Proc. First Tampere Sem. Linear Models (1983)
pp. 37—75, © Dept. of Math. Sci., Univ. Tampere, 1985

Schur Complements and Linear Statistical Models*

by

George P. H. Styan

McGill University, Montréal, Québec, Canada

CONTENTS

1. SCHUR COMPLEMENTS 39
   1.1 One Schur complement 39
   1.2 Two Schur complements 39
   1.3 Schur complements and matrix convexity 43
   1.4 Generalized Schur complements 44
   1.5 Generalized Schur complements and inertia 46
   1.6 Schur complements and statistics 48
2. CANONICAL CORRELATIONS AND THE GENERAL PARTITIONED LINEAR MODEL 49
   2.1 Canonical correlations: the number less than one and the number equal to one 49
   2.2 Canonical correlations in the general partitioned linear model 52
   2.3 Canonical correlations and testing the hypothesis that some parameters are zero 57
3. BROYDEN'S MATRIX PROBLEM AND AN ASSOCIATED ANALYSIS-OF-COVARIANCE MODEL 62
   3.1 Broyden's matrix problem and his solution 62
   3.2 An associated analysis-of-covariance model 64
   3.3 Nonnegative definiteness 65
   3.4 The special case when all the cells are filled 66
   3.5 The general case when at least one of the cells is empty 66
      3.5.1 A necessary condition for positive definiteness 67
      3.5.2 Necessary and sufficient conditions when the layout is connected 67
      3.5.3 A numerical example 70
      3.5.4 Necessary and sufficient conditions when the layout is not connected 71
4. ACKNOWLEDGEMENTS 72
   BIBLIOGRAPHY 72

---

* Based on an invited lecture series presented at The First International Tampere Seminar on Linear Statistical Models and their Applications, University of Tampere, Tampere, Finland, August 30 — September 2, 1983.

Issai Schur: 1875–1941

Photograph reproduced with permission of the publisher and the authors of ***Inequalities: Theory of Majorization and Its Applications*** by Albert W. Marshall and Ingram Olkin, pub. Academic Press, New York, 1979 (photograph appears on page 525).

1. SCHUR COMPLEMENTS

1.1 One Schur complement

If we partition the (possibly rectangular) matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \tag{1.1}$$

and if **E** is square and nonsingular, then

$$\mathbf{S} = \mathbf{H} - \mathbf{G}\mathbf{E}^{-1}\mathbf{F} = (\mathbf{A}/\mathbf{E}), \tag{1.2}$$

say, is said to be *the Schur complement of* **E** *in* **A**. The term »Schur complement» and the notation (**A**/**E**) were introduced by Haynsworth (1968). As mentioned in the survey paper by Cottle (1974), »these objects have undoubtedly been encountered from the time matrices were first used»; indeed Carlson (1984) indicates that the »idea» is due to Sylvester (1851), while in her detailed survey Ouellette (1981) cites Frobenius (1908). See also Brualdi and Schneider (1983). Bodewig (1959, Chapter 2) refers to the determinantal formula, obtained by Schur (1917, p. 217; 1973, p. 149)

$$\det \mathbf{A} = \det \mathbf{E} \cdot \det(\mathbf{A}/\mathbf{E}) \tag{1.3}$$

as »Frobenius-Schur's relation» — Issai Schur (1875–1941) was a student of Ferdinand Georg Frobenius (1849–1917), cf. e.g., Boerner (1975, p. 237). The formula (1.3) follows at once from the factorization

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{G}\mathbf{E}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & (\mathbf{A}/\mathbf{E}) \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{E}^{-1}\mathbf{F} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{1.4}$$

While (1.3) is, of course, only valid when **A** is square, in (1.4) the matrix **A** may be rectangular. It follows immediately that

$$\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{E}) + \operatorname{rank}(\mathbf{A}/\mathbf{E}), \tag{1.5}$$

which was first established by Guttman (1946).

Schur (1917) used the determinantal formula (1.3) to show that

$$\det\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \det(\mathbf{EH} - \mathbf{GF}) \tag{1.6}$$

when **E**, **F**, **G**, and **H** are all square and **EG** = **GE** (cf. Ouellette, 1981, Theorem 2.2).

In statistics »the multivariate normal distribution provides a magnificent example of how the Schur complement arises naturally» (Cottle, 1974, p. 192). Let the random vector

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \tag{1.7}$$

have covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{1.8}$$

where Σ_{11} is positive definite. (All Greek letters denoting matrices and vectors in this paper appear in light-face print.) Then the vector

$$\mathbf{x}_{2.1} = \mathbf{x}_2 - \Sigma_{21}\Sigma_{11}^{-1}\mathbf{x}_1 \tag{1.9}$$

of residuals after regressing $\mathbf{x}_2$ on $\mathbf{x}_1$ has covariance matrix

$$\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} = (\Sigma/\Sigma_{11}) \tag{1.10}$$

and is uncorrelated with the vector $\mathbf{x}_1$. When $\mathbf{x}$ follows a multivariate normal distribution then the vectors $\mathbf{x}_{2.1}$ and $\mathbf{x}_1$ are independently distributed and (Σ/Σ_{11}) is the covariance matrix of the conditional (also multivariate normal) distribution of $\mathbf{x}_2$ given $\mathbf{x}_1$, cf. e.g., Anderson (1984, Section 2.5).

When the matrix **A** in (1.4) is both square and nonsingular, then so also is the Schur complement (**A**/**E**), cf. (1.3), and

$$\begin{aligned} \mathbf{A}^{-1} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}^{-1} &= \begin{pmatrix} \mathbf{I} & -\mathbf{E}^{-1}\mathbf{F} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{E}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{A}/\mathbf{E})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{G}\mathbf{E}^{-1} & \mathbf{I} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{E}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{E}^{-1}\mathbf{F} \\ -\mathbf{I} \end{pmatrix} (\mathbf{A}/\mathbf{E})^{-1}(\mathbf{G}\mathbf{E}^{-1}, -\mathbf{I}) \\ &= \begin{pmatrix} \mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{F}(\mathbf{A}/\mathbf{E})^{-1}\mathbf{G}\mathbf{E}^{-1} & -\mathbf{E}^{-1}\mathbf{F}(\mathbf{A}/\mathbf{E})^{-1} \\ -(\mathbf{A}/\mathbf{E})^{-1}\mathbf{G}\mathbf{E}^{-1} & (\mathbf{A}/\mathbf{E})^{-1} \end{pmatrix}, \end{aligned} \tag{1.11}$$

which is due to Banachiewicz (1937a, 1937b), cf. Ouellette (1981, p. 201) and Henderson and Searle (1981, p. 55).

When the matrix **A** is both square and symmetric then **G** = **F**′ (prime denotes transpose throughout this paper and all matrices are real), and (1.4) becomes

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}' & \mathbf{H} \end{pmatrix} = \mathbf{U}' \begin{pmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & (\mathbf{A}/\mathbf{E}) \end{pmatrix} \mathbf{U}, \tag{1.12}$$

where

$$\mathbf{U} = \begin{pmatrix} \mathbf{I} & \mathbf{E}^{-1}\mathbf{F} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{1.13}$$

If follows directly from Sylvester's Law of Inertia (due to Sylvester, 1852; cf. Turnbull and Aitken, 1932, p. 99, and Mirsky, 1955, p. 377) that inertia is additive on the Schur complement (Haynsworth, 1968), in the sense that

$$\operatorname{In}\mathbf{A} = \operatorname{In}\mathbf{E} + \operatorname{In}(\mathbf{A}/\mathbf{E}), \tag{1.14}$$

where inertia is defined by the ordered triple

$$\operatorname{In}\mathbf{A} = \{\pi, \eta, \nu\}, \tag{1.15}$$

where π is the number of positive eigenvalues of $\mathbf{A}$, η is the number of negative eigenvalues of $\mathbf{A}$, and ν is the number of zero eigenvalues of $\mathbf{A}$. Thus $\pi + \eta = \operatorname{rank}(\mathbf{A})$, and $\nu = \nu(\mathbf{A})$, the nullity of $\mathbf{A}$. [The matrix $\mathbf{A}$ is real and symmetric so that all the eigenvalues are real, and rank equals the number of nonzero eigenvalues.]

When the submatrix $\mathbf{H}$ of the matrix $\mathbf{A}$ in (1.4) is square and nonsingular (instead of or in addition to the submatrix $\mathbf{E}$), then

$$\mathbf{T} = \mathbf{E} - \mathbf{F}\mathbf{H}^{-1}\mathbf{G} = (\mathbf{A}/\mathbf{H}) \tag{1.16}$$

is the Schur complement of $\mathbf{H}$ in $\mathbf{A}$. In parallel to (1.3), (1.5), (1.11), and (1.14) we obtain, therefore,

$$\det\mathbf{A} = \det\mathbf{H} \cdot \det(\mathbf{A}/\mathbf{H}), \tag{1.17}$$

$$\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{H}) + \operatorname{rank}(\mathbf{A}/\mathbf{H}), \tag{1.18}$$

$$\begin{aligned}\mathbf{A}^{-1} &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^{-1} \end{pmatrix} + \begin{pmatrix} -\mathbf{I} \\ \mathbf{H}^{-1}\mathbf{G} \end{pmatrix} (\mathbf{A}/\mathbf{H})^{-1}(-\mathbf{I}, \mathbf{F}\mathbf{H}^{-1}) \\ &= \begin{pmatrix} (\mathbf{A}/\mathbf{H})^{-1} & -(\mathbf{A}/\mathbf{H})^{-1}\mathbf{F}\mathbf{H}^{-1} \\ -\mathbf{H}^{-1}\mathbf{G}(\mathbf{A}/\mathbf{H})^{-1} & \mathbf{H}^{-1} + \mathbf{H}^{-1}\mathbf{G}(\mathbf{A}/\mathbf{H})^{-1}\mathbf{F}\mathbf{H}^{-1} \end{pmatrix},\end{aligned} \tag{1.19}$$

$$\operatorname{In}\mathbf{A} = \operatorname{In}\mathbf{H} + \operatorname{In}(\mathbf{A}/\mathbf{H}). \tag{1.20}$$

1.2 Two Schur complements

When *both* $\mathbf{E}$ *and* $\mathbf{H}$ are square and nonsingular, then we may combine (1.3) and (1.17) to yield (with $\mathbf{A}$ temporarily replaced by $\mathscr{A}$)

$$\det\mathscr{A} = \det\mathbf{E} \cdot \det(\mathscr{A}/\mathbf{E}) = \det\mathbf{H} \cdot \det(\mathscr{A}/\mathbf{H}), \tag{1.21}$$

from which with $\mathbf{E} = \lambda\mathbf{I}_m$, $\mathbf{F} = \mathbf{A}$, $\mathbf{G} = \mathbf{B}$, and $\mathbf{H} = \mathbf{I}_n$, we obtain:

$$\det(\lambda\mathbf{I}_m - \mathbf{AB}) = \lambda^{m-n} \cdot \det(\lambda\mathbf{I}_n - \mathbf{BA}), \tag{1.22}$$

and so the m eigenvalues of $\mathbf{AB}$ are equal to the n eigenvalues of $\mathbf{BA}$ plus $m - n$ zeros (assuming without loss of generality that $m \geq n$). Similarly

$$\det(\mathbf{I}_m + \mathbf{AB}) = \det(\mathbf{I}_n + \mathbf{BA}). \tag{1.23}$$

We may also combine (1.5) and (1.18) to obtain

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{E}) + \text{rank}(\mathbf{A}/\mathbf{E}) = \text{rank}(\mathbf{H}) + \text{rank}(\mathbf{A}/\mathbf{H}), \tag{1.24}$$

which yields, since both $\mathbf{E}$ and $\mathbf{H}$ are square and nonsingular,

$$\nu(\mathbf{A}) = \nu(\mathbf{A}/\mathbf{E}) = \nu(\mathbf{A}/\mathbf{H}), \tag{1.25}$$

where $\nu(\cdot)$ denotes nullity, cf. (1.15), and so the two Schur complements have the same nullity (they will, therefore, have the same rank if and only if they are of the same size).

Combining (1.11) and (1.19) yields (from the top left-hand corner of $\mathbf{A}^{-1}$)

$$\mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{F}(\mathbf{H} - \mathbf{GE}^{-1}\mathbf{F})^{-1}\mathbf{GE}^{-1} = (\mathbf{E} - \mathbf{FH}^{-1}\mathbf{G})^{-1}, \tag{1.26}$$

as noted (apparently for the first time) by Duncan (1944). A survey of the many special cases of (1.26) is given by Henderson and Searle (1981), as well as by Ouellette (1981); for example,

$$(\mathbf{E} + h\mathbf{e}_i\mathbf{e}_j')^{-1} = \mathbf{E}^{-1} - \frac{h\mathbf{E}^{-1}\mathbf{e}_i\mathbf{e}_j'\mathbf{E}^{-1}}{1 + h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{e}_i} \tag{1.27}$$

provided $h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{e}_i \neq -1$. In the formula (1.27), which was obtained by Sherman and Morrison (1950), the vector $\mathbf{e}_i$ has 1 in its ith position and zero everywhere else. When, therefore, a scalar h is added to the (i,j)th element of a nonsingular matrix $\mathbf{E}$, the new matrix is nonsingular if and only if the Schur complement $1 + h\mathbf{e}_j'\mathbf{E}^{-1}\mathbf{e}_i \neq 0$, i.e., h times the (j,i)th element of $\mathbf{E}^{-1}$ is not equal to -1. And then the inverse of the new matrix is the old inverse »corrected» as in formula (1.27) by a rank-one matrix.

If we combine (1.14) and (1.20) then we obtain

$$\text{In}\,\mathbf{A} = \text{In}\,\mathbf{E} + \text{In}(\mathbf{A}/\mathbf{E}) = \text{In}\,\mathbf{H} + \text{In}(\mathbf{A}/\mathbf{H}). \tag{1.28}$$

A special case of (1.28) is found by putting $\mathbf{E} = \mathbf{A}$, $\mathbf{F} = \mathbf{G} = \mathbf{I}$, and $\mathbf{H} = \mathbf{B}^{-1}$ (with both $\mathbf{A}$ and $\mathbf{B}$ nonsingular), and then

$$\text{In}\,\mathbf{A} + \text{In}(\mathbf{B}^{-1} - \mathbf{A}^{-1}) = \text{In}\,\mathbf{B} + \text{In}(\mathbf{A} - \mathbf{B}), \tag{1.29}$$

since $\mathbf{B}$ and $\mathbf{B}^{-1}$ have the same inertia (when $\mathbf{B}$ is nonsingular). Hence

$$\text{In}(\mathbf{B}^{-1} - \mathbf{A}^{-1}) = \text{In}(\mathbf{A} - \mathbf{B}) - [\text{In}\,\mathbf{A} - \text{In}\,\mathbf{B}]. \tag{1.30}$$

Thus $\mathbf{B}^{-1} - \mathbf{A}^{-1}$ has the same inertia as $\mathbf{A} - \mathbf{B}$ if and only if $\mathbf{A}$ and $\mathbf{B}$ have the same inertia, and so when both $\mathbf{A}$ and $\mathbf{B}$ are positive definite, then

$$\mathbf{B}^{-1} \geq \mathbf{A}^{-1} \Leftrightarrow \mathbf{A} \geq \mathbf{B}, \tag{1.31}$$

where $\mathbf{A} \geq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ nonnegative definite, i.e., $\eta(\mathbf{A} - \mathbf{B}) = 0$ or $\text{In}(\mathbf{A} - \mathbf{B}) = \{\bullet, 0, \bullet\}$, no negative eigenvalues — cf. (1.15).

1.3 Schur complements and matrix convexity

Anderson and Trapp (1976) posed the problem of showing that

$$\mathbf{Q} = \mathbf{A}^{-1} + \mathbf{B}^{-1} - 4(\mathbf{A} + \mathbf{B})^{-1} \geq \mathbf{0}, \tag{1.32}$$

where $\mathbf{A}$ and $\mathbf{B}$ are both symmetric positive definite. The two published solutions, by Moore (1977) and Lieb (1977), showed that (1.32) was a special case of a more general inequality and neither solution used Schur complements. We may prove (1.32) by noting that $\mathbf{Q}$ is the Schur complement of $\mathbf{A} + \mathbf{B}$ in

$$\begin{pmatrix} \mathbf{A} + \mathbf{B} & 2\mathbf{I} \\ 2\mathbf{I} & \mathbf{A}^{-1} + \mathbf{B}^{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{I} & \mathbf{A}^{-1} \end{pmatrix} + \begin{pmatrix} \mathbf{B} & \mathbf{I} \\ \mathbf{I} & \mathbf{B}^{-1} \end{pmatrix} \geq \mathbf{0}. \tag{1.33}$$

The nonnegative definiteness of (1.33) follows from

$$\text{In}\begin{pmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{I} & \mathbf{A}^{-1} \end{pmatrix} = \{n, 0, n\}, \tag{1.34}$$

where $\mathbf{A}$ is $n \times n$, since both Schur complements in (1.34) are the $n \times n$ zero matrix. Using (1.34) we may extend (1.33) to

$$\begin{pmatrix} \lambda\mathbf{A} + (1 - \lambda)\mathbf{B} & \mathbf{I} \\ \mathbf{I} & \lambda\mathbf{A}^{-1} + (1 - \lambda)\mathbf{B}^{-1} \end{pmatrix} = \lambda\begin{pmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{I} & \mathbf{A}^{-1} \end{pmatrix} + (1 - \lambda)\begin{pmatrix} \mathbf{B} & \mathbf{I} \\ \mathbf{I} & \mathbf{B}^{-1} \end{pmatrix} \geq \mathbf{0} \tag{1.35}$$

for all nonnegative $\lambda \leq 1$. Hence the Schur complement of $\lambda\mathbf{A} + (1 - \lambda)\mathbf{B}$,

$$\lambda\mathbf{A}^{-1} + (1 - \lambda)\mathbf{B}^{-1} - [\lambda\mathbf{A} + (1 - \lambda)\mathbf{B}]^{-1} \geq \mathbf{0}, \tag{1.36}$$

as shown by Moore (1973) using a simultaneous diagonalization argument. As noted by Moore (1977) the matrix-inverse function is »matrix convex» on the class of all symmetric positive definite matrices (see also Marshall and Olkin, 1979, pp. 469—471).

If the positive definite matrix $\mathbf{A}$ in (1.34) is random then it follows at once that

$$\mathscr{E}\begin{pmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{I} & \mathbf{A}^{-1} \end{pmatrix} = \begin{pmatrix} \mathscr{E}(\mathbf{A}) & \mathbf{I} \\ \mathbf{I} & \mathscr{E}(\mathbf{A}^{-1}) \end{pmatrix} \geq \mathbf{0}, \tag{1.37}$$

where $\mathscr{E}(.)$ denotes mathematical expectation, and so

$$\mathscr{E}(\mathbf{A}^{-1}) \geq [\mathscr{E}(\mathbf{A})]^{-1} \tag{1.38}$$

using the nonnegative definiteness of the Schur complement of $\mathscr{E}(\mathbf{A})$ in the matrix in the middle of (1.37). The inequality (1.38) was (first?) shown by Groves and Rothenberg (1969).

Kiefer (1959) showed that

$$\Sigma_1^k \lambda_i \mathbf{F}_i' \mathbf{A}_i^{-1} \mathbf{F}_i - (\Sigma_1^k \lambda_i \mathbf{F}_i')(\Sigma_1^k \lambda_i \mathbf{A}_i)^{-1}(\Sigma_1^k \lambda_i \mathbf{F}_i) \geq \mathbf{0}, \tag{1.39}$$

where the $\mathbf{A}_i$ are symmetric positive definite matrices and the scalars $\lambda_i \geq 0$ for all $i = 1, \ldots, k$. We may prove (1.39) by noting that

$$\mathbf{M}_i = \begin{pmatrix} \mathbf{A}_i & \mathbf{F}_i \\ \mathbf{F}_i' & \mathbf{F}_i' \mathbf{A}_i^{-1} \mathbf{F}_i \end{pmatrix} \geq \mathbf{0} \tag{1.40}$$

since the Schur complements $(\mathbf{M}_i/\mathbf{A}_i) = \mathbf{0}$ for all $i = 1, \ldots, k$. Hence $\Sigma_1^k \lambda_i \mathbf{M}_i \geq \mathbf{0}$ and so the Schur complement $(\Sigma_1^k \lambda_i \mathbf{M}_i/\Sigma_1^k \lambda_i \mathbf{A}_i)$, which is the left-hand side of (1.39), is nonnegative definite. When all the $\lambda_i = 1$ then (1.39) reduces to the result used by Lieb (1977) to prove (1.32), cf. also Lieb and Ruskai (1974).

1.4 Generalized Schur complements

If in the partitioned matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \tag{1.41}$$

the submatrix $\mathbf{E}$ is rectangular, or square but singular, then the definition (1.2) of Schur complement cannot be used. We may, however, define

$$\mathbf{S} = \mathbf{H} - \mathbf{G}\mathbf{E}^{-}\mathbf{F} = (\mathbf{A}/\mathbf{E}) \tag{1.42}$$

as *a generalized Schur complement of* $\mathbf{E}$ *in* $\mathbf{A}$, where $\mathbf{E}^{-}$ is a generalized inverse of $\mathbf{E}$, i.e., $\mathbf{E}\mathbf{E}^{-}\mathbf{E} = \mathbf{E}$, cf. e.g., Rao (1985, Section 1.1). In general this generalized Schur complement $\mathbf{H} - \mathbf{G}\mathbf{E}^{-}\mathbf{F}$ will depend on the choice of generalized inverse $\mathbf{E}^{-}$. If we replace $\mathbf{E}^{-1}$ with an $\mathbf{E}^{-}$ in (1.4), we obtain

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{G}\mathbf{E}_1^{-} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} - \mathbf{G}\mathbf{E}_2^{-}\mathbf{F} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{E}_3^{-}\mathbf{F} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{E} & \mathbf{E}\mathbf{E}_3^{-}\mathbf{F} \\ \mathbf{G}\mathbf{E}_1^{-}\mathbf{E} & \mathbf{H} - \mathbf{G}\mathbf{E}_2^{-}\mathbf{F} + \mathbf{G}\mathbf{E}_1^{-}\mathbf{E}\mathbf{E}_3^{-}\mathbf{F} \end{pmatrix}, \tag{1.43}$$

where $\mathbf{E}_1^-$, $\mathbf{E}_2^-$, and $\mathbf{E}_3^-$ are three (possibly different) choices of generalized inverse(s) of $\mathbf{E}$. Then (1.43) is equal to the matrix $\mathbf{A}$, cf. (1.4), if and only if

$$\mathbf{G}\mathbf{E}_1^-\mathbf{E} = \mathbf{G} \quad \text{and} \quad \mathbf{E}\mathbf{E}_3^-\mathbf{F} = \mathbf{F} \tag{1.44}$$

[we put $\mathbf{E} = \mathbf{E}\mathbf{E}_2^-\mathbf{E}$ in the bottom right-hand corner of the last matrix in (1.43)]. The conditions in (1.44), however, do not depend on the generalized inverses involved, and are equivalent, respectively, to

$$\operatorname{rank}\begin{pmatrix}\mathbf{E}\\ \mathbf{G}\end{pmatrix} = \operatorname{rank}(\mathbf{E}) \quad \text{and} \quad \operatorname{rank}(\mathbf{E}, \mathbf{F}) = \operatorname{rank}(\mathbf{E}), \tag{1.45}$$

cf. e.g., Marsaglia and Styan (1974a, Theorem 5), Ouellette (1981, Lemma 4.1). If follows that when (1.45) [or equivalently (1.44)] holds, then $(\mathbf{A}/\mathbf{E}) = \mathbf{H} - \mathbf{G}\mathbf{E}^-\mathbf{F}$ is uniquely defined and becomes *the* generalized Schur complement of $\mathbf{E}$ in $\mathbf{A}$. [To see this write $\mathbf{G}\mathbf{E}_2^-\mathbf{F} = (\mathbf{G}\mathbf{E}_1^-\mathbf{E})\mathbf{E}_2^-(\mathbf{E}\mathbf{E}_3^-\mathbf{F}) = \mathbf{G}\mathbf{E}_1^-(\mathbf{E}\mathbf{E}_2^-\mathbf{E})\mathbf{E}_3^-\mathbf{F} = \mathbf{G}\mathbf{E}_1^-(\mathbf{E}\mathbf{E}_4^-\mathbf{E})\mathbf{E}_3^-\mathbf{F} = (\mathbf{G}\mathbf{E}_1^-\mathbf{E})\mathbf{E}_4^-(\mathbf{E}\mathbf{E}_3^-\mathbf{F}) = \mathbf{G}\mathbf{E}_4^-\mathbf{F}$.]

Carlson (1984, Section 3) has pointed out that the conditions (1.45) are both necessary and sufficient [when neither $\mathbf{F}$ nor $\mathbf{G}$ is the null matrix] for the uniqueness of the generalized Schur complement $(\mathbf{A}/\mathbf{E})$, and that matrices $\mathbf{A}$ satisfying (1.45) provide the »natural setting» for results in generalized Schur complements.

Schur's determinantal formula (1.3) is not of interest when $\mathbf{E}$ is not square and nonsingular. Guttman's rank formula (1.5), however,

$$\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{E}) + \operatorname{rank}(\mathbf{A}/\mathbf{E}) \tag{1.46}$$

is of interest and does hold whenever (1.45) [or equivalently (1.44)] holds. As an example to illustrate (1.46) consider the partitioned matrix

$$\begin{pmatrix}\mathbf{E} & \mathbf{E}\mathbf{H}\\ \mathbf{H}\mathbf{E} & \mathbf{H}\end{pmatrix}. \tag{1.47}$$

Then using (1.46) and its counterpart for the »other» generalized Schur complement $(\mathbf{A}/\mathbf{H}) = \mathbf{E} - \mathbf{F}\mathbf{H}^-\mathbf{G}$, we obtain

$$\operatorname{rank}(\mathbf{E}) + \operatorname{rank}(\mathbf{H} - \mathbf{H}\mathbf{E}\mathbf{H}) = \operatorname{rank}(\mathbf{H}) + \operatorname{rank}(\mathbf{E} - \mathbf{E}\mathbf{H}\mathbf{E}); \tag{1.48}$$

when $\mathbf{H} = \mathbf{E}^-$, (1.48) reduces to [cf. Ouellette (1981, p. 247)]

$$\operatorname{rank}(\mathbf{E}^- - \mathbf{E}^-\mathbf{E}\mathbf{E}^-) = \operatorname{rank}(\mathbf{E}^-) - \operatorname{rank}(\mathbf{E}) \tag{1.49}$$

and so $\mathbf{E}^-$ is a *reflexive* generalized inverse of $\mathbf{E}$ [in that it satisfies the first *two* of the Penrose conditions (cf. e.g., Rao, 1985, Section 1.1)] if and only if $\mathbf{E}^-$ has the same rank as $\mathbf{E}$, a result due to Bjerhammar (1958). [See also Styan (1983).]

The Banachiewicz inversion formula (1.11) generalizes in the obvious manner; see Marsaglia and Styan (1974b) for details.

1.5 Generalized Schur complements and inertia

When the square matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}' & \mathbf{H} \end{pmatrix} \tag{1.50}$$

is symmetric, then the two conditions in (1.45) reduce to just

$$\operatorname{rank}(\mathbf{E}, \mathbf{F}) = \operatorname{rank}(\mathbf{E}) \tag{1.51}$$

and in this event Haynsworth's inertia formula (1.14)

$$\operatorname{In}\mathbf{A} = \operatorname{In}\mathbf{E} + \operatorname{In}(\mathbf{A}/\mathbf{E}) \tag{1.52}$$

holds. To illustrate the use of (1.52) consider the matrix

$$\mathbf{M}_1 = \begin{pmatrix} \mathbf{A} & \mathbf{A}\mathbf{A}^- \\ \mathbf{A}^-\mathbf{A} & \mathbf{B}^- \end{pmatrix}, \tag{1.53}$$

where $\mathbf{A}$, $\mathbf{A}^-$, $\mathbf{B}$, and $\mathbf{B}^-$ are all symmetric, and $\mathbf{A}^-$ and $\mathbf{B}^-$ are both reflexive generalized inverses. Then

$$\operatorname{In}\mathbf{M}_1 = \operatorname{In}\mathbf{A} + \operatorname{In}(\mathbf{B}^- - \mathbf{A}^-). \tag{1.54}$$

If

$$\mathbf{A}\mathbf{A}^- = \mathbf{B}\mathbf{B}^- \tag{1.55}$$

then we may apply the inertia formula (1.20) to the »other» generalized Schur complement $(\mathbf{M}_1/\mathbf{B}^-)$ to obtain

$$\operatorname{In}\mathbf{M}_1 = \operatorname{In}\mathbf{B}^- + \operatorname{In}(\mathbf{A} - \mathbf{B}) = \operatorname{In}\mathbf{B} + \operatorname{In}(\mathbf{A} - \mathbf{B}) \tag{1.56}$$

since the generalized inverse $\mathbf{B}^-$ is symmetric and reflexive. Hence when (1.55) holds then

$$\mathbf{A} \geq \mathbf{B} \geq \mathbf{0} \tag{1.57}$$

if and only if

$$\mathbf{B}^- \geq \mathbf{A}^- \geq \mathbf{0}. \tag{1.58}$$

On the other hand consider the matrix

$$\mathbf{M}_2 = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{B} \end{pmatrix}; \tag{1.59}$$

then

$$\operatorname{In}\mathbf{M}_2 = \operatorname{In}\mathbf{B} + \operatorname{In}(\mathbf{A} - \mathbf{B}) \tag{1.60}$$

and so when (1.57) holds then $\mathbf{M}_2 \geq \mathbf{0}$ which implies that, cf. (1.44) and (1.45),

$$\operatorname{rank}(\mathbf{A}, \mathbf{B}) = \operatorname{rank}(\mathbf{A}) \Leftrightarrow \mathbf{A}\mathbf{A}^{-}\mathbf{B} = \mathbf{B} \tag{1.61}$$

since we may write $\mathbf{A} = \mathbf{X}'\mathbf{X}$ and $\mathbf{B} = \mathbf{X}'\mathbf{Y}$ when $\mathbf{M}_2 \geq \mathbf{0}$. Similarly when (1.58) holds then

$$\operatorname{rank}(\mathbf{A}^{-}, \mathbf{B}^{-}) = \operatorname{rank}(\mathbf{B}^{-}) \Leftrightarrow \mathbf{B}^{-}\mathbf{B}\mathbf{A}^{-} = \mathbf{A}^{-} \tag{1.62}$$

since $\mathbf{B}^{-}$ is reflexive. When both (1.57) and (1.58) hold, therefore, (1.61) implies that $\mathbf{A}\mathbf{A}^{-}\mathbf{B}\mathbf{B}^{-} = \mathbf{B}\mathbf{B}^{-}$, while (1.62) implies that $\mathbf{B}^{-}\mathbf{B}\mathbf{A}^{-}\mathbf{A} = \mathbf{A}^{-}\mathbf{A}$. Since all the matrices involved are symmetric it follows that $\mathbf{B}\mathbf{B}^{-} = \mathbf{A}\mathbf{A}^{-}$, i.e., (1.57) and (1.58) imply (1.55). We have proved, therefore, the following result [due to Styan and Pukelsheim (1978), cf. Ouellette (1981, Theorem 4.13)].

THEOREM 1.1. *If any two of the following three conditions hold then all three hold:*

(1.55) . . . $\mathbf{A}\mathbf{A}^{-} = \mathbf{B}\mathbf{B}^{-}$,
(1.57) . . . $\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$,
(1.58) . . . $\mathbf{B}^{-} \geq \mathbf{A}^{-} \geq \mathbf{0}$.

This result is a direct extension of (1.31), which holds when both $\mathbf{A}$ and $\mathbf{B}$ are symmetric positive definite. Another extension of (1.31), due to Milliken and Akdeniz (1977) and Hartwig (1978), uses Moore-Penrose generalized inverses $\mathbf{A}^{+}$ and $\mathbf{B}^{+}$, which satisfy all four of the Penrose conditions (cf. e.g., Rao, 1985, Section 1.1, or Styan, 1983).

THEOREM 1.2. *If any two of the following three conditions hold then all three hold:*

(1.63) . . . $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{B})$,
(1.57) . . . $\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$,
(1.64) . . . $\mathbf{B}^{+} \geq \mathbf{A}^{+} \geq \mathbf{0}$.

Proof. Conditions (1.57) and (1.64) imply (1.63) from Theorem 1. Moreover, when (1.57) holds, then $\mathbf{A}\mathbf{A}^{+}\mathbf{B} = \mathbf{B} \Leftrightarrow \operatorname{rank}(\mathbf{A}, \mathbf{B}) = \operatorname{rank}(\mathbf{A})$, as in (1.61). When both (1.57) and (1.63) hold, therefore, $\operatorname{rank}(\mathbf{A}, \mathbf{B}) = \operatorname{rank}(\mathbf{B}) \Leftrightarrow \mathbf{B}\mathbf{B}^{+}\mathbf{A} = \mathbf{A}$. Combining yields, respectively, $\mathbf{A}\mathbf{A}^{+}\mathbf{B}\mathbf{B}^{+} = \mathbf{B}\mathbf{B}^{+}$ and $\mathbf{B}\mathbf{B}^{+}\mathbf{A}\mathbf{A}^{+} = \mathbf{A}\mathbf{A}^{+}$. Since both $\mathbf{A}\mathbf{A}^{+}$ and $\mathbf{B}\mathbf{B}^{+}$ are symmetric, it follows that $\mathbf{A}\mathbf{A}^{+} = \mathbf{B}\mathbf{B}^{+}$ and so (1.64) follows from Theorem 1.1. A similar argument shows that (1.63) and (1.64) imply (1.57). Q.E.D.

Our proof of Theorem 1.2 parallels that given by Ouellette (1981, Corollary 4.8).

When rank(**A**, **B**) = rank(**A**) then we may apply (1.20) to the symmetric matrix $\mathbf{M}_2$ in (1.59) and obtain:

$$\begin{aligned} \mathrm{In}\mathbf{M}_2 &= \mathrm{In}\mathbf{A} + \mathrm{In}(\mathbf{B} - \mathbf{B}\mathbf{A}^-\mathbf{B}) \\ &= \mathrm{In}\mathbf{B} + \mathrm{In}(\mathbf{A} - \mathbf{B}) \end{aligned} \tag{1.65}$$

from (1.60). Hence, provided rank(**A**, **B**) = rank(**A**), we find that

$$\begin{aligned} \mathrm{In}(\mathbf{B} - \mathbf{B}\mathbf{A}^-\mathbf{B}) &= \mathrm{In}[\mathbf{B}(\mathbf{B}^- - \mathbf{A}^-)\mathbf{B}] \\ &= \mathrm{In}(\mathbf{A} - \mathbf{B}) - [\mathrm{In}\mathbf{A} - \mathrm{In}\mathbf{B}], \end{aligned} \tag{1.66}$$

which extends (1.30) to possibly singular (but still symmetric) matrices **A** and **B**.

Therefore, when $\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$, then rank(**A**, **B**) = rank(**A**) from (1.61), and so from (1.66) it follows that (cf. Gaffke and Krafft, 1982, Theorem 3.5)

$$\mathbf{B} - \mathbf{B}\mathbf{A}^-\mathbf{B} \geq \mathbf{0} \tag{1.67}$$

and

$$\begin{aligned} \mathrm{rank}(\mathbf{B} - \mathbf{B}\mathbf{A}^-\mathbf{B}) &= \mathrm{rank}(\mathbf{A} - \mathbf{B}) - [\mathrm{rank}(\mathbf{A}) - \mathrm{rank}(\mathbf{B})] \\ &= \mathrm{rank}[\mathbf{B}(\mathbf{B}^- - \mathbf{A}^-)\mathbf{B}] \\ &\leq \mathrm{rank}(\mathbf{B}^- - \mathbf{A}^-). \end{aligned} \tag{1.68}$$

When rank(**A**) = rank(**B**) and $\mathbf{A} \geq \mathbf{B} \geq \mathbf{0}$, we have

$$\mathrm{rank}(\mathbf{A} - \mathbf{B}) \leq \mathrm{rank}(\mathbf{B}^- - \mathbf{A}^-), \tag{1.69}$$

with equality if (but not necessarily only if) **B** is positive definite.

1.6 Schur complements and statistics

Ouellette (1981), in her survey paper with this title, presented applications of Schur complements in the following five areas of statistics, all of which may be considered as being in multivariate analysis (Anderson, 1984):

(1) The multivariate normal distribution.
(2) Partial correlation coefficients.
(3) Special covariance and correlation structures.
(4) The chi-squared and Wishart distributions.
(5) The (multiparameter) Cramér-Rao inequality.

In this paper our applications of Schur complements to statistics will concentrate on their use in linear statistical models. Indeed, Ouellette (1981, Section 6.4) observed that in the general linear model

$$\mathscr{E}(\mathbf{y}) = \mathbf{X}\beta \tag{1.70}$$

the residual sum of squares

$$\mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{y} \tag{1.71}$$

is the generalized Schur complement of $\mathbf{X}'\mathbf{X}$ in the matrix

$$(\mathbf{X}, \mathbf{y})'(\mathbf{X}, \mathbf{y}) = \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{y} \\ \mathbf{y}'\mathbf{X} & \mathbf{y}'\mathbf{y} \end{pmatrix}. \tag{1.72}$$

Alalouf and Styan (1979a) used Schur complements in their study of estimability of $\mathbf{A}\beta$ in the general linear model, while Anderson and Styan (1982) concentrated on Cochran's theorem and tripotent matrices. Pukelsheim and Styan (1983) used Schur complements in their paper on the convexity and monotonicity properties of dispersion matrices of estimators in linear models.

In this paper we will concentrate on the general partitioned linear model

$$\mathscr{E}(\mathbf{y}) = \mathbf{X}\beta = \mathbf{X}_1\alpha + \mathbf{X}_2\gamma, \tag{1.73}$$

where the design matrix is partitioned

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2). \tag{1.74}$$

In Section 2 we study canonical correlations, and identify the numbers that are equal to one and are less than one. We also consider the canonical correlations between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$ and examine the hypothesis that $\mathscr{E}(\mathbf{y}) = \mathbf{X}_2\gamma$; our development builds on results in the paper by Latour and Styan (1985) in these *Proceedings*. In Section 3 we study the matrix problem posed and solved by Broyden (1982, 1983), and set up a closely related analysis-of-covariance linear statistical model.

2. CANONICAL CORRELATIONS AND THE GENERAL PARTITIONED LINEAR MODEL

2.1 Canonical correlations: the number less than one and the number equal to one

The canonical correlations between two random vectors (or between two sets of random variables) are the correlations between certain linear combinations of the random variables in each of the two vectors (or sets), cf. e.g., Anderson (1984, Chapter 12).

Consider the $p \times 1$ random vector

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \tag{2.1}$$

where $\mathbf{x}_1$ is $p_1 \times 1$ and $\mathbf{x}_2$ is $p_2 \times 1$, with $p_1 + p_2 = p$. Then the *first canonical correlation* between $\mathbf{x}_1$ and $\mathbf{x}_2$ is the largest correlation ϱ_1, say, between $\mathbf{a}'\mathbf{x}_1$ and $\mathbf{b}'\mathbf{x}_2$ for all possible nonrandom $p_1 \times 1$ vectors $\mathbf{a}$ and $p_2 \times 1$ vectors $\mathbf{b}$. If $\mathbf{a} = \mathbf{a}_1$ and $\mathbf{b} = \mathbf{b}_1$ are the two maximizing vectors so that

$$\varrho_1 = \text{corr}(\mathbf{a}_1'\mathbf{x}_1, \mathbf{b}_1'\mathbf{x}_2) \tag{2.2}$$

then the pair $(\mathbf{a}_1'\mathbf{x}_1, \mathbf{b}_1'\mathbf{x}_2)$ is said to be the *first pair of canonical variates.*

The second pair of canonical variates is that pair of linear combinations $(\mathbf{a}_2'\mathbf{x}_1, \mathbf{b}_2'\mathbf{x}_2)$, say, such that

$$\varrho_2 = \text{corr}(\mathbf{a}_2'\mathbf{x}_1, \mathbf{b}_2'\mathbf{x}_2) \geq \text{corr}(\mathbf{a}'\mathbf{x}_1, \mathbf{b}'\mathbf{x}_2) \tag{2.3}$$

for all $\mathbf{a}$ and $\mathbf{b}$ satisfying

$$\begin{aligned}\text{corr}(\mathbf{a}'\mathbf{x}_1, \mathbf{a}_1'\mathbf{x}_1) &= \text{corr}(\mathbf{a}'\mathbf{x}_1, \mathbf{b}_1'\mathbf{x}_2)\\ &= \text{corr}(\mathbf{b}'\mathbf{x}_2, \mathbf{a}_1'\mathbf{x}_1) = \text{corr}(\mathbf{b}'\mathbf{x}_2, \mathbf{b}_1'\mathbf{x}_2) = 0.\end{aligned} \tag{2.4}$$

The correlation ϱ_2 is called the second canonical correlation.

Higher order canonical correlations and canonical variates are defined in a similar manner. Only positive canonical correlations are defined and the number of them, as we shall see below in Theorem 2.1(b), cannot exceed the smaller of p_1 and p_2. When $\min(p_1, p_2) = 1$ then there is only one canonical correlation and this is called the multiple correlation coefficient (unless the vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are completely uncorrelated in which event there are no canonical correlations).

As has been shown, for example, by Anderson (1984, Section 12.2), the canonical correlations ϱ_h and the vectors $\mathbf{a}_h$ and $\mathbf{b}_h$ defining the canonical variates $\mathbf{a}_h'\mathbf{x}_1$ and $\mathbf{b}_h'\mathbf{x}_2$ satisfy the matrix equation

$$\begin{pmatrix}-\varrho_h\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & -\varrho_h\Sigma_{22}\end{pmatrix}\begin{pmatrix}\mathbf{a}_h\\ \mathbf{b}_h\end{pmatrix} = \mathbf{0}, \tag{2.5}$$

where the covariance matrix

$$\Sigma = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix} = \mathcal{V}ar(\mathbf{x}) = \mathcal{V}ar\begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix}. \tag{2.6}$$

Following Khatri (1976), Seshadri and Styan (1980), and Rao (1981), we have:

THEOREM 2.1. *The nonzero eigenvalues and the rank of the matrix*

$$\mathbf{P} = \Sigma_{11}^{-}\Sigma_{12}\Sigma_{22}^{-}\Sigma_{21} \tag{2.7}$$

are invariant under choices of generalized inverses Σ_{11}^{-} and Σ_{22}^{-}. Moreover

(a) *The eigenvalues of* $\mathbf{P}$ *are the squares of the canonical correlations between* $\mathbf{x}_1$ *and* $\mathbf{x}_2$.

(b) *The number of nonzero canonical correlations between* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *is*

$$\operatorname{rank}(\mathbf{P}) = \operatorname{rank}(\Sigma_{12}) \leq \min(p_1, p_2). \tag{2.8}$$

(c) *The number of canonical correlations equal to* 1 *is*

$$u = \operatorname{rank}(\Sigma_{11}) + \operatorname{rank}(\Sigma_{22}) - \operatorname{rank}(\Sigma). \tag{2.9}$$

(d) *When* Σ *is positive definite then* $u = 0$, *i.e., there are no canonical correlations equal to* 1.

Proof. Since Σ is nonnegative definite (by definition of covariance matrix), we may write

$$\Sigma = \mathbf{T}'\mathbf{T} = (\mathbf{T}_1, \mathbf{T}_2)'(\mathbf{T}_1, \mathbf{T}_2) \tag{2.10}$$

and so

$$\mathbf{P} = (\mathbf{T}_1'\mathbf{T}_1)^-\mathbf{T}_1'\mathbf{T}_2(\mathbf{T}_2'\mathbf{T}_2)^-\mathbf{T}_2'\mathbf{T}_1, \tag{2.11}$$

which has the same nonzero eigenvalues, cf. (1.22), as the matrix

$$\mathbf{Q} = \mathbf{H}_1\mathbf{H}_2, \tag{2.12}$$

where

$$\mathbf{H}_i = \mathbf{T}_i(\mathbf{T}_i'\mathbf{T}_i)^-\mathbf{T}_i'; \qquad i = 1,2, \tag{2.13}$$

is symmetric idempotent and invariant under choice of generalized inverse of $\mathbf{T}_i'\mathbf{T}_i = \Sigma_{ii}$, $i = 1,2$. Moreover

$$\begin{aligned}\operatorname{rank}(\mathbf{P}) \leq \operatorname{rank}(\mathbf{T}_1'\mathbf{T}_2) &= \operatorname{rank}[\mathbf{T}_1'\mathbf{T}_1(\mathbf{T}_1'\mathbf{T}_1)^-\mathbf{T}_1'\mathbf{T}_2(\mathbf{T}_2'\mathbf{T}_2)^-\mathbf{T}_2'\mathbf{T}_2]\\ &\leq \operatorname{rank}(\mathbf{H}_1\mathbf{H}_2)\\ &= \operatorname{rank}(\mathbf{H}_1\mathbf{H}_2\mathbf{H}_1)\\ &\leq \operatorname{rank}(\mathbf{P}),\end{aligned} \tag{2.14}$$

since $\mathbf{T}_i'\mathbf{T}_i(\mathbf{T}_i'\mathbf{T}_i)^-\mathbf{T}_i' = \mathbf{T}_i'$ $(i = 1,2)$, the rank of a product cannot exceed that of any factor, and $\mathbf{H}_2 = \mathbf{H}_2'\mathbf{H}_2 \geq \mathbf{0}$. The rank of the matrix $\mathbf{P}$, therefore, is equal to the rank of the matrix $\mathbf{T}_1'\mathbf{T}_2 = \Sigma_{12}$, for all possible choices of generalized inverses Σ_{ii}^- $(i = 1,2)$. This also proves (b).

To prove (a), we use the singular value decompositions (cf. Seshadri and Styan, 1980, p. 334)

$$\mathbf{T}_i = \mathbf{U}_i\mathbf{D}_i\mathbf{V}_i', \qquad i = 1,2, \tag{2.15}$$

where

$$\mathbf{U}_i'\mathbf{U}_i = \mathbf{V}_i'\mathbf{V}_i = \mathbf{I}_{r_i}; \qquad i = 1,2, \tag{2.16}$$

and

$$r_i = \operatorname{rank}(\mathbf{T}_i) = \operatorname{rank}(\Sigma_{ii}); \qquad i = 1,2. \tag{2.17}$$

The diagonal matrix $\mathbf{D}_i$ is $r_i \times r_i$ ($i = 1,2$). We may then write (2.5) as:

$$\mathbf{B}\begin{pmatrix} -\varrho\mathbf{I}_{r_1} & \mathbf{U}_1'\mathbf{U}_2 \\ \mathbf{U}_2'\mathbf{U}_1 & -\varrho\mathbf{I}_{r_2} \end{pmatrix}\mathbf{B}'\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} = \mathbf{0}, \tag{2.18}$$

where

$$\mathbf{B} = \begin{pmatrix} \mathbf{V}_1\mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2\mathbf{D}_2 \end{pmatrix} \tag{2.19}$$

has full column rank equal to $r_1 + r_2$. Hence (2.5) has a nontrivial solution if and only if

$$\begin{aligned} \det\begin{pmatrix} -\varrho\mathbf{I}_{r_1} & \mathbf{U}_1'\mathbf{U}_2 \\ \mathbf{U}_2'\mathbf{U}_1 & -\varrho\mathbf{I}_{r_2} \end{pmatrix} = 0 &= (-\varrho)^{r_1}\cdot\det[-\varrho\mathbf{I}_{r_2} - \mathbf{U}_2'\mathbf{U}_1(-1/\varrho)\mathbf{U}_1'\mathbf{U}_2] \\ &= (-\varrho)^{r_1 - r_2}\cdot\det(\varrho^2\mathbf{I}_{r_2} - \mathbf{U}_2'\mathbf{U}_1\mathbf{U}_1'\mathbf{U}_2), \end{aligned} \tag{2.20}$$

using the Schur determinantal formula (1.3). The canonical correlations, therefore, are the positive square roots of the nonzero eigenvalues of $\mathbf{U}_2'\mathbf{U}_1\mathbf{U}_1'\mathbf{U}_2$, or equivalently of $\mathbf{U}_1\mathbf{U}_1'\mathbf{U}_2\mathbf{U}_2' = \mathbf{H}_1\mathbf{H}_2 = \mathbf{Q}$, or of $\mathbf{P}$, cf. (2.11) and (2.12).

To prove (c) we evaluate the number u of canonical correlations equal to 1. From (2.20) we see that

$$u = \nu(\mathbf{I}_{r_2} - \mathbf{U}_2'\mathbf{U}_1\mathbf{U}_1'\mathbf{U}_2) = \nu(\mathbf{M}/\mathbf{I}_{r_1}) = \nu(\mathbf{M}), \tag{2.21}$$

using (1.25), where

$$\mathbf{M} = \begin{pmatrix} \mathbf{I}_{r_1} & \mathbf{U}_1'\mathbf{U}_2 \\ \mathbf{U}_2'\mathbf{U}_1 & \mathbf{I}_{r_2} \end{pmatrix} = (\mathbf{U}_1, \mathbf{U}_2)'(\mathbf{U}_1, \mathbf{U}_2) \tag{2.22}$$

has rank equal to the rank of $\mathbf{BMB}' = \Sigma$, cf. (2.18), because $\mathbf{B}$ has full column rank. Hence

$$u = \nu(\mathbf{M}) = r_1 + r_2 - \operatorname{rank}(\mathbf{M}) = r_1 + r_2 - \operatorname{rank}(\Sigma), \tag{2.23}$$

which proves (c). Part (d) follows trivially, and our proof is complete. Q.E.D.

2.2 Canonical correlations in the general partitioned linear model

Latour and Styan (1985), in their paper in these *Proceedings,* considered the canonical correlations between the vectors of row and column totals in the usual two-way layout without interaction:

$$\mathscr{E}(y_{ijk}) = \alpha_i + \gamma_j; \quad i = 1,\ldots,r, \;\; j = 1,\ldots,c, \;\; k = 1,\ldots,n_{ij}, \tag{2.24}$$

with possibly unequal numbers $n_{ij} \geq 0$ of observations in the cells, cf. their (1.1). They wrote this model (2.24) in matrix notation, cf. their (1.2),

$$\mathscr{E}(\mathbf{y}) = \mathbf{X}\beta = \mathbf{X}_1\alpha + \mathbf{X}_2\gamma, \tag{2.25}$$

where $\mathbf{y} = \{y_{ijk}\}$ is the $n \times 1$ vector of observations, with $n = \Sigma_{i,j}n_{ij}$, while the $(r + c) \times 1$ vector $\beta = \begin{pmatrix}\alpha\\ \gamma\end{pmatrix}$, with $\alpha = \{\alpha_i\}$ and $\gamma = \{\gamma_j\}$.

The $n \times (r + c)$ partitioned design matrix $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ satisfies

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix}\mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2\\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2\end{pmatrix} = \begin{pmatrix}\mathbf{D}_r & \mathbf{N}\\ \mathbf{N}' & \mathbf{D}_c\end{pmatrix}, \tag{2.26}$$

where $\mathbf{N} = \{n_{ij}\}$, $\mathbf{D}_r = \text{diag}\{n_{i.}\}$, and $\mathbf{D}_c = \textit{diag}\{n_{.j}\}$, with $n_{i.} = \Sigma_j n_{ij}$ and $n_{.j} = \Sigma_i n_{ij}$. Latour and Styan (1985) assumed that all the $n_{i.}$ and $n_{.j}$ were positive so that both $\mathbf{D}_r$ and $\mathbf{D}_c$ are positive definite. They also assumed, as will we, that the error vector $\mathbf{y} - \mathscr{E}(\mathbf{y})$ satisfies the white noise assumption so that

$$\mathscr{V}\!ar(\mathbf{y}) = \sigma^2\mathbf{I} \tag{2.27}$$

for some (unknown) positive scalar σ^2.

The vectors of row and column totals are $\mathbf{y}_{rt} = \mathbf{X}_1'\mathbf{y}$ and $\mathbf{y}_{ct} = \mathbf{X}_2'\mathbf{y}$, and when (2.27) holds we have

$$\mathscr{V}\!ar(\mathbf{X}'\mathbf{y}) = \mathscr{V}\!ar\begin{pmatrix}\mathbf{X}_1'\mathbf{y}\\ \mathbf{X}_2'\mathbf{y}\end{pmatrix} = \sigma^2\mathbf{X}'\mathbf{X} = \sigma^2\begin{pmatrix}\mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2\\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2\end{pmatrix}. \tag{2.28}$$

In the two-way layout defined by (2.24) the matrix (2.28) is equal to σ^2 times the matrix in (2.26).

Latour and Styan (1985) studied the canonical correlations ϱ_h between the vectors of row and column totals: $\mathbf{y}_{rt} = \mathbf{X}_1'\mathbf{y}$ and $\mathbf{y}_{ct} = \mathbf{X}_2'\mathbf{y}$ in the two-way layout. It follows from Theorem 2.1 that the ϱ_h are the positive square roots of the nonzero eigenvalues of the matrix $(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{X}_1 = \mathbf{D}_r^{-1}\mathbf{N}\mathbf{D}_c^{-1}\mathbf{N}'$. The quantities $1 - \varrho_h^2$ are called then »canonical efficiency factors» (James and Wilkinson, 1971). In an experimental design setting all the $n_{i.}$ are equal (say to s) and all the $n_{.j}$ are equal (say to k), so that $\mathbf{D}_r = s\mathbf{I}_r$ and $\mathbf{D}_c = k\mathbf{I}_c$.

With

$$1 > \varrho_1 \geq \varrho_2 \geq \ldots \geq \varrho_t > 0 \tag{2.29}$$

and

$$u = \#\{\varrho_h = 1\} \tag{2.30}$$

Latour and Styan (1985) proved [in their Theorem 1(i,ii)] that

$$\text{In}(\mathbf{S}_r - \mathbf{S}_r\mathbf{D}_r^{-1}\mathbf{S}_r) = \{t, 0, r - t\}, \tag{2.31}$$

where the Schur complement

$$\mathbf{S}_r = (\mathbf{X}'\mathbf{X}/\mathbf{D}_c) = \mathbf{D}_r - \mathbf{N}\mathbf{D}_c^{-1}\mathbf{N}'. \tag{2.32}$$

They also showed (their Theorem 3) that the eigenvalues of the matrix

$$\mathbf{L} = \mathbf{D}_r^{-1}\mathbf{S}_r - (1 - \varrho_1^2)\mathbf{S}_r^-\mathbf{S}_r \tag{2.33}$$

do not depend on the choice of generalized inverse $\mathbf{S}_r^-$ and that the eigenvalues are: 0 (multiplicity $u + 1$), ϱ_1^2 (multiplicity $r - t - u$), and $\varrho_1^2 - \varrho_h^2$ ($h = 2, \ldots, t$).

In this paper we extend these results to the general partitioned linear model (2.25), where $\mathbf{X}_i$ is $n \times p_i$ with rank equal to q_i ($i = 1,2$, or absent). In the two-way layout, therefore,

$$p_1 = q_1 = r \quad \text{and} \quad p_2 = q_2 = c, \tag{2.34}$$

while in general

$$p_1 \geq q_1 \quad \text{and} \quad p_2 \geq q_2. \tag{2.35}$$

We will keep the notation defined by (2.29) and (2.30) for our more general set-up, so that (using our Theorem 2.1)

$$t + u = \text{rank}(\mathbf{X}_1'\mathbf{X}_2) \tag{2.36}$$

and

$$\begin{aligned} u &= \text{rank}(\mathbf{X}_1) + \text{rank}(\mathbf{X}_2) - \text{rank}(\mathbf{X}) \\ &= q_1 + q_2 - q \leq p_1 + p_2 - p = \nu(\mathbf{X}). \end{aligned} \tag{2.37}$$

In the two-way layout we have equality throughout (2.37), cf. (1.14) in Latour and Styan (1985).

The generalized Schur complement

$$\begin{aligned} \mathbf{S} = (\mathbf{X}'\mathbf{X}/\mathbf{X}_2'\mathbf{X}_2) &= \mathbf{X}_1'\mathbf{X}_1 - \mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{X}_1 \\ &= \mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1, \end{aligned} \tag{2.38}$$

with

$$\mathbf{M}_2 = \mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2', \tag{2.39}$$

does not depend on the choice of generalized inverse in view of the uniqueness of the symmetric idempotent matrix (2.39). Furthermore, the matrix $\mathbf{S}$ reduces to $\mathbf{S}_r$ in the two-way layout, cf. (2.32).

We then have:

THEOREM 2.2. *The matrix*

$$\mathbf{T} = \mathbf{S} - \mathbf{S}(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}, \tag{2.40}$$

where $\mathbf{S}$ *is defined by* (2.38), *does not depend on the choice of generalized inverse* $(\mathbf{X}_1'\mathbf{X}_1)^-$, *and has inertia*

$$\mathrm{In}\mathbf{T} = \{t,\ 0,\ p_1 - t\}. \tag{2.41}$$

Proof. The matrix $\mathbf{T}$ is the generalized Schur complement of $\mathbf{X}_1'\mathbf{X}_1$ in the matrix

$$\mathbf{U} = \begin{pmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{S} \\ \mathbf{S} & \mathbf{S} \end{pmatrix}, \tag{2.42}$$

and is uniquely defined since

$$\begin{aligned} \mathrm{rank}(\mathbf{X}_1'\mathbf{X}_1,\ \mathbf{S}) &= \mathrm{rank}(\mathbf{X}_1'\mathbf{X}_1,\ \mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1) \\ &= \mathrm{rank}[\mathbf{X}_1'(\mathbf{X}_1,\ \mathbf{M}_2\mathbf{X}_1)] \\ &\le \mathrm{rank}(\mathbf{X}_1) = \mathrm{rank}(\mathbf{X}_1'\mathbf{X}_1) \le \mathrm{rank}(\mathbf{X}_1'\mathbf{X}_1,\ \mathbf{S}), \end{aligned} \tag{2.43}$$

and so equality holds throughout (2.43), cf. (1.45) and the discussion directly thereafter. Hence

$$\begin{aligned} \mathrm{In}\mathbf{T} &= \mathrm{In}\mathbf{U} - \mathrm{In}(\mathbf{X}_1'\mathbf{X}_1) = \mathrm{In}\mathbf{S} + \mathrm{In}(\mathbf{X}_1'\mathbf{X}_1 - \mathbf{S}) - \mathrm{In}(\mathbf{X}_1'\mathbf{X}_1) \\ &= \mathrm{In}(\mathbf{X}'\mathbf{X}) - \mathrm{In}(\mathbf{X}_2'\mathbf{X}_2) + \mathrm{In}[\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{X}_1] \\ &\qquad - \mathrm{In}(\mathbf{X}_1'\mathbf{X}_1) \end{aligned} \tag{2.44}$$

and so

$$\begin{aligned} \mathrm{In}\mathbf{T} &= \{q,\ 0,\ p - q\} - \{q_2,\ 0,\ p_2 - q_2\} + \{t + u,\ 0,\ p_1 - t - u\} \\ &\quad - \{q_1,\ 0,\ p_1 - q_1\} \\ &= \{t,\ 0,\ p_1 - t\}, \end{aligned} \tag{2.45}$$

since $u - q_1 - q_2 + q = 0$, cf. (2.37), and $p = p_1 + p_2$. Q.E.D.

Our Theorem 2.2 above extends Theorem 1(i,ii) of Latour and Styan (1985). We extend their Theorem 3 with our

THEOREM 2.3. *The eigenvalues of the matrix*

$$\mathbf{K} = (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S} - k\mathbf{S}^-\mathbf{S} \tag{2.46}$$

do not depend on the choices of generalized inverses $(\mathbf{X}_1'\mathbf{X}_1)^-$ *and* $\mathbf{S}^-$, *and are*

$$\left.\begin{aligned} 0 &\quad \textit{with multiplicity } u + p_1 - q_1, \\ 1 - k &\quad \textit{with multiplicity } q_1 - t - u, \\ 1 - k - \varrho_h^2; &\quad h = 1, \ldots, t, \end{aligned}\right\} \tag{2.47}$$

where the ϱ_h *are the canonical correlations between* $\mathbf{X}_1'\mathbf{y}$ *and* $\mathbf{X}_2'\mathbf{y}$ *in the general partitioned linear model* (2.25), *cf. also* (2.29) *and* (2.30).

Proof. The characteristic polynomial of $\mathbf{K}$ may be written as

$$\begin{aligned} c(\lambda) &= \det(\lambda\mathbf{I} - \mathbf{K}) = \det[\lambda\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S} + k\mathbf{S}^-\mathbf{S}] \\ &= \det[\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}(\lambda\mathbf{I} + k\mathbf{S}^-\mathbf{S})^{-1}] \cdot \det(\lambda\mathbf{I} + k\mathbf{S}^-\mathbf{S}) \end{aligned} \tag{2.48}$$

for all nonzero $\lambda \neq -k$. Since $\mathbf{S}^-\mathbf{S}$ is idempotent with the same rank as the rank of $\mathbf{S} = (\mathbf{X}'\mathbf{X}/\mathbf{X}_2'\mathbf{X}_2)$, i.e., $q - q_2$, and since

$$(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}(\lambda\mathbf{I} + k\mathbf{S}^-\mathbf{S})^{\pm 1} = (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}(\lambda + k)^{\pm 1}; \quad \lambda \neq 0 \text{ and } \lambda \neq -k, \tag{2.49}$$

we obtain

$$\begin{aligned} c(\lambda) &= \det[\mathbf{I} - (\lambda + k)^{-1}(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}] \cdot \lambda^{p_1 - q + q_2}(\lambda + k)^{q - q_2} \\ &= \det[\mu\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}] \cdot \lambda^{p_1 - q + q_2} \cdot \mu^{q - q_2 - p_1}, \end{aligned} \tag{2.50}$$

where, to ease the notation, we put

$$\mu = \lambda + k. \tag{2.51}$$

The characteristic polynomial of $(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}$ may be written as

$$\begin{aligned} d(\mu) &= \det[\mu\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}] \\ &= \det[\mu\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{X}_1 + (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1], \end{aligned} \tag{2.52}$$

since

$$\begin{aligned} \mathbf{S} &= \mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1 = \mathbf{X}_1'\mathbf{X}_1 - \mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{X}_1 \\ &= \mathbf{X}_1'\mathbf{X}_1 - \mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1, \end{aligned} \tag{2.53}$$

say, cf. (2.38) and (2.39). Since the matrix $(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{X}_1$ is idempotent and $\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{X}_1 = \mathbf{X}_1$, we obtain for all nonzero $\mu \neq 1$,

$$(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1[\mu\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{X}_1]^{\pm 1} = (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1(\mu - 1)^{\pm 1} \tag{2.54}$$

and

$$\begin{aligned} d(\mu) &= \det\{\mathbf{I} + (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1[\mu\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{X}_1]^{-1}\} \cdot \det[\mu\mathbf{I} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{X}_1] \\ &= \det[\mathbf{I} + (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1(\mu - 1)^{-1}] \cdot \mu^{p_1 - q_1} \cdot (\mu - 1)^{q_1} \\ &= \det[(\mu - 1)\mathbf{I} + (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1] \cdot \mu^{p_1 - q_1} \cdot (\mu - 1)^{q_1 - p_1}. \end{aligned} \tag{2.55}$$

The nonzero eigenvalues of $(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{X}_1$ are the squares of the canonical correlations ϱ_h between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$, cf. Theorem 2.1(a), since $\mathbf{H}_2 = \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'$, cf. (2.53). Hence, using (2.29) and (2.30), we obtain

$$\begin{aligned} d(\mu) &= (\mu - 1)^{p_1 - t - u} \cdot \Pi_1^t(\mu - 1 + \varrho_h^2) \cdot \mu^{u + p_1 - q_1} \cdot (\mu - 1)^{q_1 - p_1} \\ &= (\mu - 1)^{q_1 - t - u} \cdot \mu^{u + p_1 - q_1} \cdot \Pi_1^t(\mu - 1 + \varrho_h^2), \end{aligned} \tag{2.56}$$

and so, from (2.50),

$$c(\lambda) = (\mu - 1)^{q_1 - t - u} \cdot \mu^{u + p_1 - q_1} \cdot \Pi_1^t(\mu - 1 + \varrho_h^2) \cdot \lambda^{p_1 - q + q_2} \cdot \mu^{q + q_2 - p_1}$$
$$= \lambda^{u + p_1 - q_1} \cdot (\lambda + k - 1)^{q_1 - t - u} \cdot \Pi_1^t(\lambda + k - 1 + \varrho_h^2), \quad (2.57)$$

since $u - q_1 + q - q_2 = 0$, cf. (2.37). This completes our proof. Q.E.D.

In the two-way layout $p_1 = q_1 = r$; putting $k = 1 - \varrho_1^2$ turns our Theorem 2.3 into Theorem 3 of Latour and Styan (1985).

2.3 Canonical correlations and testing the hypothesis that some parameters are zero

In this section we will consider testing a hypothesis about a proper subset of the parameters in the general linear model. Without loss of generality we may consider testing the hypothesis:

$$H_0: \mathscr{E}(\mathbf{y}) = \mathbf{X}_2\gamma \quad (2.58)$$

in the model (2.25)

$$\mathscr{E}(\mathbf{y}) = \mathbf{X}_1\alpha + \mathbf{X}_2\gamma = \mathbf{X}\beta. \quad (2.59)$$

The hypothesis H_0 is not necessarily equivalent to the hypothesis:

$$H_0^*: \alpha = \mathbf{0}, \quad (2.60)$$

which is said to be *completely testable* whenever (cf. Roy and Roy, 1959; Alalouf and Styan, 1979a)

$$\text{rank}(\mathbf{X}) - \text{rank}(\mathbf{X}_2) = p_1, \quad (2.61)$$

the number of parameters in H_0^*. The number u of canonical correlations that are equal to one between the vectors $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$ is equal to, cf. (2.37),

$$u = \text{rank}(\mathbf{X}_1) + \text{rank}(\mathbf{X}_2) - \text{rank}(\mathbf{X})$$
$$= q_1 + q_2 - q. \quad (2.62)$$

It follows, therefore, that H_0^* is completely testable whenever

$$p_1 = q - q_2 = q_1 - u \leq p_1 - u \leq p_1 \quad (2.63)$$

since the rank of $\mathbf{X}_1$ cannot exceed the number of its columns, and $u \geq 0$. The inequality string (2.63) collapses and we find that H_0^* is completely testable if and only if

$$p_1 = q_1 \quad \text{and} \quad u = 0, \quad (2.64)$$

i.e., the matrix $\mathbf{X}_1$ has full column rank and there are no unit canonical correlations between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$, cf. Dahan and Styan (1977), Hemmerle

(1979). The hypotheses H_0 and H_0^* may be considered as equivalent whenever (2.64) holds. When (2.64) does not hold the hypothesis H_0^* is said to be *partly testable* and the hypothesis H_0 is said to be the *testable part* of H_0^* provided

$$\text{rank}(\mathbf{X}) - \text{rank}(\mathbf{X}_2) > 0. \tag{2.65}$$

When $\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}_2)$ the hypotheses H_0 and H_0^* are said to be *completely untestable* (cf. Alalouf and Styan, 1979b).

We will suppose, therefore, that

$$p_1 \geq \text{rank}(\mathbf{X}) - \text{rank}(\mathbf{X}_2) = q - q_2 \geq 1. \tag{2.66}$$

The usual numerator sum of squares in the F-test of the hypothesis H_0 in (2.58) may be written as:

$$\begin{aligned} S_h &= \mathbf{y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{y} \\ &= \mathbf{y}'\mathbf{M}_2\mathbf{X}_1(\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{M}_2\mathbf{y}, \end{aligned} \tag{2.67}$$

where $\mathbf{M}_2 = \mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'$, cf. (2.39). Following Latour and Styan (1985) let us consider also the sum of squares:

$$S_h^* = \mathbf{y}'\mathbf{M}_2\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{M}_2\mathbf{y} \tag{2.68}$$

formed from S_h by omitting the $\mathbf{M}_2$ in the middle. The sum of squares S_h^* may be easier to compute than S_h, e.g., when $\mathbf{X}_1'\mathbf{X}_1$ is diagonal and positive definite, which is so when the α_i identify row effects in the analysis of variance (the γ_j could identify column effects as in the two-way layout or covariates as in the analysis of covariance).

THEOREM 2.4. *The sums of squares S_h and S_h^* defined by* (2.67) *and* (2.68), *respectively, satisfy the inequality string:*

$$S_h^* \leq S_h \leq S_h^*/(1 - \varrho_1^2), \tag{2.69}$$

where ϱ_1 is the largest positive canonical correlation between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$ that is less than 1.

Equality holds on the left of (2.69), *with probability one, if and only if there are no positive canonical correlations between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$ that are less than* 1, *and then equality holds throughout* (2.69).

Equality holds on the right of (2.69), *with probability one, if and only if either* (a)

$$\text{rank}(\mathbf{X}_1'\mathbf{X}_2) = \text{rank}(\mathbf{X}_1) \quad \textit{and} \quad \varrho_1 = \varrho_2 = \ldots = \varrho_t \tag{2.70}$$

or (b) *there are no positive canonical correlations between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$ that are less than* 1.

Proof. The inequality on the left of (2.69) holds, with probability one, if and only if the matrix

$$\mathbf{A}_1 = \mathbf{M}_2\mathbf{X}_1\mathbf{S}^-\mathbf{X}_1'\mathbf{M}_2 - \mathbf{M}_2\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{M}_2 \geq \mathbf{0}, \tag{2.71}$$

where the Schur complement $\mathbf{S} = \mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1 = (\mathbf{X}'\mathbf{X}/\mathbf{X}_2'\mathbf{X}_2)$, cf. (2.38). If we move the matrix factor $\mathbf{M}_2\mathbf{X}_1$ from the front of $\mathbf{A}_1$ to the back, then it follows, using (1.22), that the nonzero eigenvalues of $\mathbf{A}_1$ coincide with the nonzero eigenvalues of the matrix

$$\mathbf{A}_2 = \mathbf{S}^-\mathbf{S} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}. \tag{2.72}$$

From Theorem 2.3 with $k = 1$ we see that the nonzero eigenvalues of $\mathbf{A}_2$ are ϱ_h^2; $h = 1, \ldots, t$, and so (2.71) holds. Equality will hold on the left of (2.69) if and only if $\varrho_1 = 0$ and then equality holds throughout (2.69).

To establish the inequality on the right in (2.69) it suffices to show that

$$\mathbf{A}_3 = \mathbf{M}_2\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{M}_2 - (1 - \varrho_1^2)\mathbf{M}_2\mathbf{X}_1\mathbf{S}^-\mathbf{X}_1'\mathbf{M}_2 \geq \mathbf{0}. \tag{2.73}$$

But $\mathbf{A}_3$ has the same nonzero eigenvalues as does the matrix

$$\mathbf{A}_4 = (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S} - (1 - \varrho_1^2)\mathbf{S}^-\mathbf{S} \tag{2.74}$$

and these eigenvalues are ϱ_1^2 with multiplicity $q_1 - t - u$ and $\varrho_1^2 - \varrho_h^2$; $h = 2, \ldots, t$ (we put $k = 1 - \varrho_1^2$ in Theorem 2.3). Equality will hold on the right of (2.69), therefore, if and only if $q_1 = t + u$ and $\varrho_1 = \ldots = \varrho_t$, or $\varrho_1 = 0$. Q.E.D.

As an example to illustrate Theorem 2.4, let $\mathbf{X}_1 = \mathbf{e}$, the $n \times 1$ vector with each element equal to one, and so, with $\mathbf{X}_2 = \mathbf{X}$, say, we have the usual multiple regression model with intercept. [Latour and Styan (1985, Section 3) provide another example to illustrate Theorem 2.4 involving a two-way layout with one observation in each but one of the cells.]

Then

$$\varrho_1^2 = \mathbf{e}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{e}/n, \tag{2.75}$$

while

$$S_h = (\mathbf{y}'\mathbf{M}\mathbf{e})^2/\mathbf{e}'\mathbf{M}\mathbf{e} \tag{2.76}$$

and

$$S_h^* = (\mathbf{y}'\mathbf{M}\mathbf{e})^2/n, \tag{2.77}$$

where $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ $(= \mathbf{M}_2)$. To see that $S_h^* \leq S_h$ we note that $\mathbf{e}'\mathbf{M}\mathbf{e} \leq n$ since $n - \mathbf{e}'\mathbf{M}\mathbf{e} = \mathbf{e}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}']\mathbf{e} \geq 0$, as $\mathbf{I} - \mathbf{M} = \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ is symmetric and idempotent. Moreover $S_h^*/S_h = \mathbf{e}'\mathbf{M}\mathbf{e}/n = 1 - \varrho_1^2$, and so $S_h = S_h^*/(1 - \varrho_1^2)$.

Another example which illustrates Theorem 2.4 is the balanced incomplete block (BIB) design, with r treatments (rows) and c blocks (columns). Then

$$\mathbf{X}_1'\mathbf{X}_1 = \mathbf{D}_r = s\mathbf{I}_r, \tag{2.78}$$

since a BIB design is »equireplicate» — each of the r treatments is replicated s times — and

$$\mathbf{X}_2'\mathbf{X}_2 = \mathbf{D}_c = k\mathbf{I}_c, \tag{2.79}$$

so that the design is »proper» or »equiblock» — in each of the c blocks there are k treatments — while

$$\mathbf{X}_1'\mathbf{X}_2 = \mathbf{N}, \tag{2.80}$$

the incidence matrix, satisfies

$$\mathbf{N}\mathbf{N}' = (s - \lambda)\mathbf{I}_r + \lambda\mathbf{e}\mathbf{e}', \tag{2.81}$$

where $\mathbf{e}$ is the $r \times 1$ vector with each element equal to 1. From the equality of the off-diagonal elements in (2.81) we see that each pair of treatments appears in the same block λ times.

The canonical correlations between the treatment totals and the block totals (row totals and column totals) are the positive square roots of the eigenvalues of

$$\mathbf{P}_r = \mathbf{D}_r^{-1}\mathbf{N}\mathbf{D}_c^{-1}\mathbf{N}' = \frac{1}{sk}[(s - \lambda)\mathbf{I}_r + \lambda\mathbf{e}\mathbf{e}']. \tag{2.82}$$

Postmultiplying $\mathbf{P}_r$ by $\mathbf{e}$ yields the simple (as we shall see, provided $k > 1$) eigenvalue of unity (and so $u = 1$) and

$$\lambda = \frac{s(k - 1)}{r - 1}. \tag{2.83}$$

The other $r - 1$ eigenvalues of $\mathbf{P}_r$ are

$$\varrho_h^2 = \frac{s - \lambda}{sk} = \frac{r - k}{k(r - 1)} \tag{2.84}$$

say, and so $\mathbf{X}_1'\mathbf{X}_2 = \mathbf{N}$ has full row rank $r = \text{rank}(\mathbf{X}_1)$, $t = r - 1$ canonical correlations are equal (and less than 1 provided $k > 1$), and condition (2.70) of Theorem 2.4 holds. The quantity $1 - \varrho^2 = r(k - 1)/[k(r - 1)]$ is called the »efficiency factor» of the BIB design (cf. James and Wilkinson, 1971).

In our development in this section we have seen that in general there are u canonical correlations between $\mathbf{X}_1'\mathbf{y}$ and $\mathbf{X}_2'\mathbf{y}$ that are equal to one, where, cf. (2.37) and (2.62),

$$u = \text{rank}(\mathbf{X}_1) + \text{rank}(\mathbf{X}_2) - \text{rank}(\mathbf{X}). \tag{2.85}$$

Latour and Styan (1985, Theorem 2) showed that by »adjusting» the vectors of row and column totals, $\mathbf{X}_1'\mathbf{y} = \mathbf{y}_{rt}$ and $\mathbf{X}_2'\mathbf{y} = \mathbf{y}_{ct}$ to vectors $\mathbf{z}_r = \mathbf{y}_{rt} - \mathbf{N}\mathbf{D}_c^{-1}\mathbf{y}_{ct}$ and $\mathbf{z}_c = \mathbf{y}_{ct} - \mathbf{N}'\mathbf{D}_r^{-1}\mathbf{y}_{rt}$, respectively, these u unit canonical correlations disappear while all the other canonical correlations (less than 1) remain intact. [The matrices $\mathbf{N}$, $\mathbf{D}_r$, and $\mathbf{D}_c$ are defined at (2.26).]

With the general partitioned linear model (2.25) we define

$$\mathbf{z}_1 = \mathbf{X}_1'\mathbf{M}_2\mathbf{y} \quad \text{and} \quad \mathbf{z}_2 = \mathbf{X}_2'\mathbf{M}_1\mathbf{y} \tag{2.86}$$

where

$$\mathbf{M}_i = \mathbf{I} - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^-\mathbf{X}_i'; \quad i = 1,2. \tag{2.87}$$

With the white noise assumption (2.27), therefore, the joint covariance matrix of $\mathbf{z}_1$ and $\mathbf{z}_2$ is

$$\mathscr{V}ar\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix} = \sigma^2\begin{pmatrix}\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1 & \mathbf{X}_1'\mathbf{M}_2\mathbf{M}_1\mathbf{X}_2\\ \mathbf{X}_2'\mathbf{M}_1\mathbf{M}_2\mathbf{X}_1 & \mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2\end{pmatrix}; \tag{2.88}$$

the cross-covariance matrix between $\mathbf{z}_1$ and $\mathbf{z}_2$ may be written as

$$\sigma^2\mathbf{X}_1'\mathbf{M}_2\mathbf{M}_1\mathbf{X}_2 = -\sigma^2\mathbf{X}_1'\mathbf{H}_2\mathbf{M}_1\mathbf{X}_2 = -\sigma^2\mathbf{X}_1'\mathbf{M}_2\mathbf{H}_1\mathbf{X}_2 \tag{2.89}$$

where

$$\mathbf{H}_i = \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^-\mathbf{X}_i' = \mathbf{I} - \mathbf{M}_i; \quad i = 1,2. \tag{2.90}$$

We then obtain:

THEOREM 2.5. *The canonical correlations between the vectors* $\mathbf{z}_1$ *and* $\mathbf{z}_2$, *defined by* (2.86), *are all less than* 1 *and are precisely the* t *positive canonical correlations* ϱ_h *between the vectors* $\mathbf{X}_1'\mathbf{y}$ *and* $\mathbf{X}_2'\mathbf{y}$ *that are not equal to* 1.

Proof. Using Theorem 2.1(a) we are concerned with the eigenvalues of the matrix

$$\begin{aligned}\mathbf{B} &= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{M}_2\mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{M}_1\mathbf{M}_2\mathbf{X}_1\\ &= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{M}_1\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_1\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{M}_1\mathbf{H}_2\mathbf{X}_1\\ &= (\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1)^-\mathbf{X}_1'\mathbf{H}_2\mathbf{M}_1\mathbf{H}_2\mathbf{X}_1\end{aligned} \tag{2.91}$$

using (2.89) and (2.90). Using (2.90) again and writing $\mathbf{S} = \mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1$ we obtain

$$\begin{aligned}\mathbf{B} &= \mathbf{S}^-\mathbf{X}_1'(\mathbf{I} - \mathbf{M}_2)\mathbf{M}_1(\mathbf{I} - \mathbf{M}_2)\mathbf{X}_1\\ &= \mathbf{S}^-\mathbf{X}_1'\mathbf{M}_2\mathbf{M}_1\mathbf{M}_2\mathbf{X}_1\\ &= \mathbf{S}^-\mathbf{X}_1'\mathbf{M}_2(\mathbf{I} - \mathbf{H}_1)\mathbf{M}_2\mathbf{X}_1\\ &= \mathbf{S}^-\mathbf{S} - \mathbf{S}^-\mathbf{S}(\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S},\end{aligned} \tag{2.92}$$

since $\mathbf{X}_1'\mathbf{M}_1 = \mathbf{0}$. Moving $\mathbf{S}^-\mathbf{S}$ from the front to the back of $\mathbf{B}$ shows that the nonzero eigenvalues of $\mathbf{B}$ coincide with those of

$$\mathbf{A}_2 = \mathbf{S}^-\mathbf{S} - (\mathbf{X}_1'\mathbf{X}_1)^-\mathbf{S}, \tag{2.93}$$

cf. (2.72), and these eigenvalues are ϱ_h^2 ; $h = 1, \ldots, t$. Q.E.D.

We may interpret Theorem 2.5 in the following way. Suppose that the random vector

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \tag{2.94}$$

has covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \geq \mathbf{0}, \tag{2.95}$$

cf. (1.7) and (1.8). Then the canonical correlations between the two »residual» vectors

$$\mathbf{x}_{1.2} = \mathbf{x}_1 - \Sigma_{12}\Sigma_{22}^-\mathbf{x}_2 \tag{2.96}$$

and

$$\mathbf{x}_{2.1} = \mathbf{x}_2 - \Sigma_{21}\Sigma_{11}^-\mathbf{x}_1, \tag{2.97}$$

cf. (1.9), are all less than 1 and are precisely the positive canonical correlations between $\mathbf{x}_1$ and $\mathbf{x}_2$ that are not equal to 1.

3. BROYDEN'S MATRIX PROBLEM AND AN ASSOCIATED ANALYSIS-OF-COVARIANCE MODEL

3.1 Broyden's matrix problem and his solution

In the »Problems and Solutions» section of *SIAM Review*, Broyden (1982, 1983) posed and solved »A Matrix Problem» about the inertia of a certain matrix which we will show to be a Schur complement associated with a particular analysis-of-covariance linear statistical model.

Let the $r \times c$ matrix $\mathbf{Z} = \{z_{ij}\}$ have no null rows and no null columns, and let the $r \times c$ binary incidence matrix $\mathbf{N} = \{n_{ij}\}$ be defined by

$$\left.\begin{aligned} n_{ij} = 1 \Leftrightarrow z_{ij} \neq 0 \\ n_{ij} = 0 \Leftrightarrow z_{ij} = 0 \end{aligned}\right\} i = 1, \ldots, r;\ j = 1, \ldots, c. \tag{3.1}$$

Let $n_{i.} = \Sigma_j n_{ij}$ and $n_{.j} = \Sigma_i n_{ij}$ ($i = 1, \ldots, r$; $j = 1, \ldots, c$) as in Section 2.2, and let the $r \times r$ diagonal matrix

$$\mathbf{D}_r = \operatorname{diag}\{n_{i.}\} \tag{3.2}$$

as in (2.26). Introduce the $c \times c$ diagonal matrix

$$\mathbf{D}_z = \operatorname{diag}(\mathbf{Z}'\mathbf{Z}) = \operatorname{diag}\{\sum_{k=1}^{r} z'_{jk} z_{kj}\} = \operatorname{diag}\{\sum_{k=1}^{r} z_{kj}^2\}. \tag{3.3}$$

When $\mathbf{Z} \equiv \mathbf{N}$ then $\mathbf{D}_z \equiv \mathbf{D}_c$ as in (2.26).

Broyden (1982) sought »conditions under which the matrix

$$\mathbf{B} = \mathbf{D}_z - \mathbf{Z}'\mathbf{D}_r^{-1}\mathbf{Z} \tag{3.4}$$

is (a) positive definite, (b) positive semidefinite. This problem arose in connection with an algorithm for scaling examination marks».

In his solution, Broyden (1983) established the nonnegative definiteness of the matrix $\mathbf{B}$ in (3.4) from the nonnegativity of the quadratic form

$$\mathbf{u}'\mathbf{B}\mathbf{u} = \mathbf{e}'\mathbf{D}_u(\mathbf{D}_z - \mathbf{Z}'\mathbf{D}_r^{-1}\mathbf{Z})\mathbf{D}_u\mathbf{e} = \sum_{i=1}^{r} \mathbf{e}_i'\mathbf{Z}\mathbf{D}_u\mathbf{E}_i\mathbf{D}_u\mathbf{Z}'\mathbf{e}_i, \tag{3.5}$$

where the $c \times c$ matrix

$$\mathbf{E}_i = \mathbf{I} - (n_{i.})^{-1}\mathbf{N}'\mathbf{e}_i\mathbf{e}_i'\mathbf{N} \tag{3.6}$$

is symmetric idempotent with rank equal to $c - 1$ (for all $i = 1, \ldots, r$). In (3.5) $\mathbf{D}_u = \operatorname{diag}(\mathbf{u})$ is the $c \times c$ diagonal matrix formed from the $c \times 1$ vector $\mathbf{u}$, while $\mathbf{e}$ is the $c \times 1$ vector with each element equal to 1; the vector $\mathbf{e}_i$ is the $c \times 1$ vector with its ith element equal to 1 and the rest zero.

In addition, Broyden (1983) showed that his matrix $\mathbf{B}$, as defined by (3.4), is positive semidefinite when there exist scalars $a_1, \ldots, a_r, u_1, \ldots, u_c$, all nonzero, so that

$$z_{ij}u_j = a_i n_{ij} \quad \text{for all } i = 1, \ldots, r \text{ and all } j = 1, \ldots, c. \tag{3.7}$$

He also stated that $\mathbf{B}$ is positive definite unless there exist scalars $a_1, \ldots, a_r, u_1, \ldots, u_c$, with at least one of the u_j's nonzero, so that (3.7) holds. [At least one of the a_i's must then also be nonzero when (3.7) holds or else $\mathbf{Z}$ would have to have a null column.]. These conditions do not, however, completely characterize the singularity of Broyden's matrix $\mathbf{B}$. Moreover, Broyden does not mention the rank (or nullity) of $\mathbf{B}$.

We will solve Broyden's matrix problem by constructing an analysis-of-covariance linear statistical model in which the matrix $\mathbf{B}$ in (3.4) arises naturally as a Schur complement. This will enable us to completely characterize the rank of $\mathbf{B}$ from the structure of the matrix $\mathbf{Z}$ and its associated binary incidence matrix $\mathbf{N}$. When $\mathbf{Z} \equiv \mathbf{N}$ our analysis-of-covariance model reduces to the usual two-way layout as considered by Latour and Styan (1985), see also Section 2.2.

3.2 An associated analysis-of-covariance model

Consider the linear statistical model defined by

$$\mathscr{E}(y_{ij}) = \alpha_i n_{ij} + z_{ij}\gamma_j \quad (i = 1, \ldots, r; j = 1, \ldots, c), \tag{3.8}$$

where the n_{ij} are as defined by (3.1) — (0,1)-indicators of the z_{ij} — and so the n_{ij} and z_{ij} are zero only simultaneously and then the corresponding observation y_{ij} has zero mean [we could just as well have replaced y_{ij} in (3.8) with $n_{ij}y_{ij}$ and then the (i,j)th cell of the $r \times c$ layout would be missing whenever $n_{ij} = z_{ij} = 0$; such y_{ij} play no role in what follows].

The observations y_{ij} in (3.8) may be arranged in a two-way layout with r rows and c columns. The α_i may be taken to represent »row effects», but the »column effects» in the usual two-way layout (cf. e.g., Latour and Styan, 1985, and our Section 2.2) are here replaced by »regression coefficients» γ_j on each of c »covariates» on each of which we have (at most) r observations. This is the analysis-of-covariance model considered, for example, by Scheffé (1959, page 200); in many analysis-of-covariance models, however, the γ_j are all taken to be equal (to γ, say).

We may rewrite (3.8) as

$$\mathscr{E}(\mathbf{y}_j) = \mathbf{Q}_j\alpha + \gamma_j\mathbf{z}_j \quad (j = 1, \ldots, c), \tag{3.9}$$

where the $r \times 1$ vectors $\alpha = \{\alpha_i\}$, $\mathbf{y}_j = \{y_{ij}\}$ and $\mathbf{z}_j = \{z_{ij}\}$. The $r \times r$ diagonal matrix

$$\mathbf{Q}_j = \text{diag}\{n_{ij}\} = \text{diag}(n_{1j}, \ldots, n_{rj}) \tag{3.10}$$

is symmetric idempotent with rank equal to $n_{.j}$ $(j = 1, \ldots, c)$. Moreover

$$\mathbf{Q}_j\mathbf{z}_j = \mathbf{z}_j \quad (j = 1, \ldots, c) \tag{3.11}$$

and, cf. (3.2),

$$\Sigma_1^c\mathbf{Q}_j = \text{diag}\{n_{i.}\} = \mathbf{D}_r. \tag{3.12}$$

We may then write (3.9) and (3.8) as

$$\mathscr{E}(\mathbf{y}) = \mathbf{X}_1\alpha + \mathbf{X}_2\gamma = \mathbf{X}\beta, \tag{3.13}$$

cf. (2.25), where

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \ldots \\ \mathbf{y}_c \end{pmatrix}, \quad \mathbf{X}_1 = \begin{pmatrix} \mathbf{Q}_1 \\ \ldots \\ \mathbf{Q}_c \end{pmatrix}, \quad \mathbf{X}_2 = \begin{pmatrix} \mathbf{z}_1 & & & \\ & \mathbf{z}_2 & & \\ & & \ldots & \\ & & & \mathbf{z}_c \end{pmatrix}, \tag{3.14}$$

and

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2), \quad \beta = \begin{pmatrix}\alpha\\ \gamma\end{pmatrix}. \tag{3.15}$$

Then

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix}\mathbf{D}_r & \mathbf{Z}\\ \mathbf{Z}' & \mathbf{D}_z\end{pmatrix}, \tag{3.16}$$

and so Broyden's matrix as defined in (3.4) is

$$\mathbf{B} = (\mathbf{X}'\mathbf{X}/\mathbf{D}_r) = \mathbf{D}_z - \mathbf{Z}'\mathbf{D}_r^{-1}\mathbf{Z}, \tag{3.17}$$

the Schur complement of $\mathbf{D}_r$ in $\mathbf{X}'\mathbf{X}$. The see that (3.16) follows from (3.13) we note that

$$\mathbf{X}_1'\mathbf{X}_1 = \Sigma_1^c\mathbf{Q}_j'\mathbf{Q}_j = \Sigma_1^c\mathbf{Q}_j = \mathbf{D}_r \tag{3.18}$$

using $\mathbf{Q}_j' = \mathbf{Q}_j = \mathbf{Q}_j^2$ and (3.12). Moreover

$$\mathbf{X}_1'\mathbf{X}_2 = (\mathbf{Q}_1'\mathbf{z}_1, \ldots, \mathbf{Q}_c'\mathbf{z}_c) = \mathbf{Z}, \tag{3.19}$$

since $\mathbf{Q}_j'\mathbf{z}_j = \mathbf{Q}_j\mathbf{z}_j = \mathbf{z}_j$, cf. (3.11), while

$$\mathbf{X}_2'\mathbf{X}_2 = \mathrm{diag}\{\mathbf{z}_j'\mathbf{z}_j\} = \mathrm{diag}(\mathbf{Z}'\mathbf{Z}) = \mathbf{D}_z, \tag{3.20}$$

cf. (3.3).

3.3 Nonnegative definiteness

Let u denote the nullity of the $(r + c) \times (r + c)$ matrix $\mathbf{X}'\mathbf{X}$ defined by (3.16). Then using Haynsworth's inertia formula (1.14) we have

$$\begin{aligned}\mathrm{In}\,\mathbf{B} &= \mathrm{In}(\mathbf{X}'\mathbf{X}/\mathbf{D}_r) = \mathrm{In}(\mathbf{X}'\mathbf{X}) - \mathrm{In}\,\mathbf{D}_r\\ &= \{r + c - u, 0, u\} - \{r, 0, 0\}\\ &= \{c - u, 0, u\}\end{aligned} \tag{3.21}$$

and so Broyden's matrix $\mathbf{B}$ is nonnegative definite with nullity

$$u = \nu(\mathbf{B}) = \nu(\mathbf{X}'\mathbf{X}) = \nu(\mathbf{X}), \tag{3.22}$$

the number of unit canonical correlations between the $r \times 1$ vector of row totals $\mathbf{y}_{rt} = \mathbf{X}_1'\mathbf{y} = \{y_{i.}\}$ and the $c \times 1$ vector of »weighted» column totals $\mathbf{y}_{ct}^{(z)} = \mathbf{X}_2'\mathbf{y} = \{\Sigma_i^r y_{ij}z_{ij}\}$, cf. (2.30) and (1.25).

We may also consider the »other» Schur complement

$$\mathbf{S} = (\mathbf{X}'\mathbf{X}/\mathbf{D}_z) = \mathbf{D}_r - \mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}' ; \tag{3.23}$$

if $\mathbf{Z} \equiv \mathbf{N}$ then $\mathbf{S} = \mathbf{S}_r$ as defined by (2.32). Moreover, using (1.25),

$$\nu(\mathbf{S}) = \nu(\mathbf{X}'\mathbf{X}/\mathbf{D}_z) = \nu(\mathbf{X}'\mathbf{X}/\mathbf{D}_r) = u = \nu(\mathbf{B}), \tag{3.24}$$

cf. (3.22). The Schur complement $\mathbf{S}$ is, of course, also nonnegative definite.

3.4 The special case when all the cells are filled

When all the cells are filled, i.e., when

$$z_{ij} \neq 0 \Leftrightarrow n_{ij} = 1 \text{ for } \textit{all } i = 1, \ldots, r \text{ and } \textit{all } j = 1, \ldots, c \tag{3.25}$$

then

$$\mathbf{D}_r = c\mathbf{I}_r \tag{3.26}$$

and the Schur complement

$$\begin{aligned} \mathbf{S} = (\mathbf{X}'\mathbf{X}/\mathbf{D}_z) &= \mathbf{D}_r - \mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}' \\ &= c\mathbf{I}_r - \mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}' = c(\mathbf{I}_r - c^{-1}\mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}'), \end{aligned} \tag{3.27}$$

and

$$u = \nu(\mathbf{S}) \tag{3.28}$$

is the number of unit eigenvalues of $c^{-1}\mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}'$. Its trace, however, is

$$\begin{aligned} \operatorname{tr}(c^{-1}\mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}') &= \operatorname{tr}(\mathbf{Z}'\mathbf{Z}\mathbf{D}_z^{-1})/c \\ &= \operatorname{tr}\{\mathbf{Z}'\mathbf{Z}[\operatorname{diag}(\mathbf{Z}'\mathbf{Z})]^{-1}\}/c = 1 \end{aligned} \tag{3.29}$$

and since $\mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}' \geq \mathbf{0}$ it follows at once that

$$\left.\begin{aligned} u \geq 1 \Leftrightarrow u = 1 &\Leftrightarrow \operatorname{rank}(\mathbf{Z}) = 1 \\ u = 0 &\Leftrightarrow \operatorname{rank}(\mathbf{Z}) > 1 \end{aligned}\right\} \tag{3.30}$$

and so, *when all the cells are filled,* Broyden's matrix **B** is

$$\text{positive} \left\{\begin{aligned} &\text{definite and } \operatorname{rank}(\mathbf{B}) = c \Leftrightarrow \operatorname{rank}(\mathbf{Z}) > 1 \\ &\text{semidefinite and singular} \Leftrightarrow \operatorname{rank}(\mathbf{Z}) = 1 \\ &\qquad\qquad \Leftrightarrow \operatorname{rank}(\mathbf{B}) = c - 1. \end{aligned}\right\} \tag{3.31}$$

3.5 The general case when at least one of the cells is empty

When at least one of the cells is empty, i.e., when

$$z_{ij} = 0 \Leftrightarrow n_{ij} = 0 \text{ for } \textit{at least one } i = 1, \ldots, r \text{ and } \textit{at least one } j = 1, \ldots, c, \tag{3.32}$$

then the characterization of the positive (semi)definiteness of Broyden's matrix **B** is much more complicated than when all the cells are filled, cf. (3.30) and (3.31).

3.5.1 A necessary condition for positive definiteness

Using (3.27) we may write

$$\mathbf{S} = (\mathbf{Z}'\mathbf{Z}/\mathbf{D}_z) = \mathbf{D}_r - \mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}' = \Sigma_1^c(\mathbf{Q}_j - \mathbf{z}_j\mathbf{z}_j'/\mathbf{z}_j'\mathbf{z}_j) = \Sigma_1^c\mathbf{G}_j, \tag{3.33}$$

say, using (3.12), (3.19) and (3.20). In (3.33) the matrix $\mathbf{G}_j$ is symmetric and idempotent with rank equal to rank$(\mathbf{Q}_j) - 1 = n_{.j} - 1$ $(j = 1, \ldots, c)$.

When the $r \times r$ Schur complement $\mathbf{S}$ is positive definite,

$$r = \text{rank}(\mathbf{S}) = \text{rank}(\Sigma_1^c\mathbf{G}_j) \leq \Sigma_1^c\text{rank}(\mathbf{G}_j) = \Sigma_1^c(n_{.j} - 1) = n_{..} - c, \tag{3.34}$$

where $n_{..}$ is the number of »filled» cells. The inequality (3.34) then shows that a *necessary condition* for Broyden's matrix $\mathbf{B}$ to be positive definite is that there be at least $r + c$ filled cells. [Recall that $\nu(\mathbf{S}) = \nu(\mathbf{B})$, cf. (3.24).]

3.5.2 Necessary and sufficient conditions when the layout is connected

We will divide our presentation of necessary and sufficient conditions for the nullity of Broyden's matrix $\mathbf{B}$ to be a particular number u into the two situations when the layout is either (a) connected, or (b) not connected. In this subsection we will suppose that the layout is connected; the situation when the layout is not connected will be discussed in subsection 3.5.4.

The two-way layout defined by the binary incidence matrix $\mathbf{N}$ is said to be *connected* if, given any two rows i and i', say, it is possible to construct a chain of rows $i = i_1, i_2, \ldots, i_k = i'$, say, such that for each $h = 1, \ldots, k - 1$ there exists a column $j = j_h$ so that the product

$$n(i_h, j_h) \cdot n(i_{h+1}, j_h) = 1 \tag{3.35}$$

for all $h = 1, \ldots, k - 1$, where (to ease the notation) we have written $n(i,j)$ for n_{ij}, the (i,j)th element of the binary incidence matrix $\mathbf{N}$. [Cf. Raghavarao (1971, page 49) and Latour and Styan (1985, page 228).]

It follows at once from this definition that the layout is, therefore, *not connected* if the rows and columns of the $r \times c$ incidence matrix $\mathbf{N}$ can be arranged so that

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{N}_1 \end{pmatrix}, \tag{3.36}$$

where $\mathbf{N}_0$ is $r_0 \times c_0$ with $1 \leq r_0 < r$ and $1 \leq c_0 < c$, for then the product

$$n(i,j) \cdot n(i',j) = 0 \text{ for all } i = 1, \ldots, r_0 \text{ and all } i' = r_0 + 1, \ldots, r \text{ and all } j = 1, \ldots, c. \tag{3.37}$$

For the rest of this subsection, then, we will suppose that the layout *is* connected; this is characterized (cf. Raghavarao, 1971, page 50) by the nullity

$$\nu(\mathbf{S}_r) = \nu(\mathbf{D}_r - \mathbf{N}\mathbf{D}_c^{-1}\mathbf{N}') = 1, \tag{3.38}$$

where $\mathbf{D}_c = \text{diag}\{n_{.j}\}$. The Schur complement $\mathbf{S}_r = \mathbf{D}_r - \mathbf{N}\mathbf{D}_c^{-1}\mathbf{N}'$ is also known as the »C-matrix of experimental design» (cf. e.g., Raghavarao, 1971, page 49).

The Schur complement $\mathbf{S} = \mathbf{D}_r - \mathbf{Z}\mathbf{D}_z^{-1}\mathbf{Z}'$, as defined in (3.33), is positive semidefinite and singular (with rank at most equal to $r - 1$) if and only if there exists a nonnull vector $\mathbf{a}$, say, so that

$$\begin{aligned}\mathbf{S}\mathbf{a} = \mathbf{0} \Leftrightarrow \mathbf{a}'\mathbf{S}\mathbf{a} = 0 \Leftrightarrow \mathbf{a}'\mathbf{G}_j\mathbf{a} &= 0 \text{ for all } j = 1, \ldots, c \\ \Leftrightarrow \quad \mathbf{G}_j\mathbf{a} &= \mathbf{0} \text{ for all } j = 1, \ldots, c,\end{aligned} \tag{3.39}$$

since the $\mathbf{G}_j$ and $\mathbf{S}$ are all nonnegative definite. Hence $\mathbf{S}$ is positive semidefinite and singular if and only if

$$\mathbf{Q}_j\mathbf{a} = \mathbf{z}_j(\mathbf{z}_j'\mathbf{a}/\mathbf{z}_j'\mathbf{z}_j) \quad \text{for all } j = 1, \ldots, c, \tag{3.40}$$

cf. (3.33). Since $\mathbf{z}_j \neq \mathbf{0}$ (there being no null column in $\mathbf{Z}$) we see that

$$\mathbf{Q}_j\mathbf{a} = \mathbf{0} \Leftrightarrow \mathbf{z}_j'\mathbf{a} = 0 \; (j = 1, \ldots, c). \tag{3.41}$$

We may, without loss of generality, assume that the columns of $\mathbf{Z}$ are so arranged that

$$\begin{aligned}\mathbf{z}_j'\mathbf{a} &\neq 0 \text{ for all } j = 1, \ldots, c_0 \\ &= 0 \text{ for all } j = c_0 + 1, \ldots, c.\end{aligned} \tag{3.42}$$

[A rearrangement of the columns of $\mathbf{Z}$ will not effect the rank of $\mathbf{Z}$ or the positive (semi)definiteness of the associated Schur complements.] Then $c_0 \geq 1$ since if $\mathbf{z}_j'\mathbf{a} = 0$ for all $j = 1, \ldots, c$ then from (3.41) we would have $\mathbf{Q}_j\mathbf{a} = \mathbf{0}$ for all $j = 1, \ldots, c$ and so $\Sigma_1^c\mathbf{Q}_j\mathbf{a} = \mathbf{0} = \mathbf{D}_r\mathbf{a}$, which implies that $\mathbf{a} = \mathbf{0}$ since $\mathbf{D}_r > \mathbf{0}$. But we have assumed that $\mathbf{a} \neq \mathbf{0}$ and so $c_0 \geq 1$.

We may then define

$$b_j = \mathbf{z}_j'\mathbf{z}_j/\mathbf{z}_j'\mathbf{a} \neq 0 \quad (j = 1, \ldots, c_0) \tag{3.43}$$

and so from (3.40) we have

$$\mathbf{z}_j = d_j\mathbf{Q}_j\mathbf{a} \quad (j = 1, \ldots, c_0). \tag{3.44}$$

We may, without loss of generality, assume that the rows of $\mathbf{Z}$ are so arranged [cf. comment below (3.42)] that (with $\mathbf{a} = \{a_i\}$)

$$\begin{aligned}a_i &\neq 0 \text{ for all } i = 1, \ldots, r_0 \\ &= 0 \text{ for all } i = r_0 + 1, \ldots, r.\end{aligned} \tag{3.45}$$

Then $r_0 \geq 1$ since $\mathbf{a} \neq \mathbf{0}$. Thus, using the last $r - r_0$ rows of (3.44) we obtain

$$z_{ij} = 0 \text{ for all } i = r_0 + 1, \ldots, r \text{ and all } j = 1, \ldots, c_0. \tag{3.46}$$

Moreover from (3.42) we see that for all $j = c_0 + 1, \ldots, c$, we have

$$\begin{aligned} \mathbf{z}_j'\mathbf{a} = 0 \Leftrightarrow \mathbf{Q}_j\mathbf{a} = \mathbf{0} &\Leftrightarrow a_i n_{ij} = 0 \text{ for all } i = 1, \ldots, r \\ &\Leftrightarrow a_i = 0 \text{ or } n_{ij} = 0 \text{ for all } i = 1, \ldots, r \\ &\Leftrightarrow n_{ij} = 0 \text{ for all } i = 1, \ldots, r_0 \end{aligned} \tag{3.47}$$

using (3.41) and (3.45).

We may, therefore, write the matrix $\mathbf{Z}$ in the partitioned form

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_1 \end{pmatrix} \tag{3.48}$$

where the leading principal submatrix

$$\mathbf{Z}_0 = \{a_i b_j n_{ij}\} \tag{3.49}$$

is $r_0 \times c_0$. We have, however, supposed that the layout is connected, and so the arrangement (3.48) is possible if and only if $r_0 = r$ and $c_0 = c$, cf. (3.36), and so we must have that the (full) $r \times c$ matrix

$$\mathbf{Z} = \{z_{ij}\} = \{a_i b_j n_{ij}\} = \mathbf{D}_a \mathbf{N} \mathbf{D}_b, \tag{3.50}$$

cf. (3.7), where the diagonal matrices $\mathbf{D}_a = \operatorname{diag}\{a_i\}$ and $\mathbf{D}_b = \operatorname{diag}\{b_j\}$ are both nonsingular.

The nullity u of Broyden's matrix $\mathbf{B}$ is equal to

$$u = \nu(\mathbf{B}) = \nu(\mathbf{X}), \tag{3.51}$$

cf. (3.22). We will now prove that $u \geq 1 \Leftrightarrow u = 1$, as in the special case when all the cells are filled, cf. (3.30).

When the matrix $\mathbf{Z}$ has the special form given by (3.50) we may write the jth column of $\mathbf{Z}$,

$$\mathbf{z}_j = b_j \mathbf{D}_a \mathbf{N} \mathbf{e}_j = \mathbf{D}_a \mathbf{Q}_j \mathbf{e} b_j, \quad j = 1, \ldots, c, \tag{3.52}$$

where the $r \times r$ diagonal matrix $\mathbf{Q}_j = \operatorname{diag}\{n_{ij}\}$, cf. (3.10). In (3.52) the $r \times 1$ vector $\mathbf{e}_j$ has each element equal to zero except for the jth which is equal to one, while the $r \times 1$ vector $\mathbf{e}$ has every element equal to one. And so, using (3.15) and (3.14), we may write

$$\begin{aligned} \mathbf{X} &= \begin{pmatrix} \mathbf{Q}_1 & \mathbf{z}_1 & & \\ \cdots & & \cdots & \\ \mathbf{Q}_c & & & \mathbf{z}_c \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{D}_a\mathbf{Q}_1\mathbf{e}b_1 & & \\ \cdots & & \cdots & \\ \mathbf{Q}_c & & & \mathbf{D}_a\mathbf{Q}_c\mathbf{e}b_c \end{pmatrix} \\ &= (\mathbf{I}_c \otimes \mathbf{D}_a) \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_1\mathbf{e} & & \\ \cdots & & \cdots & \\ \mathbf{Q}_c & & & \mathbf{Q}_c\mathbf{e} \end{pmatrix} \begin{pmatrix} \mathbf{D}_a^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_b \end{pmatrix} \\ &= (\mathbf{I}_c \otimes \mathbf{D}_a)\mathbf{X}_0 \begin{pmatrix} \mathbf{D}_a^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_b \end{pmatrix}, \end{aligned} \tag{3.53}$$

say, where $\mathbf{D}_b$ is defined as in (3.50). The factorization (3.53) is possible since $\mathbf{D}_a$ is nonsingular and diagonal (so that $\mathbf{D}_a\mathbf{Q}_j = \mathbf{Q}_j\mathbf{D}_a$).

Hence

$$u = \nu(\mathbf{X}) = \nu(\mathbf{X}_0) = 1 \tag{3.54}$$

since $\mathbf{X}_0$ is the design matrix for the connected layout based on the binary incidence matrix $\mathbf{N}$.

We have, therefore, proved

THEOREM 3.1. *When the layout is connected Broyden's matrix* $\mathbf{B}$ *is positive semidefinite and singular if and only if it is possible to write the matrix*

$$\mathbf{Z} = \{z_{ij}\} = \{a_i b_j n_{ij}\} = \mathbf{D}_a\mathbf{N}\mathbf{D}_b \tag{3.55}$$

for some nonsingular diagonal matrices $\mathbf{D}_a$ *and* $\mathbf{D}_b$, *and then the nullity of* $\mathbf{B}$ *is equal to* 1.

We may interpret the condition (3.55) as follows. There exist nonzero scalars a_i and b_j for each empty cell (i,j) with $z_{ij} = n_{ij} = 0$ so that the matrix $\mathbf{Z}^*$, say, formed from $\mathbf{Z}$ by replacing each empty cell (i,j) with $a_i b_j$ has rank equal to 1. Conversely, there exists a matrix $\mathbf{Z}^*$, say, with rank equal to 1, so that $\mathbf{Z}$ can be formed from $\mathbf{Z}^*$ by changing some of the entries into 0. [We assume that neither $\mathbf{Z}$ nor $\mathbf{Z}^*$ have any null rows or null columns.] When all the cells are filled the condition (3.55) reduces to $\operatorname{rank}(\mathbf{Z}) = 1$, cf. (3.31).

The formula (3.55) in Theorem 3.1 also yields the following *sufficient condition* for positive definiteness of Broyden's matrix $\mathbf{B}$: If there exists in the matrix $\mathbf{Z}$ a 2×2 nonsingular submatrix with all 4 elements nonzero then the matrix $\mathbf{B}$ is positive definite (since then no matrix $\mathbf{Z}^*$ of rank 1 can be constructed from $\mathbf{Z}$, which has rank at least equal to 2).

When (3.55) holds then $\operatorname{rank}(\mathbf{Z}) = \operatorname{rank}(\mathbf{N})$, but this equality of ranks is not sufficient to imply (3.55) when at least one of the cells is empty. When all the cells are filled, however, $\operatorname{rank}(\mathbf{N}) = 1$ and so $\operatorname{rank}(\mathbf{Z}) = \operatorname{rank}(\mathbf{N})$ does imply (3.55), but when at least one of the cells is empty, e.g., when $\mathbf{Z} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 5 \end{pmatrix}$ then $\operatorname{rank}(\mathbf{Z}) = 2 = \operatorname{rank}(\mathbf{N})$, but (3.55) does not hold since the leading 2×2 submatrix in $\mathbf{Z}$ has all 4 elements nonzero and is nonsingular.

3.5.3 A numerical example

To illustrate Theorem 3.1 and the condition (3.55) let us consider the numerical excample with

$$\mathbf{Z} = \begin{pmatrix} 0 & 2 & 3 & 4 \\ 2 & 0 & 6 & 0 \\ 3 & 6 & 0 & 12 \end{pmatrix} \tag{3.56}$$

with $r = 3$ and $c = 4$. The associated layout is connected. Then $n_{..} = 8 > 7 = r + c$ so that **B** might be positive definite, cf. (3.34). But we can form the rank-one matrix

$$\mathbf{Z}^* = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (1, 2, 3, 4) \tag{3.57}$$

from **Z** by replacing its empty cells with appropriate nonzero scalars, and so by Theorem 3.1, Broyden's matrix **B** has nullity one. In fact

$$\mathbf{B} = \frac{1}{3} \begin{pmatrix} 24 & -18 & -18 & -36 \\ -18 & 80 & -6 & -80 \\ -18 & -6 & 72 & -12 \\ -36 & -80 & -12 & 320 \end{pmatrix}, \tag{3.58}$$

while the other Schur complement

$$\mathbf{S} = \frac{1}{65} \begin{pmatrix} 169 & -26 & -39 \\ -26 & 58 & -30 \\ -39 & -30 & 33 \end{pmatrix}. \tag{3.59}$$

Clearly

$$\mathbf{B} \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \\ 1/4 \end{pmatrix} = \mathbf{0} \quad \text{and} \quad \mathbf{S} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \mathbf{0}, \tag{3.60}$$

and so **B** and **S** have nullity at least equal to 1. To confirm that the nullities are both actually equal to 1 we compute the determinant of the leading 3×3 submatrix of **B** to be equal to $3120 \neq 0$, and the determinant of the leading 2×2 submatrix of **S** to be equal to $9126/4225 \neq 0$.

The equations in (3.60) generalize to:

$$\mathbf{B}\mathbf{D}_b^{-1}\mathbf{e} = \mathbf{0} \quad \text{and} \quad \mathbf{S}\mathbf{D}_a\mathbf{e} = \mathbf{S}\mathbf{a} = \mathbf{0}, \tag{3.61}$$

i.e., the vector with elements equal to $1/b_j$ spans the null space of **B**, while the vector with elements a_i spans the null space of **S**.

3.5.4 Necessary and sufficient conditions when the layout is not connected

When the layout associated with the matrix **Z** is not connected, then the rows and columns of **Z** may be arranged so that

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & & & \\ & \mathbf{Z}_2 & & \\ & & \ddots & \\ & & & \mathbf{Z}_l \end{pmatrix}, \tag{3.62}$$

with $l \geq 2$, where the layout corresponding to each submatrix $\mathbf{Z}_h$ ($h = 1, \ldots, l$) is connected (cf. e.g., Searle, 1971, page 320).

Broyden's matrix **B** will then be positive definite if and only if none of the submatrices $\mathbf{Z}_h$ can be expressed in the form (3.55), and positive semidefinite with nullity u ($1 \leq u \leq l$) if and only if precisely u of the submatrices $\mathbf{Z}_h$ can be expressed in the form (3.55).

4. ACKNOWLEDGEMENTS

I am extremely grateful to Tarmo Pukkila and to Simo Puntanen for their most generous invitation to speak at this First International Tampere Seminar on Linear Statistical Models and their Applications.

I am also very grateful to T. W. Anderson, Dominique Latour, Timo Mäkeläinen, Diane V. Ouellette, and Friedrich Pukelsheim, for most helpful discussions concerning various aspects of the research reported in this paper. The final version of this paper was completed while I was on sabbatical leave in the Department of Mathematics and Statistics at The University of Auckland: I am particularly appreciative to Alastair J. Scott for his remarkably generous provision of facilities.

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and by the Gouvernement du Québec, Programme de formation de chercheurs et d'action concertée.

NOTE ADDED IN PROOF

I am very grateful to Professor C.G. Broyden for drawing my attention to his paper "A mark-scaling algorithm", *The Computer Journal,* vol. 26 (1983), pp. 108—112, which describes the algorithm for scaling examination marks that led to the matrix problem discussed in detail in Section 3.

BIBLIOGRAPHY

Alalouf, I. S., and Styan, G. P. H. (1979a). Characterizations of estimability in the general linear model. *Ann. Statist.,* 7, 194—200.

Alalouf, I. S., and Styan, G. P. H. (1979b). Estimability and testability in restricted linear models. *Math. Operationsforsch. Statist., Ser. Statistics,* 10, 189—201.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* Second Edition. Wiley, New York.

Anderson, T. W., and Styan, George P. H. (1982). Cochran's theorem, rank additivity and tripotent matrices. *Statistics and Probability: Essays in Honor of C. R. Rao* (G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh, *eds.*). North-Holland, Amsterdam, pp. 1—23.

Anderson, W., Jr., and Trapp, G. (1976). Problem 76—8: A matrix inequality. *SIAM Review,* 18, p. 295.

Banachiewicz, Th. (1937a). Sur l'inverse d'un cracovien et une solution générale d'un système d'équations linéaires. *Comptes Rendus Mensuels des Séances de la Classe des Sciences Mathématiques et Naturelles de l'Académie Polonaise des Sciences et des Lettres,* no. 4 (séance du lundi, 5 avril 1937), pp. 3—4.

with $r = 3$ and $c = 4$. The associated layout is connected. Then $n_{..} = 8 > 7 = r + c$ so that **B** might be positive definite, cf. (3.34). But we can form the rank-one matrix

$$\mathbf{Z}^* = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (1, 2, 3, 4) \tag{3.57}$$

from **Z** by replacing its empty cells with appropriate nonzero scalars, and so by Theorem 3.1, Broyden's matrix **B** has nullity one. In fact

$$\mathbf{B} = \frac{1}{3} \begin{pmatrix} 24 & -18 & -18 & -36 \\ -18 & 80 & -6 & -80 \\ -18 & -6 & 72 & -12 \\ -36 & -80 & -12 & 320 \end{pmatrix}, \tag{3.58}$$

while the other Schur complement

$$\mathbf{S} = \frac{1}{65} \begin{pmatrix} 169 & -26 & -39 \\ -26 & 58 & -30 \\ -39 & -30 & 33 \end{pmatrix}. \tag{3.59}$$

Clearly

$$\mathbf{B} \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \\ 1/4 \end{pmatrix} = \mathbf{0} \quad \text{and} \quad \mathbf{S} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \mathbf{0}, \tag{3.60}$$

and so **B** and **S** have nullity at least equal to 1. To confirm that the nullities are both actually equal to 1 we compute the determinant of the leading 3×3 submatrix of **B** to be equal to $3120 \neq 0$, and the determinant of the leading 2×2 submatrix of **S** to be equal to $9126/4225 \neq 0$.

The equations in (3.60) generalize to:

$$\mathbf{B}\mathbf{D}_b^{-1}\mathbf{e} = \mathbf{0} \quad \text{and} \quad \mathbf{S}\mathbf{D}_a\mathbf{e} = \mathbf{S}\mathbf{a} = \mathbf{0}, \tag{3.61}$$

i.e., the vector with elements equal to $1/b_j$ spans the null space of **B**, while the vector with elements a_i spans the null space of **S**.

3.5.4 Necessary and sufficient conditions when the layout is not connected

When the layout associated with the matrix **Z** is not connected, then the rows and columns of **Z** may be arranged so that

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Z}_1 & & & \\ & \mathbf{Z}_2 & & \\ & & \cdots & \\ & & & \mathbf{Z}_l \end{pmatrix}, \tag{3.62}$$

with $l \geq 2$, where the layout corresponding to each submatrix $\mathbf{Z}_h$ ($h = 1, \ldots, l$) is connected (cf. e.g., Searle, 1971, page 320).

Broyden's matrix **B** will then be positive definite if and only if none of the submatrices $\mathbf{Z}_h$ can be expressed in the form (3.55), and positive semidefinite with nullity u ($1 \leq u \leq l$) if and only if precisely u of the submatrices $\mathbf{Z}_h$ can be expressed in the form (3.55).

4. ACKNOWLEDGEMENTS

I am extremely grateful to Tarmo Pukkila and to Simo Puntanen for their most generous invitation to speak at this First International Tampere Seminar on Linear Statistical Models and their Applications.

I am also very grateful to T. W. Anderson, Dominique Latour, Timo Mäkeläinen, Diane V. Ouellette, and Friedrich Pukelsheim, for most helpful discussions concerning various aspects of the research reported in this paper. The final version of this paper was completed while I was on sabbatical leave in the Department of Mathematics and Statistics at The University of Auckland: I am particularly appreciative to Alastair J. Scott for his remarkably generous provision of facilities.

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and by the Gouvernement du Québec, Programme de formation de chercheurs et d'action concertée.

NOTE ADDED IN PROOF

I am very grateful to Professor C.G. Broyden for drawing my attention to his paper "A mark-scaling algorithm", *The Computer Journal*, vol. 26 (1983), pp. 108—112, which describes the algorithm for scaling examination marks that led to the matrix problem discussed in detail in Section 3.

BIBLIOGRAPHY

Alalouf, I. S., and Styan, G. P. H. (1979a). Characterizations of estimability in the general linear model. *Ann. Statist.*, 7, 194—200.

Alalouf, I. S., and Styan, G. P. H. (1979b). Estimability and testability in restricted linear models. *Math. Operationsforsch. Statist., Ser. Statistics,* 10, 189—201.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* Second Edition. Wiley, New York.

Anderson, T. W., and Styan, George P. H. (1982). Cochran's theorem, rank additivity and tripotent matrices. *Statistics and Probability: Essays in Honor of C. R. Rao* (G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh, *eds.*). North-Holland, Amsterdam, pp. 1—23.

Anderson, W., Jr., and Trapp, G. (1976). Problem 76—8: A matrix inequality. *SIAM Review,* 18, p. 295.

Banachiewicz, Th. (1937a). Sur l'inverse d'un cracovien et une solution générale d'un système d'équations linéaires. *Comptes Rendus Mensuels des Séances de la Classe des Sciences Mathématiques et Naturelles de l'Académie Polonaise des Sciences et des Lettres,* no. 4 (séance du lundi, 5 avril 1937), pp. 3—4.

Banachiewicz, T. (1937b). Zur Berechnung der Determinanten, wie auch der Inversen, und zur darauf basierten Auflösung der Systeme linearer Gleichungen. *Acta Astronomica, Série C,* 3, 41—67.

Bjerhammar, A. (1958). A generalized matrix algebra. *Kungl. Tekniska Högskolans Handlingar (Stockholm),* 124, 1—32.

Bodewig, E. (1959). *Matrix Calculus.* Second revised and enlarged edition. North-Holland, Amsterdam.

Boerner, H. (1975). Issai Schur. *Dictionary of Scientific Biography* (C. C. Gillispie, *ed.*). Scribner's, New York, vol. 12, p. 237.

Broyden, C. G. (1982). Problem 82—6: A matrix problem. *SIAM Review,* 24, p. 223.

Broyden, C. G. (1983). Solution to Problem 82—6. *SIAM Review,* 25, p. 405.

Brualdi, Richard A., and Schneider, Hans (1983). Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, and Cayley. *Linear Algebra Appl.,* 52/53, 769—791.

Carlson, David (1984). What are Schur complements, anyway? *Linear Algebra Appl.,* 59, 188—193.

Cottle, R. W. (1974). Manifestations of the Schur complement. *Linear Algebra Appl.,* 8, 189—211.

Dahan, Simon, and Styan, George P. H. (1977). Estimability of a proper subset of the regression coefficients in the general linear model with less than full rank. Technical Report, Department of Mathematics, McGill University. [Abstract: *Inst. Math. Statist. Bull.,* 6, p. 48.]

Duncan, W. J. (1944). Some devices for the solution of large sets of simultaneous linear equations (with an appendix on the reciprocation of partitioned matrices). *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Seventh Series,* 35, 660—670.

Frobenius. F. G. (1908). Über Matrizen aus positiven Elementen. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin,* 471—476. [Reprinted in *Ferdinand Georg Frobenius Gesammelte Abhandlungen* (J. P. Serre, *ed.*), Springer-Verlag, Berlin, vol. 3(1968), pp. 404—409.]

Gaffke, N., and Krafft, O. (1982). Matrix inequalities in the Löwner ordering. *Modern Applied Mathematics: Optimization and Operations Research* (Bernhard Korte, *ed.*). North-Holland, Amsterdam, pp. 595—622.

Groves, Theodore, and Rothenberg, Thomas (1969). A note on the expected value of an inverse matrix. *Biometrika,* 56, 690—691.

Guttman, Louis (1946). Enlargement methods for computing the inverse matrix. *Ann. Math. Statist.,* 17, 336—343.

Hartwig, Robert E. (1978). A note on the partial ordering of positive semi-definite matrices. *Linear and Multilinear Algebra,* 6, 223—226.

Haynsworth, E. V. (1968). Determination of the inertia of a partitioned Hermitian matrix. *Linear Algebra Appl.,* 1, 73—81.

Hemmerle, William J. (1979). Balanced hypotheses and unbalanced data. *J. Amer. Statist. Assoc.,* 74, 794—798.

Henderson, H. V., and Searle, S. R. (1981). On deriving the inverse of a sum of matrices. *SIAM Review,* 23, 53—60.

James, A. T., and Wilkinson, G. N. (1971). Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance. *Biometrika,* 58, 279—294.

Khatri, C. G. (1976). A note on multiple and canonical correlation for a singular covariance matrix. *Psychometrika,* 41, 465—470.

Latour, Dominique, and Styan, George P.H. (1985). Canonical correlations in the two-way layout. *Proc. First International Tampere Seminar on Linear Statistical Models and their Applications,* 225—243.

Lieb, E. H. (1977). Solution to Problem 76—8. *SIAM Review,* 19, p. 330.

Lieb, E. H., and Ruskai, M. B. (1974). Some operator inequalities of the Schwarz type. *Advances in Math.,* 12, 269—273.

Marsaglia, George, and Styan, George P. H. (1974a). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra,* 2, 269—292.

Marsaglia, George, and Styan, George P. H. (1974b). Rank conditions for generalized inverses of partitioned matrices. *Sankhyā Ser. A,* 36, 437—442.

Marshall, Albert W., and Olkin, Ingram (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

Milliken, G. A., and Akdeniz, F. (1977). A theorem on the difference of the generalized inverses of two nonnegative matrices. *Comm. Statist. A — Theory Methods,* 6, 73—79.

Mirsky, L. (1955). *An Introduction to Linear Algebra.* Oxford Univ. Press. [Reprinted (1982), Dover, New York.]

Moore, M. H. (1973). A convex matrix function. *Amer. Math. Monthly,* 80, 408—409.

Moore, M. H. (1977). Solution to Problem 76—8. *SIAM Review,* 19, 329—330.

Ouellette, Diane Valérie (1981). Schur complements and statistics. *Linear Algebra Appl.,* 36, 187—295.

Pukelsheim, Friedrich, and Styan, George P. H. (1983). Convexity and monotonicity properties of dispersion matrices of estimators in linear models. *Scand. J. Statist.,* 10, 145—149.

Raghavarao, Damaraju (1971). *Constructions and Combinatorial Problems in Design of Experiments.* Wiley, New York.

Rao, C. Radhakrishna (1981). A lemma on g-inverse of a matrix and computation of correlation coefficients in the singular case. *Comm. Statist. A — Theory Methods,* 10, 1—10.

Rao, C. Radhakrishna (1985). A unified approach to inference from linear models. *Proc. First International Tampere Seminar on Linear Statistical Models and their Applications,* 9—36.

Roy, S. N., and Roy, J. (1959). A note on a class of problems in 'normal' multivariate analysis of variance. *Ann. Math. Statist.,* 30, 577—581.

Scheffé, Henry (1959). *The Analysis of Variance.* Wiley, New York.

Schur, J. [Issai] (1917). Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, I. *J. Reine Angew. Math.,* 147, 205—232. [Reprinted in *Issai Schur Gesammelte Abhandlungen* (Alfred Brauer & Hans Rohrbach, *eds.*), Springer-Verlag, Berlin, vol. 2 (1973), pp. 137—164.]

Searle, S. R. (1971). *Linear Models.* Wiley, New York.

Seshadri, V., and Styan, G. P. H. (1980). Canonical correlations, rank additivity and characterizations of multivariate normality. *Colloquia Math. Soc. János Bolyai, vol. 21: Analytic Function Methods in Probability Theory (Debrecen, Hungary, August 1977).* János Bolyai, Budapest, and North-Holland, Amsterdam, pp. 331—344.

Sherman, J., and Morrison, W. J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *Ann. Math. Statist.,* 21, 124—127. [Abstract: vol. 20 (1949), p. 621.]

Styan, George P. H. (1983). Generalized inverses. *Encyclopedia of Statistical Sciences, vol. 3: Faà di Bruno's Formula to Hypothesis Testing* (Samuel Kotz, Norman L. Johnson, and Campbell B. Read, *eds.*). Wiley, New York, pp. 334—337.

Styan, George P. H., and Pukelsheim, Friedrich (1978). On the inertia of matrix differences and Schur complements. Technical Report, Department of Mathematics, McGill University. [Abstract: *Inst. Math. Statist. Bull.,* 7, p. 364.]

Sylvester, J. J. On the relation between the minor determinants of linearly equivalent quadratic functions. *Philosophical Magazine,* 1, 295—305. [Reprinted in *The Collected Mathematical Papers of James Joseph Sylvester* (H. F. Baker, *ed.*), Cambridge Univ. Press, vol. 1 (1904), pp. 241—250, plus »Editor's Note», pp. 647—650.]

Sylvester, J. J. (1852). A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitutions to the form of a sum of positive and negative squares. *Philosophical Magazine,* 4, 138—142. [Reprinted in *The Collected Mathematical Papers of James Joseph Sylvester* (H. F. Baker, *ed.*), Cambridge Univ. Press, vol. 1 (1904), pp. 378—381.]

Turnbull, H. W., and Aitken, A. C. (1932). *An Introduction to the Theory of Canonical Matrices.* Blackie, London. [Reprinted (1961), Dover, New York.]

December 1984

Department of Mathematics and Statistics
McGill University
805 ouest, rue Sherbrooke
Montréal, Québec, Canada H3A 2K6

and

Department of Mathematics and Statistics
The University of Auckland
Private Bag
Auckland, New Zealand

Pukkila
Puntanen
Editors

Proceedings of the First International Tampere Seminar on
Linear Statistical Models and their Applications

PROCEEDINGS

of the

First International Tampere Seminar on Linear Statistical Models and their Applications

EDITED BY Tarmo Pukkila and Simo Puntanen
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF TAMPERE, FINLAND

PROCEEDINGS

of the

First International Tampere Seminar on Linear Statistical Models and their Applications

University of Tampere, Tampere, Finland
August 30th to September 2nd, 1983

Edited by

Tarmo Pukkila and Simo Puntanen

Published by

Department of Mathematical Sciences/Statistics, University of Tampere
P.O. BOX 607, SF-33101 Tampere, Finland

1985

Report No. A 138

ISBN 951-44-1693-7
ISSN 0356-4231

Printed by the University of Tampere Printing Service
Composition by Vammalan Kirjapaino Oy, Vammala

Preface

The First International Tampere Seminar on Linear Statistical Models and their Applications was held at the University of Tampere, Tampere, Finland, during the period August 30—September 2, 1983. The seminar brought together, from at least nine different countries, more than 100 researchers in linear statistical models and related areas.

The main speakers in the seminar were Professor C. Radhakrishna Rao (University of Pittsburgh, U.S.A., and the Indian Statistical Institute, New Delhi, India) and George P. H. Styan (McGill University, Montréal, Canada). There is no doubt that the whole seminar audience greatly enjoyed the excellent lecture series given by these two outstanding statisticians. Professor Rao's topic was »A Unified Approach to Inference from Linear Models» and Professor Styan's was »Schur Complements and Linear Statistical Models». The editors of these *Proceedings* are extremely grateful to Professors Rao and Styan for the effort they put into their fine articles for this volume; we are very proud to publish them.

This was the first time Professor Rao had visited Finland and certainly his visit was a great honour for the whole statistical community of Finland. It is also a great honour to the University of Tampere that Professor Rao, one of the fathers of modern statistics, has consented to accept an Honorary Doctorate in connection with the University's 60th anniversary in May 1985.

Both Professor Rao's and Professor Styan's stimulating personalities and their vital participation in all seminar activities made a great impression on the organizers. The role of Professor Styan in planning the seminar was indeed invaluable. Also our sincere thanks go to him for his very useful advice and cooperation in preparing the *Proceedings* for publication.

The invited talks at the seminar were given by Professors R. W. Farebrother (University of Manchester, U.K.), Johan Fellman (Swedish School of Economics, Helsinki), Hannu Niemi (University of Helsinki) and Bimal Kumar Sinha (University of Pittsburgh, U.S.A.). My special thanks are due to these speakers, as well as to the speakers in the contributed paper sessions, for an excellent series of talks and for their kind cooperation in preparing the papers for publication. I also wish to thank Professors Jerzy K. Baksalary (Academy of Agriculture, Poznań, Poland), J. A. Melamed (Tbilisi, U.S.S.R.) and Der-Shin Chang (Hsinchu, Taiwan), who, unfortunately, for unavoidable reasons, were not able to participate in the seminar. The last two, however, do have their contributions in the *Proceedings.*

All papers in this volume have been refereed and I therefore wish to thank the anonymous referees for their efforts. The deadline for the first versions of the papers was September 2, 1983. On the front page of each paper is the affiliation of the author as of August 1983 and at the end of the paper the current affiliation and the date when the final version was received.

Financially the seminar was supported by the University of Tampere, the Academy of Finland, the City of Tampere, and various associations and companies whose names are listed on page viii. All deserve our sincerest thanks for their invaluable support. The City of Tampere also very kindly invited all participants to a Civic Reception in the Town Hall.

The seminar was organized by a local committee within the Department of Mathematical Sciences/Statistics, consisting of Paula Hietala, Pentti Huuhtanen, Päivi Laurila, Erkki Liski, Simo Puntanen and myself. Also Raija Leppälä, Olavi Stenman and Pirkko Welin very kindly gave their help in various arrangements. I am deeply grateful to each and every one of my colleagues for this tremendous cooperation.

Jointly with Simo Puntanen, as the editors of these *Proceedings,* I wish to give special thanks to Pirkko Welin for her kind assistance in preparing this volume.

I would also like to thank Professor and Mrs. Eino Haikala for their warm hospitality in connection with the seminar. The organizers of the seminar are also grateful to the University Rector J. K. Visakorpi for his address during the inauguration of the seminar.

Our sincere thanks also go to all of the participants in the seminar, whether reading papers or not, for it was their participation which made the seminar a success. The organizers were also glad to see the active interest shown in the social programme of the seminar. It is my impression that everyone will have warm memories of the Seminar Dinner and Civic Reception — even hot ones of the Sauna Party.

The success of the seminar is due also to Professors Gustav Elfving (University of Helsinki) and Timo Mäkeläinen (University of Helsinki). Both acted as chairmen during several sessions in the seminar. Besides this, Professor Elfving gave an inspiring opening address entitled »Finnish Mathematical Statistics in the Past». I am deeply grateful for their help and for their cooperation in carrying out the seminar.

While preparing this volume for publication, we were deeply saddened that Professor Gustav Elfving passed away on March 25, 1984. He was instrumental in the promotion of statistics in Finland as well as of our seminar. As the seminar organizers we were very impressed not only by his active participation in mathematical and statistical discussions but also by his great and warm sociability: he was one of the real old-time gentlemen. It is to Professor Elfving that we dedicate this volume.

Tampere, February 1985

Tarmo Pukkila
Seminar Director

THE FINANCIAL SUPPORT
OF THE FOLLOWING INSTITUTIONS
IS GRATEFULLY ACKNOWLEDGED:

THE ACADEMY OF FINLAND
THE TAMPERE UNIVERSITY FOUNDATION
THE TAMPERE CITY SCIENCE FOUNDATION

KANSALLIS-OSAKE-PANKKI

FINNAIR OY

OY INTERNATIONAL BUSINESS MACHINES AB
THE FINNISH STATISTICAL SOCIETY
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY

Contents

PREFACE v

Timo Mäkeläinen: In Memoriam: Gustav Elfving 1

Opening Address

Gustav Elfving: Finnish mathematical statistics in the past 3

Invited Lecture Series

C. Radhakrishna Rao: A unified approach to inference from linear models. 9

George P.H. Styan: Schur complements and linear statistical models. 37

Invited Talks

R.W. Farebrother: The statistical estimation of the standard linear model, 1756–1853 77

Johan Fellman: Topics on singular linear models with special reference to experimental design 101

Hannu Niemi: On time dependent covariance and parameter structures. 123

B.K. Sinha and H. Drygas: Robustness in linear models 143

Contributed Papers

H. Bunke and U. Schulze: Approximation of change points in regression models. . 161

Kamal C. Chanda: Asymptotic expansions for densities of sample means for linear processes 179

D.S. Chang: On the asymptotic distribution of the multivariate regression estimator 189

L.B. Klebanov, G.M. Manija and J.A. Melamed: On the estimation of parameters of type of stable laws 203

Dominique Latour and George P.H. Styan: Canonical correlations in the two-way layout 225

Erkki P. Liski: Choosing a shrinkage estimator – a test of the mean square error criterion 245

Jorma Kaarlo Merikoski: On a matrix inequality in genetics 263

Simo Puntanen: Properties of the canonical correlations between the least squares fitted values and the residuals 269

Burkhard Schaffrin: A note on linear prediction within a Gauss-Markoff model linearized with respect to a random approximation 285

P. Stahlecker and G. Trenkler: On heterogeneous versions of the best linear and the ridge estimator 301

Yves Thiebaudeau and George P.H. Styan: Bounds for Chakrabarti's measure of imbalance in experimental design 323

Dedicated to the memory of

Gustav Elfving

IN MEMORIAM

Gustav Elfving
1908—1984

While these *Proceedings* were in print we learned that the prominent contributor to the Seminar, Professor Gustav Elfving had died on the 25th of March 1984 at his home in Helsinki.

Erik Gustav Elfving was born in 1908 in Helsinki. He graduated in 1930 in Mathematics and took his doctoral degree in 1934, both at the University of Helsinki. He was appointed Professor of Mathematics at the same university in 1948 and retired in 1975.

Elfving wrote his doctoral thesis under the guidance of Rolf Nevanlinna on the latter's value distribution theory. Elfving's first paper in probability in 1937 marks the beginning of a long and distinguished career in probability and statistics. His writings published for the international professional audience contain papers on Markov chains, point processes, sufficiency and complete classes, design of linear experiments and sample designs, test item selection, order statistics, exact distributions and expansions of distributions, quality control, nonparametric statistics, and Bayesian statistics. Characteristic to them are many original ideas, clear formulation, and elegant presentation stressing central ideas.

Perhaps the single most influential piece of Elfving's writings — and one in a subject very close to this Seminar — is the paper »Optimum allocation in linear regression theory» published in 1952 in the *Annals of Mathematical Statistics*. The problem of optimal design of linear experiments is here brought, for the first time in some generality, before a wide statistical audience. Today an extensive field of study, optimal design of experiments has since undergone profound technical development. Yet, Elfving's work remains part of its foundations. (Compare with Professor Fellman's lecture in this volume.)

During his retirement Elving took the role of a historian of mathematics. His lecture in this volume partly draws on a major work »*The History of Mathematics in Finland 1828—1918*» (Societas Scientiarum Fennica, Helsinki 1981, 195 pp.).

September 1984

Timo Mäkeläinen
Department of Mathematics
University of Helsinki
Helsinki, Finland