COMMUN. STATIST.-THEOR. METH., A10(1), 1-10 (1981)

A LEMMA ON G-INVERSE OF A MATRIX AND COMPUTATION OF CORRELATION COEFFICIENTS IN THE SINGULAR CASE

C. Radhakrishna Rao
University of Pittsburgh
Pittsburgh, Pa. 15260

Key Words and Phrases: g-inverse, singular covariance matrix, multiple, partial, canonical correlations.

ABSTRACT

Formulae for multiple, partial and canonical correlation coefficients are generally expressed in terms of the elements of inverse covariance matrix. They are not applicable when the covariance matrix is singular. In this paper, a unified approach is presented to cover the singular and non-singular cases. The formulae involve g-inverse of singular matrices and the results are derived from a lemma on the structure of the idempotent matrix AA^- where A^- is any g-inverse (i.e., $AA^-A=A$). The conditions under which some columns of AA^- are unit vectors are obtained. The formulae for canonical correlations and the canonical transformations of two sets of variables in the singular case are shown to be of the same form as in the non-singular case, with the convention that only the proper eigen values and vectors of determinantal equations are considered.

1. A BASIC LEMMA ON G-INVERSE

Let A be $m \times n$ matrix and A^- (of order $n \times m$) be a g-inverse of A, i.e., $AA^-A=A$ (Rao, 1962). If A is a square matrix of order m admitting a regular inverse A^{-1}, then $AA^{-1}=I_m$. In general $Q = AA^- \neq I_m$. The following basic lemma provides the conditions under which some columns of Q are unit vectors.

We use the following notations. A' is the transpose of A, A^* is the conjugate transpose of A, $\rho(A)$ is the rank of A, $S(A)$ is the linear space generated by the columns of A and e_i is the i-th unit column vector, i.e., with unity in the i-th position and zeroes elsewhere. A vector a_1 is said to be independent of $a_2, a_3, \ldots$ if there do not exist scalars $b_2, b_3, \ldots$ such that $a_1 = b_2a_2+b_3a_3+\ldots$. By definition the null vector is always a dependent vector.

Lemma. Let $A' = (A_1' : A_2')$ be a partition of A and $A^- = (B_1 : B_2)$ be any g-inverse of A. Then

$$S(A_1') \cap S(A_2') = 0 \iff \tag{1.1}$$

$$A_1B_1A_1 = A_1,\ A_2B_2A_2 = A_2,\ A_1B_2A_2 = 0,\ A_2B_1A_1 = 0. \tag{1.2}$$

Proof. By definition

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} (B_1 : B_2) \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \iff$$

$$A_1B_1A_1 + A_1B_2A_2 = A_1,\ A_2B_1A_1 + A_2B_2A_2 = A_2$$

$$A_1'(I - B_1'A_1') = A_2'B_2'A_1',\quad A_1'B_1'A_2' = A_2'(I - B_2'A_2'). \tag{1.3}$$

If $S(A_1') \cap S(A_2') = 0$, then all the expressions in (1.3) must be zero, which proves the if part.

Conversely let (1.2) hold and suppose that there exist vectors a_1 and a_2 such that $A_1'a_1 = A_2'a_2$. Then using the conditions (1.2)

$$A_1'a_1 = A_1'B_1'A_1'a_1 = A_1'B_1'A_2'a_2 = 0$$

which implies (1.1).

Corollary 1. Let A_1 be $p \times n$ matrix. Then

$$S(A_1') \cap S(A_2') = 0,\ \ \rho(A_1) = p \tag{1.4}$$

$$\iff \quad A_1B_1 = I_p,\ A_2B_1 = 0 \tag{1.5}$$

i.e., the first p columns of $Q = AA^-$ are unit vectors $e_1, \ldots, e_p$ for any g-inverse A^-.

Proof. If (1.4) holds, then (1.5) follows by post multiplying (1.2) by A_1^* and observing that $A_1 A_1^*$ is non-singular. It is easy to see that if (1.5) holds then $\rho(A_1) = p$ and the vector spaces $S(A_1')$ and $S(A_2')$ have no non-null intersection.

Note 1. Corollary 1 shows that a necessary and sufficient condition for Q_i, the i^{th} column of Q, to be e_i is that the i^{th} row vector of A is non-null and independent of the other row vectors of A.

Corollary 2. Let each of p rows of A be independent of all the other rows and each of q columns of A be independent of all the other columns. We may assume without loss of generality that these are the first p rows and q columns. Further let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad A^- = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{1.6}$$

where A^- is any g-inverse of A, A_{11} is $p \times q$ and B_{11} is $q \times p$ matrices. Then B_{11} is unique for all g-inverses.

Proof. By hypothesis $\rho(A_{11}\ A_{12}) = p$ and from Corollary 1,

$$A_{11}B_{11} + A_{12}B_{21} = I \tag{1.7}$$

$$A_{21}B_{11} + A_{22}B_{21} = 0. \tag{1.8}$$

If there is an alternative differing form B_{ij} by C_{ij}, then

$$\begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix} C_{11} + \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} C_{21} = 0$$

which implies $C_{11} = 0$ using the conditions on the columns.

Note 2. Corollary 2 says that if the i-th row of A is independent of the others and the j-th column is independent of the others, then the (j, i)-th element of A^- is unique.

Corollary 3. Let A be an n.n.d. (non-negative definite) matrix and consider the partitions (1.6) with p = q. Then

$$\rho(A_1) = p \text{ and } S(A_1') \cap S(A_2') = 0 \tag{1.9}$$

$$\Longleftrightarrow \rho(A_{11} - A_{12}A_{22}^{-}A_{21}) = p \tag{1.10}$$

in which case $B_{11} = (A_{11} - A_{12}A_{22}^{-}A_{21})^{-1}$.

Proof. If (1.9) holds, Corollary 2 applies giving (1.7) and (1.8). Multiplying (1.8) by $A_{12}A_{22}^{-}$ and subtracting from (1.7) we have $(A_{11} - A_{12}A_{22}^{-}A_{21})B_{11} = I$ noting that $A_{12} = A_{12}A_{22}^{-}A_{22}$ since $S(A_{21}) \subset S(A_{22})$ for an n.n.d. matrix. The converse is easy to establish noting that (1.10) implies $\rho(A_{11}) = p = \rho(A_1)$ and if there exist vectors a_1 and a_2 such that $a_1'A_1 = a_2'A_2$, then $a_1 = 0$.

2. CORRELATION COEFFICIENTS

Let the joint covariance matrix of two sets of variables $x_1,\ldots,x_p$ and $x_{p+1},\ldots,x_{p+q}$ be

$$\Sigma = (\sigma_{ij}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}. \tag{2.1}$$

If $\rho(\Sigma) \neq p + q$, then some rows are dependent on the others. If the i-th row is dependent on the others, then x_i is a linear function (may be non-homogeneous) of the other variables with probability 1.

It is shown in Rao (1973, p. 522) that the covariance matrix of $x_{i.(p+1)\ldots(p+q)}, i = 1,\ldots,p$, i.e., of $x_1,\ldots,x_p$ eliminating the regression due to $x_{p+1},\ldots,x_{p+q}$ is

$$\Sigma_{11} - \Sigma_{21}\Sigma_{22}^{-}\Sigma_{12} \tag{2.2}$$

for any g-inverse of Σ_{22}. Let $\Sigma^{-} = (\sigma^{ij})$ be any g-inverse of Σ and define $\Sigma\,\Sigma^{-} = (Q_1:\ldots:Q_{p+q})$. We shall express the multiple and partial correlation coefficients using only the elements of Σ^{-} and the nature of Q_i (whether it is e_i or not).

2.1 SQUARED MULTIPLE CORRELATION

Theorem 2.1. The squared multiple correlation of x_1 on the rest is

$$R^2_{1.23\ldots} = 1 \quad \text{if } Q_1 \neq e_1 \text{ and } \sigma_{11} \neq 0 \tag{2.3}$$

$$= 1 - \frac{1}{\sigma_{11}\sigma^{11}} \quad \text{if } Q_1 = e_1. \tag{2.4}$$

Proof. Theorem 3.1 is proved by Tucker et al (1973), and in a more general form by Khatri (1976), using the correlation matrix. We shall prove the results using the basic lemma and its corollaries.

Consider (2.2) with $p = 1$. If $Q_1 = e_1$, then $\sigma_{11} \neq 0$ and by Corollary 3, $\sigma^{11} = (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21})^{-1}$. By definition

$$R^2_{1.23\ldots} = 1 - \frac{\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21}}{\sigma_{11}}$$

$$= 1 - \frac{1}{\sigma_{11}\sigma^{11}}$$

which proves (2.4). If $Q_1 \neq e_1$, $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21}$ must zero by Corollary 3. If $\sigma_{11} \neq 0$, the formula (2.5) shows that $R^2_{1.23\ldots} = 1$, which proves (2.3). (If $\sigma_{11} = 0$, then x_1 is a constant with probability 1).

2.2 PARTIAL CORRELATIONS

Consider the partition of the covariance matrix (2.1) with $p = 2$ and let the corresponding matrix (2.2) be

$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{21} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$$

Then the covariance between $x_{1.3\ldots}$ and $x_{2.3\ldots}$ is γ_{12}, and the partial correlation is $\gamma_{12}/\sqrt{\gamma_{11}\gamma_{22}}$ provided $\gamma_{11} \neq 0$, $\gamma_2 \neq 0$. We may define the partial correlation to zero if $\gamma_{12} = 0$ and at least one of γ_{11} and γ_{22} is not zero, and as unity if $\gamma_{11} = \gamma_{22} = \gamma_{12} = 0$. With this convention, we have the following theorem, which expresses the results in terms of the elements of Σ^{-} only.

Theorem 2.2 The partial correlation between x_1 and x_2 eliminating $x_3, x_4, \ldots$ is

$$r_{12.34\ldots} = \frac{-\sigma^{12}}{\sqrt{\sigma^{11}\sigma^{22}}} \quad \text{if } Q_1 = e_1 \text{ and } Q_2 = e_2. \tag{2.5}$$

$$= 0 \text{ if } Q_1 \neq e_1 \text{ and } Q_2 = e_2, \text{ or if } Q_1 = e_1 \text{ and } Q_2 \neq e_2. \tag{2.6}$$

$$= 1 \text{ if } Q_1 \neq e_1 \text{ and } Q_2 \neq e_2. \tag{2.7}$$

Proof. If $Q_1 = e_1$ and $Q_2 = e_2$, then by Corollary 3

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{pmatrix}$$

which proves (2.5).

If $Q_1 = e_1$ and $Q_2 \neq e_2$ then $\gamma_{11} \neq 0$, but $\gamma_{22} = \gamma_{12} = 0$, i.e., the partial covariance and one of the partial variances is zero, which by our convention proves (2.6). If $Q \neq e_1$ and $Q_2 = e_2$, then $\gamma_{22} \neq 0$, but $\gamma_{12} = \gamma_{11} = 0$.

If $Q_1 \neq e_1$ and $Q_2 \neq e_2$, then the residuals $x_{1.3\ldots}$, $x_{2.3\ldots}$ are either degenerate random variables (in which case $\gamma_{11} = \gamma_{22} = \gamma_{12} = 0$) or non-degenerate linearly related variables (in which case $\gamma_{12}^2 = \gamma_{11}\gamma_{22}$, $\gamma_{11} \neq 0$, $\gamma_{22} \neq 0$). By convention in the former case and by the usual definition in the latter, the partial correlation is unity, which proves (2.7).

2.3 CANONICAL CORRELATIONS

Consider the partition of the covariance matrix (2.1) corresponding to two sets of p and q variables. If Σ_{11} and Σ_{22} are non-singular, Khsirsager (1972) has shown that transformations of the p and q variables into canonical variates can be obtained by considering the singular value decomposition of

$$\Sigma_{11}^{-\frac{1}{2}} \Sigma_{12} \Sigma_{22}^{-\frac{1}{2}} = P \Delta Q' \tag{2.8}$$

where $\Sigma_{11}^{-\frac{1}{2}}$ and $\Sigma_{22}^{-\frac{1}{2}}$ are Gramian square roots of Σ_{11} and Σ_{22}. We show that the same method works when Σ_{11} and Σ_{22} are possibly singular by suitably defining the pre and post multiplying matrices of Σ_{12}.

Consider the spectral decompositions

$$\Sigma_{11} = (P_1 : P_2) \begin{pmatrix} \Delta_1^2 & 0 \\ 0 & 0 \end{pmatrix} (P_1 : P_2)'$$

$$\Sigma_{22} = (Q_1 : Q_2) \begin{pmatrix} \Delta_2^2 & 0 \\ 0 & 0 \end{pmatrix} (Q_1 : Q_2)'.$$

Note that P_1 is of order $p \times r$ where $r = \rho(\Sigma_{11})$ and Q_1 is of order $q \times t$ where $t = \rho(\Sigma_{22})$. Now, consider the singular value decomposition of

$$\Delta_1^{-1} P_1' \Sigma_{12} Q_1 \Delta_2^{-1} = (W_1 : W_2) \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} (F_1 : F_2)' \tag{2.9}$$

where Λ is a non-singular diagonal matrix of rank $s = \rho(\Sigma_{12})$, with diagonal elements $\lambda_1,\ldots,\lambda_s$. The transformations on $Y_1' = (x_1,\ldots,x_p)'$ and $Y_2' = (x_{p+1},\ldots,x_{p+q})'$ can be broken up into three parts

$$S_1 = W_1' \Delta_1^{-1} P_1' Y_1 \ , \ T_1 = F_1' \Delta_2^{-1} Q_1' Y_2, \tag{2.10}$$

$$S_2 = W_2' \Delta_1^{-1} P_1' Y_1 \ , \ T_2 = F_2' \Delta_2^{-1} Q_1' Y_2, \tag{2.11}$$

$$S_3 = P_2' Y_1 \qquad , \ T_3 = Q_2' Y_2 \ . \tag{2.12}$$

The joint covariance matrix of S_1, S_2, S_3, T_1, T_2, T_3 is

$$\begin{pmatrix} I_s & 0 & 0 & \Lambda & 0 & 0 \\ 0 & I_{r-s} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \Lambda & 0 & 0 & I_s & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{t-s} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.13}$$

<u>Note 1</u>. From (2.9) it is seen that λ_i^2 are the non-zero roots of

$$|P_1'\Sigma_{12}\, \Sigma_{22}^{-} \Sigma_{21}\, P_1 - \lambda^2 \Delta_1^2| = 0 \tag{2.14}$$

which are the same as the non-zero roots of

$$|\Sigma_{12}\ \Sigma_{22}^{-}\ \Sigma_{21}\ \Sigma_{11}^{-} - \lambda^2\ I| = 0 \tag{2.15}$$

where Σ_{11}^{-} and Σ_{22}^{-} are any g-inverses of Σ_{11} and Σ_{22} respectively. As shown in Rao and Mitra (1971, p. 125), the non-zero roots of (2.15) are the proper non-zero roots of

$$|\Sigma_{12}\ \Sigma_{22}^{-}\ \Sigma_{21} - \lambda^2\ \Sigma_{11}| = 0\ . \tag{2.16}$$

[λ^2 is said to be a proper root if there exists a non-null vector Y such that $\Sigma_{12}\ \Sigma_{22}^{-}\ \Sigma_{21}\ Y = \lambda^2\ \Sigma_{11}\ Y$ and $\Sigma_{11}\ Y \neq 0$. The vector Y is said to be a proper eigen vector.] Further (2.16) is equivalent to

$$\begin{vmatrix} -\lambda\ \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \lambda\ \Sigma_{22} \end{vmatrix} = 0 \tag{2.17}$$

provided we consider only the proper non-zero roots. The equations (2.16) and (2.17) are again of the same form as the determinantal equations for canonical correlations in the singular case (see 8 f.1.2 and 8 f.1.6 on p. 583, Rao, 1973).

Seshadri and Styan (1977) give different versions of (2.16) and (2.17) based on some matrices computed from Σ. But (2.16) and (2.17) use the original Σ.

Note 2. The number of unit canonical correlations is

$$\rho(\Sigma_{11}) - \rho(P_1'\ \Sigma_{12}\ \Sigma_{22}\ \Sigma_{21}\ P_1 - \Delta_1^2) \text{ from (2.14)} \tag{2.18}$$

$$= p - \rho(\Sigma_{12}\ \Sigma_{22}^{-}\ \Sigma_{21}\ \Sigma_{11}^{-} - I) \text{ from (2.15)} \tag{2.19}$$

$$= \rho(\Sigma_{11}) - \rho(\Sigma_{11} - \Sigma_{12}\ \Sigma_{22}^{-}\ \Sigma_{21}) \text{ from (2.16)} \tag{2.20}$$

$$= \rho(\Sigma_{11}) + \rho(\Sigma_{22}) - \rho(\Sigma) \text{ from (2.17)} \tag{2.21}$$

remembering that in (2.16) and (2.17), the multiplicity of a root is counted by the dimension of the proper eigen space.

The result (2.20) was given earlier by Khatri (1976) and the result (2.21) by Seshadri and Styan (1977) using different arguments. The case of singular covariance matrix was also considered by Hoschel (1974) in his general treatment of correlation and linear dependence between random vector variables.

Note 3. The canonical transformations (2.10) - (2.12) can be directly obtained from the roots and associated vectors of the determinantal equation (2.17). The equations for a given λ are

$$-\lambda \Sigma_{11} L + \Sigma_{12} M = 0 \qquad (2.22)$$

$$\Sigma_{21} L - \lambda \Sigma_{22} M = 0$$

The compounding vectors in the transformation (2.12) are the orthonormal solutions of $\Sigma_{11} L = 0$ and $\Sigma_{22} M = 0$; the vectors in (2.11) are the solutions of $\Sigma_{11} L \neq 0$, $\Sigma_{21} L = 0$ and $\Sigma_{22} M \neq 0$, $\Sigma_{12} M = 0$; and the vectors in (2.10) are those obtained by substituting the non-zero proper roots of (2.17) in the equation (2.22). Thus we have a unified approach to the theory of canonical correlations, using the original covariance matrix Σ.

ACKNOWLEDGEMENTS

The work was partially supported by the Air Force Office of Scientific Research, Air Force Systems Command under Contract F49620-79-C-0161.

BIBLIOGRAPHY

Hoschell, Hans-Peter (1974). A general approach to correlation and linear dependence between random vectors with regular or singular covariance matrix. Math. Oper. u. Statist. 6, 487-507.

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

Kshirsagar, A. M. (1972). Multivariate Analysis. New York: Marcel Dekker, Inc.

Rao, C. R. (1962). A note on a generalized inverse of a matrix with applications to problems in statistics. J. Roy. Statist. Soc. Ser. B 24, 152-158.

Rao, C. R. (1973). Linear Statistical Inference and its Applications (Second Edition). New York: Wiley and Sons, Inc.

Rao, C. R. and Mitra, S. K. (1971). Generalized Inverse of Matrices and its Applications. New York: Wiley and Sons, Inc.

Seshadri, V. and Styan, G. P. H. (1977). Canonical correlations, rank additivity and characterization of multivariate normality. Proc. Coll. on Analytic Function Methods in Probability Theory. (Debrecen).

Tucker, Ledyard R., Cooper, Lee G. and Meredith, William (1972). Obtaining squared multiple correlations from a correlation matrix which may be singular. Psychometrika, 1972, 37, 143-148.

Received February, 1980.

Refereed by Patrick L. Odell, University of Texas at Dallas, Richardson, TX.