GENERALIZED CORRELATIONS IN THE SINGULAR CASE

BY

ASHIS SEN GUPTA

TECHNICAL REPORT NO. 46
NOVEMBER 1980

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
GENERALIZED CORRELATIONS IN THE SINGULAR CASE

by

ASHIS SEN GUPTA
Stanford University

TECHNICAL REPORT NO. 46
NOVEMBER 1980

PREPARED UNDER CONTRACT NO0014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government. Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
GENERALIZED CORRELATIONS IN THE SINGULAR CASE

Ashis Sen Gupta

1. Introduction. When the covariance matrix is singular, the usual expressions for the multiple, partial, canonical [see, e.g., Rao (1973)] and some generalized canonical correlations [for a review, see Sen Gupta (1979)] need to be revised. Tucker, et al. (1972), Khatri (1976) and Rao (1979), (1980) have provided formulae for some of these correlation coefficients in the general case by using g-inverses. A review of their results on multiple, partial and canonical correlations is given first. Next, it is shown that there exists a general representation which covers several generalized canonical correlations and as special cases the multiple, partial and canonical correlations, too. Then a general theory is formulated which deals with the singular case for the representation. Previous results on multiple, partial and canonical correlations follow as special cases of this theory. Further, appropriate formulae are also provided through this formulation for various generalized canonical correlations in the singular case. Finally, the numbers of various critical generalized correlations are derived for the general case.

2. Multiple, partial and canonical correlations in the singular case. Let \( R = (R_1: \ldots : R_p) \) be the correlation matrix of \( p \) variables. Further, let \( R^- = (r_{ij}^{-}) = (T_1: \ldots : T_p) \) be any g-inverse of \( R \). Define, \( RR^- = Q = (Q_1: \ldots : Q_p) \). Let
I_p have the unit vector e_i, as its i-th column, i = 1, ..., p.

**Result 1.** The squared multiple correlation of X_1 on X_2, ..., X_p is

\[ R_1^2 \| 2 \ldots p \| = 1 \text{ if } Q_1 \neq e_1 \]

\[ = 1 - (r_{11})^{-1} \text{ if } Q_1 = e_1. \]

**Result 2.** The partial correlation between X_1 and X_2 eliminating X_3, X_4, ..., X_p is

\[ r_{12} \| 34 \ldots p \| = 0 \text{ if } Q_1 \neq e_1 \text{ and } Q_2 = e_2 \text{ or if } Q_1 = e_1 \text{ and } Q_2 \neq e_2 \]

\[ = 1 \text{ if } Q_1 \neq e_1 \text{ and } Q_2 \neq e_2 \]

\[ = -r_{12} / (r_{11} r_{22})^{1/2} \text{ if } Q_1 = e_1 \text{ and } Q_2 = e_2. \]

Let X_1 and X_2 be two sets of variables with the joint dispersion matrix \( \Sigma \), partitioned accordingly.

**Result 3.** The squared canonical correlations are the non-zero roots of the determinantal equation

\[ | \Sigma_{11}^{-} \Sigma_{12}^{-} \Sigma_{22}^{-} \Sigma_{21}^{-} - \rho^2 I | = 0 \text{ where } \Sigma_{11}^{-} \text{ and } \Sigma_{22}^{-} \text{ are any g-inverses of } \Sigma_{22} \text{ and } \Sigma_{11} \text{ respectively.} \]

For proofs and further discussions on the results see Rao [(1979), (1980).]

3. **Generalized canonical correlations in the singular case.** Canonical correlations have been generalized in various ways. Formulae in the general case will be provided here for those obtained by extending the concepts of tests of independence for two sets of
variates--giving rise to partial, part and bipartial canonical correlations [see Timm (1975) pp. 352-353] and $g_1$-bipartial canonical correlations [see Lee (1978)] and some association measures [see McKeon (1965), pp. 16-19]. Various other authors [Horst (1961); Edgerton and Kolbe (1936); Wilks (1938); Lord (1958)] arrived at the same solution as that of McKeon for the particular case of a single variable per set. Appropriate formula will also be provided for the new generalized canonical correlation arising out of the concept of minimum generalized variance [proposed by Anderson (1958) Problem 5, pp. 305-306 and derived by the author {see SenGupta (1979)} under constraint of equi-correlation structure of the generalized canonical variables].

Let $X = (X_1, \ldots, X_k)$, $X_i : p_i \times 1$, $p_1 + \cdots + p_k = p$, $\text{Disp}(X) = k\Sigma$, $\text{Cov}(X_i, X_j) = \Sigma_{ij}$ and non-zero $p$s be the generalized canonical correlations. Starting with the defining equations it can be easily seen that for all the above cases, the generalized canonical correlations are obtained from the eigen values of $k\Sigma^*$ in the metric of $k\Sigma^*_d$, i.e. from the solutions of

\[(3.1) \quad |k\Sigma^* - \rho^* k\Sigma^*_d| = 0\]

where $\rho^* = 1 + (k - 1)\rho$ and $k\Sigma^*_d$ is a diagonal super matrix with elements $\Sigma^*_{ii}$ $i = 1, \ldots, k$. In the notation of Lee, $k\Sigma^*$, with $k = 2$, is the covariance matrix of the residual vectors $(\tilde{e}_{1.34}, \tilde{e}_{2.35})$ and $(e_{1.34}, e_{2.35})$ for the $g_1$- and
\( g_2 \)-bipartite canonical correlations, respectively. In the notation of Timm, \( k \Sigma^* = \Sigma \cdot 3^* \Sigma_1(2 \cdot 3) \) and \( \Sigma_1(1 \cdot 4)(2 \cdot 3) \) with \( k = 2 \) for partial, part and bipartial canonical correlations respectively. Also for McKeon's and the new generalized canonical correlations, \( k \Sigma^* = k \Sigma^* \).

**Theorem.** The generalized canonical correlations, for the methods quoted above, are given by \( \rho = (\rho^* - 1)/(k - 1) \) where \( \rho^* \)'s are the non-zero roots of \( |k \Sigma^* k d^* - \rho^* I| = 0 \), \( k \Sigma d^* \) being any \( g \)-inverse of \( k \Sigma d^* \).

**Proof.** First note the representation (3.1). Consider next the following Lemmas.

**Lemma 1.** Let \( A \) be a hermitian matrix of order \( n \) and rank \( s \), and \( B \) be non-negative definite matrix of order \( n \) and rank \( r \) such that \( S(A) \subseteq S(B) \) [where \( S(M) \) represents the vector space spanned by the column vectors of \( M \)]. Then

(i) There exists a matrix \( L \) of order \( nxr \) such that \( L^* A L = \Lambda \), \( L^* B L = I_r \), where \( \Lambda \) is a diagonal matrix with \( s \) non-zero elements, some of which may be repeated and \( I_r \) is the identity matrix of order \( r \).

(ii) The non-zero elements of \( \Lambda \) are the same as the roots of the determinantal equation, \( |A B^* - \lambda I| = 0 \) with repetitions allowed, for any \( g \)-inverse \( B^* \) of \( B \).

**Proof of Lemma 1.** See Lemma 3, Rao (1979).

**Lemma 2.** \( S(k \Sigma^*) \subseteq S(k \Sigma^*) \).

**Proof of Lemma 2.** Note that \( S(\Sigma_{ij}^*) \subseteq S(\Sigma_{ii}^*) \) for all the \( k \Sigma^* \) considered above. This follows immediately from the
result [see Proposition 3.31, pp. 3.15-3.16 of Eaton]

that, if \( A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \geq 0 \) then \( N(A_{22}) \subset N(A_{12}) \) and

and \( N(A_{11}) \subset N(A_{21}) \) where \( N(M) \) is the null space of \( M \).

Then, there exist matrices \( B_{ij} \) such that \( B_{ij} \Sigma^*_{jj} = \Sigma^*_{ij}, i, j = 1, \ldots, k \).

Hence, there exist matrices \( C_i \) such that \( (C_1 \ldots C_k)k^d = C_k \Sigma^* = k^d \Sigma^* \)

which proves Lemma 2.

Coupling Lemma 2 with Lemma 1 proves the Theorem.

Note: For \( k = 2 \), if \( p_1 = 1, p_2 > 1 \) and if \( p_1 > 1, p_2 > 1 \) then we have the cases of multiple and canonical correlations respectively. Further, with \( k = 2 \), consideration of residual variables leads to partial correlation. Thus the above Theorem unifies the Results 1 through 3 and also considers simple (and not squared) multiple, partial and canonical correlations.


Let \( A \) and \( B \) be two hermitian matrices and \( B \) be non-negative definite. If \( \lambda \) is a constant and \( v \) a vector such that

\( Av = \lambda Bv, Bv \neq 0 \), then \( \lambda \) is called a proper eigen value and \( v \) a proper eigen vector of \( A \) with respect to \( B \). In the context of Lemma 1, the elements of \( \lambda \) are called the proper eigen values and the corresponding columns of \( L \), the proper eigen vectors of \( A \) with respect to \( B \). For the generalized correlations, we consider from (3.1) only the proper eigen
values of $k^*_{\Sigma}$ with respect to $k^*_{\hat{d}}$. Also note that for $k \geq 2$, 1 and $-1/(k - 1)$ are the maximum and minimum possible values, respectively, for the generalized correlations. Let $k^*_{\od}$ be the super off-diagonal matrix such that $k^*_{\Sigma} = k^*_{\hat{d}} + k^*_{\od}$. Also let $R(M)$ denote the rank of the matrix $M$.

**Lemma 3.** The numbers of zero, unit and $-1/(k - 1)$-valued generalized correlations are given by $r = R(k^*_{\od})$, $r = R([k^*_{\od} - (k - 1)k^*_{\hat{d}}])$ and $r = R(k^*)$ respectively,

where $r = R(k^*_{\hat{d}})$.

**Proof:** The proof follows by rewriting (3.1) as

$$
|k^*_{\Sigma} - k^*_{\hat{d}| = 0 \text{ where}
$$

$$
(k^*_{\Sigma}, \lambda) = [k^*_{\od}, (k - 1)\rho], [k^*_{\od} - (k - 1)k^*_{\hat{d}} , (k - 1)(\rho - 1)]
$$

and $[k^*_{\Sigma}, (k - 1)\rho + 1]$ for the zero, unit and $-1/(k - 1)$-valued generalized correlations respectively and noting the one-one relationship between $\lambda$ and $\rho$.

**Acknowledgements.** I am grateful to Professor C. R. Rao and Professor T. W. Anderson for their kind remarks on the subject.
REFERENCES


[14] Tucker, L. R., Cooper, L. G. and Meredith, W. Obtaining squared multiple correlations from a correlation matrix which may be singular. Psychometrika 37, 143-148.


17. "General Exponential Models for Discrete Observations," 


Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)


Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," 
Mituaki Huzii, February 1976.


28. "Estimation of the Parameters of Finite Location and Scale Mixtures," 
Javad Behboodian, October 1976.


31. "Principal Components in the Nonnormal Case: The Test for Sphericity," 
Christine M. Wateraux, October 1977.


**REPORT DOCUMENTATION PAGE**

1. **REPORT NUMBER**
   46

2. **GOVT ACCESSION NO.**

3. **RECIPIENT'S CATALOG NUMBER**

4. **TITLE (and Subtitle)**
   GENERALIZED CORRELATIONS IN THE SINGULAR CASE

5. **TYPE OF REPORT & PERIOD COVERED**
   Technical Report

6. **PERFORMING ORG. REPORT NUMBER**

7. **AUTHOR(s)**
   ASHIS SEN GUPTA

8. **CONTRACT OR GRANT NUMBER(s)**
   N00014-75-C-0442

9. **PERFORMING ORGANIZATION NAME AND ADDRESS**
   Department of Statistics
   Stanford University
   Stanford, California

10. **PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS**
    (NR-042-034)

11. **CONTROLLING OFFICE NAME AND ADDRESS**
    Office of Naval Research
    Statistics and Probability Program Code 436
    Arlington, Virginia 22217

12. **REPORT DATE**
    November 1980

13. **NUMBER OF PAGES**
    8

14. **MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)**

15. **SECURITY CLASS. (of this report)**
    UNCLASSIFIED

15a. **DECLASSIFICATION/Downgrading Schedule**

16. **DISTRIBUTION STATEMENT (of this Report)**
    APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. **DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)**

18. **SUPPLEMENTARY NOTES**

19. **KEY WORDS (Continue on reverse side if necessary and identify by block number)**
   Generalized canonical correlations, g-inverse, hermitian matrix, singular covariance matrix.

20. **ABSTRACT (Continue on reverse side if necessary and identify by block number)**
    A general result is given which provides appropriate formulae for various generalizations of canonical correlations in the singular case. This covers as special cases the results for multiple correlation due to Tucker, Cooper and Meredith (1972) and Khatri (1976) and for partial and canonical correlations due to Rao (1979), (1980). The numbers of zero, unit and other critical generalized correlations are also given for the general case.