1})'=\mathbf{A}\circ\mathbf{IA}^{-1}, \tag{A.7}$$

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

$$\begin{aligned}(\mathbf{A}\circ\mathbf{B})\circ(\mathbf{IA}^{-1}\circ\mathbf{IB}^{-1})&=(\mathbf{B}'\mathbf{A})\circ\{(\mathbf{A}^{-1})'\circ(\mathbf{B}^{-1})'\}\\ &=(\mathbf{B}'\mathbf{A})\circ\{\mathbf{B}^{-1}(\mathbf{A}^{-1})'\}\\ &=\mathbf{A}^{-1}(\mathbf{B}^{-1})'\mathbf{B}'\mathbf{A}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}.\end{aligned} \tag{A.8}$$

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

I am also grateful to Professor Emilie Haynsworth, who suggested an extension of Theorem 4.14 and the comments at the end of Sec. V, as well as to Professor Hans Schneider, who drew my attention to the paper by Frobenius [29] and thus Theorem 2.3.

REFERENCES

1 M. Ahsanullah and M. Rahman, A singular inverse of a matrix by rank annihilation, *Canad. Math. Bull.* 16:1–4 (1973).

2 A. C. Aitken, *Determinants and Matrices*, Oliver & Boyd, Edinburgh, 1939.

3 T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

4 T. Ando, Commutativity formulas for generalized Schur complements, *Linear Algebra and Appl.* 27:173–186 (1979).

5 T. Banachiewicz, Zur Berechnung der Determinanten, wie auch der Inversen, und zur darauf basierten Auflösung der Systeme linearer Gleichungen, *Acta Astronom. Sér. C* 3:41–67 (1937).

6 M. S. Bartlett, An inverse matrix adjustment arising in discriminant analysis, *Ann. Math. Statist.* 22:107–111 (1951).

7 P. Bhimasankaram, On generalized inverses of partitioned matrices, *Sankhyā Ser. A* 33:311–314 (1971).

8 A. Bjerhammar, A generalized matrix algebra, *Kungl. Tekniska Högskolans Handlingar (Stockholm)*, 124:1–32 (1958).

9 A. Bjerhammar, *Theory of Errors and Generalized Matrix Inverses*, Elsevier Scientific, Amsterdam, 1973.

10 E. Bodewig, Comparison of some direct methods for computing determinants and inverse matrices, *Koninklijke Nederlandse Akademie van Wetenschappen, Proc. Section of Sciences* 50:49–57 (1947).

11 E. Bodewig, *Matrix Calculus*, North-Holland, Amsterdam, 1956; 2nd revised and enlarged ed., 1959.

12 H. Boerner, Issai Schur, in *Dictionary of Scientific Biography* (C. C. Gillispie, Ed.), Scribner's, New York, 1975, Vol. XII, p. 237.

13 F. Burns, D. Carlson, E. Haynsworth, and T. Markham, Generalized inverse formulas using the Schur complement, *SIAM J. Appl. Math.* 26:254–259 (1974).

14 D. Carlson, Matrix decompositions involving the Schur complement, *SIAM J. Appl. Math.* 28:577–587 (1975).

15 D. Carlson, E. Haynsworth, and T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, *SIAM J. Appl. Math.* 26:169–175 (1974).

16 A. L. Cauchy, Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opérées entre les variables qu'elles renferment. *J. École Polytech.* 10:29–112 (1812). [See also *Oeuvres Complètes*, vol. 2: pp. 1, 91–169.]

17 R. E. Cline and R. E. Funderlic, The rank of a difference of matrices and associated generalized inverses, *Linear Algebra and Appl.* 24:185–215 (1979).

18 R. W. Cottle, Manifestations of the Schur complement, *Linear Algebra and Appl.* 8:189–211 (1974). [See also *Rend. Sem. Mat. Fis. Milano* 45:31–40 (1975).]

19 D. E. Crabtree and E. V. Haynsworth, An identity for the Schur complement of a matrix. *Proc. Amer. Math. Soc.* 22:364–366 (1969).

20 J. Dias da Silva, Review of [26], *Math. Reviews* 51, #543 (1976).
21 W. J. Duncan, 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 (1944).
22 D. J. Edelblute, Matrix inversion by rank annihilation, *Math. Comp.* 20:149–151 (1966).
23 J. H. Ellenberg, The joint distribution of the standardized least squares residuals from a general linear regression, *J. Amer. Statist. Assoc.* 68:941–943 (1973).
24 D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra*, Freeman, San Francisco, 1963.
25 V. N. Faddeeva, *Computational Methods of Linear Algebra*, Dover, New York, 1959.
26 M. Fiedler, Eigenvalues of nonnegative symmetric matrices. *Linear Algebra and Appl.* 9:119–142 (1974).
27 R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary Matrices, and Some Applications to Dynamics and Differential Equations*, Cambridge U.P., 1938.
28 M. Fréchet, Sur l'extension de certaines évaluations, statistiques au cas de petits échantillons, *Rev. Inst. Internat. Statist.* 11:183–205 (1943).
29 F. G. Frobenius, Über Matrizen aus positiven Elementen, *S.-B. Königlich Preussischen Akad. Wiss. zu Berlin*, 1908, pp. 471–476; reprinted in Ferdinand Georg Frobenius, *Gesammelte Abhandlungen*, Springer, Berlin, 1968, vol. 3, pp. 404–409.
30 L. S. Goddard and H. Schneider, Pairs of matrices with a non-zero commutator, *Proc. Cambridge Philos. Soc.* 51:551–553 (1955).
31 Irwin Guttman, A Bayesian analogue of Paulson's lemma and its use in tolerance region construction when sampling from the multi-variate normal, *Ann. Inst. Statist. Math.* 23:67–76 (1971).
32 Louis Guttman, Enlargement methods for computing the inverse matrix, *Ann. Math. Statist.* 17:336–343 (1946).
33 D. J. Hartfiel, An extension of Haynsworth's determinant inequality, *Proc. Amer. Math. Soc.* 41:463–465 (1973).
34 R. E. Hartwig, Singular value decomposition and the Moore-Penrose inverse of bordered matrices, *SIAM J. Appl. Math.* 31:31–41 (1976).
35 R. E. Hartwig, Rank factorization and Moore-Penrose inversion. *J. Industrial Math. Soc.* 26:49–63 (1976).
36 E. V. Haynsworth, Determination of the inertia of a partitioned Hermitian matrix, *Linear Algebra and Appl.* 1:73–81 (1968).
37 E. V. Haynsworth, Reduction of a matrix using properties of the Schur complement, *Linear Algebra and Appl.* 3:23–29 (1970). [See also: On the Schur complement, Basle Mathematical Notes, BMN 20, 1968, 17 pp.]
38 E. V. Haynsworth, Applications of an inequality for the Schur complement, *Proc. Amer. Math. Soc.* 24:512–516 (1970).
39 I. N. Herstein and L. W. Small, An extension of a theorem of Schur, *Linear and Multilinear Algebra* 3:41–43 (1975).

40 H. Hotelling, Some new methods in matrix calculation, *Ann. Math. Statist.* 14:1–34 (1943).

41 A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.

42 C. G. J. Jacobi, 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, *J. Reine Angew. Math.* 12:1–69 (1834).

43 F. Jossa, Risoluzione progressiva di un sistema di equazioni lineari. Analogia con un problema meccanico, *Rend. Accad. Sci. Fis. Mat. Napoli Ser. 4* 10:346–352 (1940).

44 C. G. Khatri, A note on some results on a generalised inverse of a matrix, *J. Indian Statist. Assoc.* 7:38–45 (1969).

45 C. Lanczos, Linear systems in self-adjoint form, *Amer. Math. Monthly* 65:665–679 (1958).

46 M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber & Schmidt, Boston, 1964.

47 G. Marsaglia and G. P. H. Styan, Equalities and inequalities for ranks of matrices, *Linear and Multilinear Algebra* 2:269–292 (1974).

48 G. Marsaglia and G. P. H. Styan, Rank conditions for generalized inverses of partitioned matrices, *Sankhyā Ser. A* 36:437–442 (1974).

49 C. D. Meyer, Jr., Generalized inverses of block triangular matrices, *SIAM J. Appl. Math.* 19:741–750 (1970).

50 C. D. Meyer, Jr., Generalized inverses and ranks of block matrices, *SIAM J. Appl. Math.* 25:597–602 (1973).

51 G. A. Milliken and F. Akdeniz, A theorem on the difference of the generalized inverses of two nonnegative matrices, *Comm. Statist.—Theory and Methods* A6:73–79 (1977).

52 L. Mirsky, *An Introduction to Linear Algebra*, Oxford U.P., 1955.

53 S. K. Mitra, A density-free approach to the matrix variate beta distribution, *Sankhyā Ser. A* 32:81–88 (1970).

54 M. Newman, Matrix computations, in *Survey of Numerical Analysis* (J. Todd, Ed.), McGraw-Hill, New York, 1962, pp. 222–254.

55 A. Ostrowski, A new proof of Haynsworth's quotient formula for Schur complements, *Linear Algebra and Appl.* 4:389–392 (1971).

56 R. M. Pringle and A. A. Rayner, Expressions for generalized inverses of a bordered matrix with application to the theory of constrained linear models, *SIAM Rev.* 12:107–115 (1970).

57 C. R. Rao, A note on a generalized inverse of a matrix with applications to problems in mathematical statistics, *J. Roy. Statist. Soc. Ser. B* 24:152–158 (1962).

58 C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York, 1973.

59 C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, New York, 1971.

60 C. A. Rohde, Generalized inverses of partitioned matrices, *J. Soc. Industrial and Appl. Math.* 13:1033–1035 (1965).
61 M. Ruohonen, Generalization of a matrix inversion lemma, Technical Report No. 44, Institute for Applied Mathematics, Univ. of Turku, Turku, Finland, 1973.
62 M. Ruohonen, Proof of Wishartness of a residual matrix, Technical Report No. 69, Institute for Applied Mathematics, Univ. of Turku, Turku, Finland, 1976.
63 A. R. Sampson, Simple BAN estimators of correlations for certain multivariate normal models with known variances. *J. Amer. Statist. Assoc.*, 73:859–862 (1978).
64 L. J. Savage, *The Foundations of Statistics*, Dover, New York, 1972.
65 H. Scheffé, *The Analysis of Variance*, Wiley, New York, 1959.
66 J. [Issai] Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, *J. Reine Angew. Math.* 147:205–232 (1917).
67 G. E. Sharpe and G. P. H. Styan, Circuit duality and the general network inverse, *IEEE Trans. Circuit Theory* CT-12:22–27 (1965).
68 J. Sherman and W. J. Morrison, Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix (Abstract), *Ann. Math. Statist.* 20:621 (1949).
69 J. Sherman and W. J. Morrison, Adjustment of an inverse matrix corresponding to a change in one element of a given matrix, *Ann. Math. Statist.* 21:124–127 (1950).
70 G. P. H. Styan and F. Pukelsheim, On the nonnegative definiteness of the difference between two matrices (Abstract), *Inst. Math. Statist. Bull.* 7:364 (1978).
71 E. Sverdrup, *Basic Concepts of Statistical Inference*, Vol. II, North-Holland, Amsterdam, 1967.
72 F. V. Waugh, A note concerning Hotelling's method of inverting a partitioned matrix, *Ann. Math. Statist.* 16:216–217 (1945).
73 J. H. M. Wedderburn, *Lectures on Matrices*, Amer. Math. Soc. Colloquium Publications, Vol. XVII, 1934; republished, Dover, New York, 1964.
74 J. R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations*, Wiley, New York, 1968.
75 H. S. Wilf, Matrix inversion by the annihilation of rank. *J. Soc. Industrial and Appl. Math.* 7:149–151 (1959).
76 M. A. Woodbury, Inverting modified matrices, Memorandum Report 42, Statistical Research Group, Princeton, N.J., 1950.