Res. Lett. Inf. Math. Sci., 2005, Vol.**8**, pp53-76
Available online at http://iims.massey.ac.nz/research/letters/

Antieigenvalues and antisingularvalues of a matrix and applications to problems in statistics

C. Radhakrishna Rao

Statistics Department
Joab Thomas Building
Pennsylvania State University
University Park, PA 16802

Let A be $p \times p$ positive definite matrix. A p-vector x such that $Ax = \lambda x$ is called an eigenvector with the associated with eigenvalue λ. Equivalent characterizations are:

(i) $\cos\theta = 1$, where θ is the angle between x and Ax.

(ii) $(x'Ax)^{-1} = xA^{-1}x$.

(iii) $\cos\Phi = 1$, where ϕ is the angle between $A^{1/2}x$ and $A^{-1/2}x$.

We ask the question what is x such that $\cos\theta$ as defined in (i) is a minimum or the angle of separation between x and Ax is a maximum. Such a vector is called an anti-eigenvector and $\cos\theta$ an anti-eigenvalue of A. This is the basis of operator trigonometry developed by K. Gustafson and P.D.K.M. Rao (1997), *Numerical Range: The Field of Values of Linear Operators and Matrices*, Springer. We may define a measure of departure from condition (ii) as $\min[(x'Ax)(x'A^{-1}x)]^{-1}$ which gives the same anti-eigenvalue. The same result holds if the maximum of the angle Φ between $A^{1/2}x$ and $A^{-1/2}x$ as in condition (iii) is sought. We define a hierarchical series of anti-eigenvalues, and also consider optimization problems associated with measures of separation between an $r(< p)$ dimensional subspace S and its transform AS.

Similar problems are considered for a general matrix A and its singular values leading to anti-singular values.

Other possible definitions of anti-eigen and anti-singular values, and applications to problems in statistics will be presented.

1 Introduction

Let A be $p \times p$ positive definite (pd) matrix. A p-vector x such that $Ax = \lambda x$ is called an eigenvector of A associated with the eigenvalue λ of A. Equivalent characterizations, among many others, are

$$(i) \quad \cos(x, Ax) = (x'Ax)/\sqrt{(x'x)(x'A^2x)} = 1 \tag{1.1}$$

$$(ii) \quad x'Ax - (x'A^{-1}x)^{-1} = 0 \text{ with } x'x = 1. \tag{1.2}$$

where $\cos(x, Ax)$ is the cosine of the angle between the vectors x and Ax.

In statistical and computational problems, we are interested in a vector x for which there is maximum departure from the equations (1.1) and (1.2). For instance

$$\min_x \frac{x'Ax}{\sqrt{(x'x)(x'A^2x)}} = \frac{2\sqrt{\lambda_1\lambda_p}}{\lambda_1 + \lambda_p} \tag{1.3}$$

is attained at

$$x^* = \frac{\sqrt{\lambda_p}}{\sqrt{\lambda_1 + \lambda_p}} x_1 \pm \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1 + \lambda_p}} x_p \tag{1.4}$$

where $x_1, \ldots, x_p$ are eigenvectors corresponding to eigenvalues $\lambda_1 \geq \ldots \geq \lambda_p$. The result (1.3) is independently discovered by various authors, [Frucht (1943), Kantorovich (1952), Wielandt (1953), Gustafson (1968) and Krein (1969)]. Gustafson calls x^* in (1.4) an antieigenvector and the right hand side of (1.3), an antieigenvalue of A. In a series of papers Gustafson (1999, 2000a, 2000b, 2002) and Gustafson and Rao (1997) gave a geometrical interpretation of antieigenvalues and antieigenvectors and their applications to several problems. Khattree (2001, 2002, 2003) made some extensions of these results and provided some statistical and computational details involving antieigenvalues.

Shisha and Mond (1967) derived the inequality

$$\max_{x'x=1} \left[x'Ax - (x'A^{-1}x)^{-1}\right] = \left(\sqrt{\lambda_1} - \sqrt{\lambda_p}\right)^2 \tag{1.5}$$

which provides the maximum departure from the relation (1.2). The maximum is attained at

$$x^* = \left[\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1} + \sqrt{\lambda_p}}\right]^{1/2} x_1 \pm \left[\frac{\sqrt{\lambda_p}}{\sqrt{\lambda_1} + \sqrt{\lambda_p}}\right]^{1/2} x_p. \tag{1.6}$$

A generalization of (1.1)-(1.4) is as follows. Let X be a $p \times r$ matrix of rank r. The cosines of the angles between the vector spaces generated by the column vectors of X and those of AX are square roots of the eigenvalues of

$$R = (X'X)^{-1/2}X'AX(X'A^2X)^{-1}X'AX(X'X)^{-1/2} \tag{1.7}$$

which reduces to identity matrix I of order r when the columns of X are eigenvectors of A. Two measures of departure of R from I are the product and sum of the eigenvalues of (1.7). The minimum values of these measures and their applications are considered in Section 2.2.

Another characterization of a matrix X of any set of eigenvectors of A is

$$X'AX - \left(X'A^{-1}X\right)^{-1} = 0 \tag{1.8}$$

and a measure of departure from (1.8) is

$$\text{trace}\left(X'AX - (X'A^{-1}X)^{-1}\right), \text{ with } X'X = I. \tag{1.9}$$

The maximum value of (1.9) is obtained in Section 4.3.

A general problem of interest is the extension of the concepts of antieigenvalues and antieigenvectors of a matrix A to eigenvalues and eigenvectors of a matrix B with respect to a positive definite matrix A arising from the determinantal equation $|B - \lambda A| = 0$. This leads to minimization of a function of the type

$$(x'Cx)^2/(x'Ax)(x'Bx) \tag{1.10}$$

which is considered in Section 6.

Some problems in statistics require optimization of expressions like

$$(x'Ay)^2/(x'Ax)(y'Ay) \tag{1.11}$$

and functions of

$$(X'AX)^{-1/2}X'AY(YAY')^{-1}Y'AX(X'AX)^{-1/2} \tag{1.12}$$

where X and Y are matrices. These are considered in Section 3.2.

Notation: Throughout this paper $S(X)$, where X is a $p \times r$ matrix, represents a subspace of R^p spanned by the column vectors of X. The eigenvalues of a $p \times p$ positive definite (pd) matrix A are represented by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ and the corresponding eigenvectors by $x_1, \ldots, x_p$.

2 Kantorovich inequality and generalizations

2.1 Antieigenvalues and antieigenvectors

Let A be a $p \times p$ positive definite (pd) matrix. Then the cosine of the angle θ between a vector x and Ax is

$$\cos\theta = \frac{x'Ax}{\sqrt{(x'x)(x'A^2x)}} \tag{2.1}$$

which has the value unity if x is an eigenvector of A, i.e., $Ax = \lambda x$ for some λ. We raise the question: For what vector x, $\cos\theta$ takes the minimum value or the angle of separation between x and Ax is a maximum. The answer is provided by the Kantorovich inequality

$$1 \geq \frac{x'Ax}{\sqrt{(x'x)(x'A^2x)}}{}' \geq \frac{2\sqrt{\lambda_1\lambda_p}}{\lambda_1+\lambda_p} = \mu_1 \tag{2.2}$$

and the minimum value is attained at

$$x = \frac{\sqrt{\lambda_p}x_1 \pm \sqrt{\lambda_1}x_p}{\sqrt{\lambda_1+\lambda_p}} = (u_1, u_2). \tag{2.3}$$

The pair of vectors in (2.3) represented by (u_1, u_2), are called the first antieigenvectors and μ_1 in (2.2), the first antieigenvalue of A. The terminology was introduced by Gustafson (1968). The angle $\theta_1 = \cos^{-1}\mu_1$ is called an angle of the operator of A.

Now, we define

$$\mu_2 = \min_{x \perp x_1, x_p} \frac{x'Ax}{\sqrt{(x'x)(x'A^2x)}} \tag{2.4}$$

as the second antieigenvalue of A and the associated vectors (u_3, u_4), as the second antieigenvectors of A. Expressing $x = a_2x_2 + \ldots + a_{p-1}x_{p-1}$

$$\frac{x'Ax}{\sqrt{(x'x)(x'A^2x)}} = \frac{\lambda_2a_2^2 + \ldots + \lambda_{p-1}a_{p-1}^2}{\sqrt{(\sum a_i^2)(\sum \lambda_i^2a_i^2)}}$$

and applying Kantorovich inequality we find

$$\mu_2 = \frac{2\sqrt{\lambda_2\lambda_{p-1}}}{\lambda_2+\lambda_{p-1}}$$

and

$$(u_3, u_4) = \frac{\sqrt{\lambda_{p-1}}x_2 \pm \sqrt{\lambda_2}x_{p-1}}{\sqrt{\lambda_2 + \lambda_p}}. \tag{2.5}$$

We seek now the minimum of (2.1) subject to the condition $x \perp x_1, x_2, x_{p-1}, x_p$ which yields the third antieigenvalue μ_3 and the antieigenvectors (u_4, u_5), and so on. Thus we have

$$\mu_1 \le \mu_2 \le \ldots \le \mu_r,\ r = [p/2],\ \mu_i = 2\sqrt{\lambda_i \lambda_{p-i+1}}/(\lambda_i + \lambda_{p-i+1}) \tag{2.6}$$

the ordered antieigenvalues and the corresponding antieigenvectors $(u_1, u_2), \ldots,$ (u_{2r-1}, u_{2r}). When p is odd, the antieigenvalue of order $(p+1)/2$ is unity with the corresponding antieigenvector $x_{(p+1)/2}$.

2.2 Antieigensupspace

Consider the subspace $S(X)$ spanned by the columns of a matrix X of order $p \times r$ and rank $r \le p$. The squared cosines of the angles between the supspaces $S(X)$ and $S(AX)$ are eigenvalues of

$$(X'X)^{-1/2}X'AX(X'A^2X)^{-1}X'AX(X'X)^{-1/2} \tag{2.7}$$

which reduces to I_r (identity matrix of order r) when $S(X)$ is spanned by r eigenvectors of A.

Making the transformation $Y = A^{1/2}X$, the expression (2.7) can be written in a familiar form:

$$(Y'A^{-1}Y)^{-1/2}Y'Y(Y'AY)^{-1}Y'Y(Y'A^{-1}Y)^{-1/2}. \tag{2.8}$$

A measure of departure of (2.8) from I_r is the determinant of (2.8)

$$|Y'Y|^2/|Y'A^{-1}Y||Y'AY| \tag{2.9}$$

which is less than unity. We seek the minimum of (2.9). There are a number of proofs showing that

$$\frac{|Y'Y|^2}{|Y'A^{-1}Y||Y'AY|} \ge \mu_1\mu_2 \ldots \mu_r \tag{2.10}$$

where $\mu_1, \ldots, \mu_r$ are defined in (2.6) and the minimum is attained if $S(X)$ is spanned by the first r antieigenvectors

$$\frac{\sqrt{\lambda_{p-i+1}}}{\sqrt{\lambda_i + \lambda_{p-i+1}}}x_i + \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i + \lambda_{p-i+1}}}x_{p-i+1} \tag{2.11}$$

$i = 1, \ldots, r$.[Bloomfield and Watson (1975), Knott (1975)].

2.3 A statistical application: Efficiency of least squares estimator

2.3.1 A linear model with one covariate

Consider the linear model with one covariate

$$y = x\beta + \varepsilon, \ C(\varepsilon, \varepsilon) = \sigma^2 A \tag{2.12}$$

where y, x and ε are p-vectors, β is the regression coefficient and $C(u, v)$ represents the covariance of random variables u and v. Least squares estimator of β (assuming $A = I_p$) is

$$\tilde{\beta} = (x'x)^{-1}x'y \text{ with } V(\tilde{\beta}) = \sigma^2 x'Ax/(x'x)^2.$$

The minimum variance linear estimator of β is

$$\hat{\beta} = (x'A^{-1}x)^{-1}x'A^{-1}y \text{ with } V(\hat{\beta}) = \sigma^2(x'A^{-1}x)^{-1}.$$

The efficiency of $\tilde{\beta}$ compared to that of $\hat{\beta}$ is

$$\frac{V(\hat{\beta})}{V(\tilde{\beta})} = \frac{(x'x)^2}{(x'Ax)(x'A^{-1}x)} \geq \frac{4\lambda_1\lambda_p}{(\lambda_1 + \lambda_p)^2} = \mu_1^2 \tag{2.13}$$

so that the efficiency is above μ_1^2 for any covariate, using the result (2.2).

Another measure of efficiency is the squared correlation of $\tilde{\beta}$ and $\hat{\beta}$, which is the same as $V(\hat{\beta})/V(\tilde{\beta})$ with the minimum value μ_1^2.

2.3.2 A linear model with s covariates

Consider the linear model

$$y = X\beta + \varepsilon, \ C(\varepsilon, \varepsilon) = \sigma^2 A$$

where y and ε are p-vectors, X is a $p \times s$ matrix of rank s and β is an s-vector of unknown parameters. The least squares estimator of β is

$$\tilde{\beta} = (X'X)^{-1}X'y \text{ with } C(\tilde{\beta}, \tilde{\beta}) = \sigma^2(X'X)^{-1}X'AX(X'X)^{-1}. \tag{2.14}$$

The minimum covariance linear estimator of β is

$$\hat{\beta} = (X'A^{-1}X)^{-1}X'A^{-1}y \text{ with } C(\hat{\beta}, \hat{\beta}) = \sigma^2(X'A^{-1}X)^{-1}. \tag{2.15}$$

(i) A measure of relative efficiency is

$$\frac{|C(\hat{\beta},\hat{\beta})|}{|C(\tilde{\beta},\tilde{\beta})|} = \frac{|X'X|^2}{|X'AX||X'A^{-1}X|} \geq (\mu_1 \ldots \mu_s)^2 \tag{2.16}$$

using the result (2.10).

(ii) Another way of measuring efficiency is to consider the product of the squared canonical correlations between $\tilde{\beta}$ and $\tilde{\beta}$, which are the eigenvalues of

$$\left[C(\hat{\beta},\hat{\beta})\right]^{-1/2} C(\hat{\beta},\tilde{\beta}) \left[C(\tilde{\beta},\tilde{\beta})\right]^{-1} C(\tilde{\beta},\hat{\beta}) \left[C(\hat{\beta},\hat{\beta})\right]^{-1/2}. \tag{2.17}$$

Substituting

$$\begin{aligned} C(\hat{\beta},\tilde{\beta}) &= \sigma^2(X'A^{-1}X)^{-1} = C(\hat{\beta},\hat{\beta}) \\ C(\tilde{\beta},\tilde{\beta}) &= \sigma^2(X'X)^{-1}X'AX(X'X)^{-1} \end{aligned}$$

the expression (2.17) becomes

$$(X'A^{-1}X)^{-1/2}X'X(X'AX)^{-1}X'X(X'A^{-1}X)^{-1/2} \tag{2.18}$$

and the product of the eigenvalues of (2.18) is

$$\frac{|X'X|^2}{|X'A^{-1}X||X'AX|}$$

which is the same as (2.16) and hence has the same lower limit.

(iii) Another way of looking at the problem is to find the condition for $\tilde{\beta}$ and $\hat{\beta}$ to be equal. Using the expressions (2.14) and (2.15)

$$\tilde{\beta} - \hat{\beta} = \left[(X'X)^{-1}X' - (X'A^{-1}X)^{-1}X'A^{-1}\right] y.$$

The difference will be zero for all y if

$$\begin{aligned} &(X'X)^{-1}X' = (X'A^{-1}X)^{-1}X'A^{-1} \\ \Rightarrow &X'A = (X'X)(X'A^{-1}X)^{-1}X' \Rightarrow X'AZ = 0 \end{aligned} \tag{2.19}$$

where $S(Z)$ is the orthogonal complement of $S(X)$. Conversely if $X'AZ = 0$, then $(\tilde{\beta} - \hat{\beta}) = 0$.

Note 1. The condition $X'AZ = 0$ in (2.19) implies that A has the structure

$$A = XUX' + ZVZ' \tag{2.20}$$

for some matrices u and V. The result (2.20) was established in Rao (1967).

Note 2. The condition for equality of $\tilde{\beta}$ and $\hat{\beta}$ can also be derived from Theorem (i) proved in Rao (1973, p.317) that $\tilde{\beta}$ is a minimum covariance estimator iff it has zero covariance with the linear functions $Z'y$ which have zero expectation. Now

$$C(\tilde{\beta}, Z'y) = (X'X)^{-1}X'AZ = 0 \Rightarrow X'AZ = 0. \tag{2.21}$$

Note 3. The condition

$$X'AZ = 0 \Rightarrow P_X A^2 P_X - (P_X A P_X)^2 = 0 \tag{2.22}$$

which provides another measure of efficiency

$$\max_X \operatorname{trace}\left[P_X A^2 P_X - (P_X A P_X)^2\right] = \frac{1}{4}\sum_1 (\lambda_i - \lambda_{p-i+1})^2 \tag{2.23}$$

derived by Bloomfield and Watson (1975). [See Bartman and Bloomfield (1981) for related work].

The condition (2.22) can also be written as

$$\begin{aligned} &X'A^2X = X'AX(X'X)^{-1}X'AX \\ \Rightarrow &(X'A^2X)^{-1/2}X'AX(X'X)^{-1}X'AX(X'A^2X)^{-1/2} = I \end{aligned} \tag{2.24}$$

which provides measures of efficiency

$$|X'AX|^2/|X'X||X'A^2X| \tag{2.25}$$

and

$$\operatorname{trace}\left[(XA^2X)^{-1/2}X'AX(X'X)^{-1}X'AX(X'A^2X)^{-1/2}\right]. \tag{2.26}$$

3 Wielandt inequality and applications

3.1 Proof of Wielandt inequality

Consider two p-vectors x, y such that $x'y = 0$ (i.e., x and y are orthogonal and the problem:

$$\max_{x \perp y} \frac{(x'Ay)^2}{(x'x)(y'A^2y)} \tag{3.1}$$

i.e., minimizing the angle between x and Ay. Making the transformation

$$u = A^{-1/2}x, \ v = A^{1/2}y$$

the problem reduces to

$$\max_{u \perp v} \frac{(u'Av)^2}{(u'Au)(v'Av)}. \tag{3.2}$$

The following is known as Wielandt inequality

$$\frac{(u'Av)^2}{(u'Au)(v'Av)} \leq \left(\frac{\lambda_1 - \lambda_p}{\lambda_1 + \lambda_p}\right)^2 = 1 - \mu_1^2 \tag{3.3}$$

under the condition $u'v = 0$. [See the references, Wielandt (1953), Alparger (1996) and Davis and Schneider (1996)]. A simple proof of (3.3) is as follows.

Let U be a matrix of order $p \times (p-1)$ such that $u'U = 0$. Then $v = Ua$ for some $(p-1)$-vector a. Then (3.2) becomes

$$\frac{(u'AUa)^2}{(u'Au)(a'U'AUa)}. \tag{3.4}$$

By Cauchy-Schwarz inequality

$$(3.4) \leq \frac{u'AU(U'AU)^{-1}U'Au}{u'Au}. \tag{3.5}$$

Using the result (Rao (1973), p.77, example 32)

$$AU(U'AU)^{-1}U'A = A - u(u'A^{-1}u)^{-1}u'.$$

(3.5) becomes

$$1 - \frac{(u'u)^2}{(u'Au)(u'A^{-1}u)} \leq 1 - \mu_1^2 = \frac{(\lambda_1 - \lambda_p)^2}{(\lambda_1 + \lambda_p)^2} \tag{3.6}$$

which proves (3.3).

3.2 Generalization of Wielandt inequality

Let us consider the subspaces $S(X)$ and $S(Y)$ spanned by the columns of X of order $p \times r$ and of Y of order $p \times s$ respectively, such that $X'Y = 0$. What are X and Y such that $S(X)$ and $S(AY)$ are as close as possible? The squared cosines of the angles, $\rho_1^2, \ldots, \rho_m^2 [m = \min(r, s)]$ between $S(X)$ and $S(AY)$ are the eigenvalues of

$$(X'X)^{-1/2} X'AY(Y'A^2Y)^{-1}Y'AX(X'X)^{-1/2} \tag{3.7}$$

which, after suitable transformation can be written as

$$\Phi(X, Y) = (X'AX)^{-1/2} X'AY(Y'AY)^{-1}Y'AX(X'AX)^{-1/2}. \tag{3.8}$$

Let

$$G(X) = (X'AX)^{-1/2} X'X(X'A^{-1}X)^{-1}X'X(X'AX)^{-1/2}$$
$$P(X, Z) = (X'AX)^{-1/2} X'A^{1/2}Z(Z'Z)^{-1}Z'A^{1/2}X(X'AX)^{-1/2}$$

where $S(Z)$ is the orthogonal compliment of $S(A^{-1/2}X)$ and $S(A^{1/2}Y)$. Using the identity

$$I = A^{-1/2}X(X'A^{-1}X)^{-1}X'A^{-1/2} + A^{1/2}Y(Y'AY)^{-1}Y'A^{1/2} + Z(Z'Z)^{-1}Z'.$$

$\Phi(X, Y)$ in (3.8) can be written as

$$I - G(X) - P(X, Z) \tag{3.9}$$

and

$$I - \Phi(X, Y) = G(X) + P(X, Z) \geq G(X) \tag{3.10}$$

since $P(X, Z)$ is nnd. Denoting the squared cosines of angles between $S(X)$ and $S(AY)$ by $\rho_1^2, \ldots, \rho_m^2, [m = \min(r, s)]$,

$$|I - \Phi(X, Y)| = \prod_{i=1}^{m} (1 - \rho_i^2) = |G(X) + P(X, Z)|$$
$$\geq |G(X)| \geq \prod_{i=1}^{m} \frac{4\lambda_i \lambda_{p-i+1}}{(\lambda_i + \lambda_{p-i+1})^2} = \prod_{i=1}^{m} \mu_i^2 \tag{3.11}$$

using (2.10). Also

$$|\Phi(X,Y)| = \prod_{i=1}^{m} \rho_i^2 = |I - G(X) - P(X,Z)|$$
$$\leq |I - G(X)| \leq \prod_{i=1}^{m} \frac{(\lambda_i - \lambda_{p-m+i})^2}{(\lambda_i + \lambda_{p-m+i})^2}. \tag{3.12}$$

The results (3.11) and (3.12) are given in Khatri (1978), Khatri and Rao (1981) and Khatri and Rao (1982). For related work when X is a vector, reference may be made to Eaton (1976).

3.3 Statistical application: Sphericity tests

Wielandt's inequality (3.3) is used for constructing some test criteria in multivariate analysis. Let x be a p-vector variable with mean μ and variance covariance matrix $\Sigma = E[(x-\mu)(x-\mu)']$. We want to test the hypothesis H_0 against H_1,

$$H_0 : \Sigma = \sigma^2 I_p \ (I_p \text{ is identity matrix}),$$
$$H_1 : \Sigma \text{ is arbitrary.}$$

Test 1: If $\Sigma = \sigma^2 I_p$, then $k_1' A k_2 = 0$ for any two orthogonal vectors k_1 and k_2 (i.e., $k_1' k_2 = 0$). This condition can be used to construct the test criterion

$$C_1 = \max_{k_1 \perp k_2} \frac{(k_1' A k_2)^2}{(k_1' A k_1)(k_2' A k_2)} \tag{3.13}$$

where A is an estimate of Σ based on a sample, and k_1 and k_2 are normalized vectors. Using Wielandt inequality (3.3), the test criterion is

$$C_1 = \frac{(\lambda_1 - \lambda_p)^2}{(\lambda_1 + \lambda_p)^2} = 1 - \mu_1^2 \tag{3.14}$$

where λ_1 and λ_p are the largest and smallest eigenvalues of A. A large value of C_1 or a small value of μ_1 indicates departure from H_0.

Test 2: Let K_1 be a $p \times q$ matrix of rank q and K_2 be a $p \times (p-q)$ matrix of rank $(p-q)$ such that $K_1' K_2 = 0$ (null matrix). Also let $q \leq p - q$.

Consider random variables $u = K_1' x$ and $v = K_2' x$. If the covariance matrix of x is $\sigma^2 I$, then $K_1' x$ and $K_2' x$ are uncorrelated. If A is the estimated covariance matrix of x based on a sample of observations on x, then the estimated squared canonical correlations between $K_1' x$ and $K_2' x$ are the eigenvalues of

$$
\begin{aligned}
&(K_1'AK_1)^{-1/2}K_1'AK_2(K_2'AK_2)^{-1}K_2'AK_1(K_1'AK_1)^{-1/2} \\
&= I-(K_1'AK_1)^{-1/2}(K_1'A^{-1}K_1)^{-1}(K_1'AK_1)^{-1/2} = \Phi(K_1). \qquad (3.15)
\end{aligned}
$$

Then

$$
\begin{aligned}
|\Phi(K_1)| &= \hat{\rho}_1 \dots \hat{\rho}_q^2 \\
|I-\Phi(K_1)| &= (1-\hat{\rho}_1^2)\dots(1-\hat{\rho}_q^2) = |(K_1'A^{-1}K_1)(K_1'AK_1)|^{-1}
\end{aligned}
$$

where $\hat{\rho}_1^2, \dots, \hat{\rho}_q^2$ are the estimated squared canonical correlations. We may choose the test statistic as

$$
C_2 = \max_{K_1} |K_1'A^{-1}K_1||K_1'AK_1| = \prod_{i=1}^{q} \frac{(\lambda_i+\lambda_{p-i+1})^2}{4\lambda_i\lambda_{p-i+1}}. \qquad (3.16)
$$

A large value of C_2 indicates rejection of H_0.

Test 3: Another possible test criterion is

$$
\begin{aligned}
C_3 = \max_{K_1}\Phi(K_1) &= \max_{K_1}\prod_{i=1}^{q}\rho_i^2 \\
&= \prod_{i=1}^{q}\frac{(\lambda_i-\lambda_{n-q+i})^2}{(\lambda_i+\lambda_{n-q+i})^2}. \qquad (3.17)
\end{aligned}
$$

A large value of C_3 indicates departure from the hypothesis H_0 of sphericity. For a discussion of these tests reference may be made to Venebles (1976).

4 Shisha-Mond inequality and generalizations

4.1 (SM)-antieigenvalues and antieigenvectors

An eigenvector x of a pd matrix A can be characterized in many ways other than that the angle between x and Ax is zero. An interesting characterization is

$$
x'Ax = (x'A^{-1}x)^{-1}, \; x'x = 1. \qquad (4.1)
$$

By Cauchy-Schwarz inequality,

$$
(x'Ax)(x'A^{-1}x) \geq 1
$$

so that

$$d = x'Ax - (x'A^{-1}x)^{-1} \geq 0 \tag{4.2}$$

with equality holding when x is an eigenvector of A. Shisha and Mond (1967) have shown that [see also Styan (1983)],

$$\max_{|x|=1} \left[x'Ax - (x'A^{-1}x)^{-1}\right] = \left(\sqrt{\lambda_1} - \sqrt{\lambda_p}\right)^2 = \nu_1 \tag{4.3}$$

so that (4.2) has an upper bound and is attained when x is

$$(z_1 \text{ or } z_2) = \left[\sqrt{\lambda_1}/(\sqrt{\lambda_1} + \sqrt{\lambda_p})\right]^{1/2} x_1 \pm \left[\sqrt{\lambda_p}/(\sqrt{\lambda_1} + \sqrt{\lambda_p})\right]^{1/2} x_p. \tag{4.4}$$

We call ν_1 as the first (SM)-antieigenvalue and z_1 or z_2 as the first (SM)-antieigenvector. [(SM) stands for Shisha-Mond. By analogy, Gustafson antieigenvalue may be called (K)-antieigenvalue as it is based on Kolmogorov inequality].

4.2 Higher order (SM)-antieigenvalues and antieigenvectors

As in Section 2.1, we seek

$$\max_{\substack{x \perp z_1, z_2 \\ x'x=1}} \left[x'Ax - (x'A^{-1}x)^{-1}\right].$$

Using the same type of argument as in Section 2.1 we find

$$\max_{\substack{x \perp z_1, z_2 \\ x'x=1}} \left[x'Ax - (x'A^{-1}x)^{-1}\right] = \left(\sqrt{\lambda_2} - \sqrt{\lambda_{p-1}}\right)^2 = \nu_2 \tag{4.5}$$

and the maximum is attained at x equal to

$$(z_3 \text{ or } z_4) = \left[\sqrt{\lambda_2}/(\sqrt{\lambda_2} + \sqrt{\lambda_{p-1}})\right]^{1/2} x_2 \pm \sqrt{\lambda_{p-1}}/\left[(\sqrt{\lambda_2} + \sqrt{\lambda_{p-1}})\right]^{1/2} x_{p-1}. \tag{4.6}$$

We this build up the series

$$(\nu_1; z_1, z_2), (\nu_2; z_3, z_4), \ldots, (\nu_r; z_{2r-1}, z_{2r}) \tag{4.7}$$

where $r = [p/2]$ and $\nu_i = \left(\sqrt{\lambda_i} - \sqrt{\lambda_{p-i+1}}\right)^2$.

4.3 Generalized Shisha-Mond inequality

Let X be a matrix of order $p \times r$ such that $X'X = I_r$. For any given X,

$$X'AX - (X'A^{-1}X)^{-1} \text{ is nnd} \tag{4.8}$$

and is a null matrix if and only if $S(X)$ is spanned by r eigenvectors of A. Now consider

$$\begin{aligned}
&\left(X'A^{1/2} - (X'A^{-1}X)^{-1}X'A^{-1/2}\right)\left(A^{1/2}X - A^{-1/2}X(X'A^{-1}X)^{-1}\right)\\
&= X'AX - (X'A^{-1}X)^{-1} = 0 \Rightarrow A^{1/2}X - A^{-1/2}X(X'A^{-1}X)^{-1} = 0\\
&\Rightarrow AX = XB,\ B = (X'A^{-1}X)^{-1}.
\end{aligned}$$

Using the spectral decomposition of $B = Q\Delta Q'$,

$$AX = XB \Rightarrow AX = XQ\Delta Q' \Rightarrow AXQ = XQ\Delta \tag{4.9}$$

i.e., XQ is a matrix whose columns are eigenvectors of A. This implies that $S(X)$ is spanned by a set of eigenvectors of A and the if part if proved. The only if part follows easily.

A measure of departure from $X'AX - (X'A^{-1}X)^{-1}$ is

$$\underset{X'X=I}{\text{trace}} \left[X'AX - (X'A^{-1}X)^{-1}\right]. \tag{4.10}$$

It is shown in Rao (1985) that

$$(4.10) \leq \nu_1 + \ldots + \nu_m \tag{4.11}$$

where ν_i is the i-th (SM) antieigenvalue of A as defined in (4.7) and $m = \min(r, p-r)$. In Drury, Liu, Lu, Puntanen and Styan (2000), the inequality (4.11) is referred to as Rao inequality.

An interesting result arising out of (4.10) is the inequality

$$\begin{aligned}
\text{trace}\left[A_{11} - (A^{11})^{-1}\right] &= \text{trace}\left(A_{12}A_{22}^{-1}A_{21}\right)\\
&\leq \sum_1^m \left(\sqrt{\lambda_i} - \sqrt{\lambda_{p-i+1}}\right)^2
\end{aligned}$$

where A_{ij} and A^{ij} are the parts of

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},\ A^{-1} = \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix}.$$

5 Some related inequalities

5.1 (K)-antisingularvalues and vectors

Let A be a matrix of order $p \times p$ with the singular value decomposition (SVD)

$$A = \delta_1 x_1 y_1' + \ldots + \delta_p x_p y_p' \tag{5.1}$$

where $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_p > 0$ are singularvalues of A, x_i and y_i such that $x_i' x_i = y_i' y_i = 1$, $i = 1, \ldots, p$ are left and right singularvectors. A natural extension of Kantorovich inequality is (with $x'x = 1 = y'y$),

$$(x'Ay)(x'A^{-1}y) \leq \frac{(\delta_1 + \delta_p)^2}{4\delta_1\delta_p} = \omega_1 \tag{5.2}$$

and the maximum is attained at

$$x = (u_1 \text{ or } u_2) = \frac{1}{\sqrt{2}}(x_1 \pm x_p),\ y = (v_1 \text{ or } v_2) = \frac{1}{\sqrt{2}}(v_1 \pm v_p). \tag{5.3}$$

We call ω_1 as the first (K)-antisingularvalue with $(u_1, u_2), (v_1, v_2)$ as (K)-antisingularvectors.

As in Section 2.1, we can show that

$$\max_{\substack{x \perp u_1 u_2\, y \perp v_1, v_2 \\ |x| = |y| = 1}} (x'Ay)(y'A^{-1}x) = \frac{(\delta_2 + \delta_{p-1})^2}{4\delta_2\delta_{p-1}} = \omega_2 \tag{5.4}$$

and the maximum is attained at

$$x = (u_3 \text{ or } u_4) = \frac{1}{\sqrt{2}}(x_2 \pm x_{p-1}),\ y = (v_3 \text{ or } v_4) = \frac{1}{\sqrt{2}}(y_2 \pm y_{p-1}). \tag{5.5}$$

Thus we obtain the sequence of (K)-antisingularvalues and singularvectors

$$(\omega_1; u_1, u_2, v_1, v_2), (\omega_2; u_3, u_4, v_3, v_4), \ldots, (\omega_r; u_{2r-1}, u_{2r}, v_{2r-1}, v_{2r}) \tag{5.6}$$

where $r = \left[\frac{p}{2}\right]$, and $\omega_i = (\delta_i + \delta_{p-i+1})^2 / 4\delta_i\delta_{p-i+1}$.

5.2 A generalization

Let X and Y be $p \times r$ and $p \times s$ matrices of ranks r and s respectively with $r \geq s$ and $X'X = I_r,\ Y'Y = I_s$ and A be a nonsingular matrix with the SVD

$$A = \delta_1 x_1 y_1' + \ldots + \delta_p x_p y_p'. \tag{5.7}$$

The following inequalities have been proved in Khatri and Rao (1981, 1982).

$$(i) \quad |X'AYY'A^{-1}X| \leq \prod_{i=1}^{\min(r,n-s)} \omega_i \tag{5.8}$$

$$\begin{aligned}(ii) \quad \text{trace}(X'AYY'A^{-1}X) &\leq \sum_{i=1}^{r} \omega_i \text{ if } p \geq r+s \\ &\leq \left(\sum_{i=1}^{p-s} \omega_i\right) + (r+s-p) \text{ if } p < r+s \end{aligned} \tag{5.9}$$

where

$$\omega_i = (\delta_i + \delta_{p-i+1})^2 / 4\delta_i \delta_{p-i+1}.$$

The result (5.8) generalizes an inequality due to Strang (1960).

Other inequalities proved by Khatri and Rao (1981, 1982) are as follows. let B and C be symmetric nonsingular matrices such that $BC = CB$ and BC^{-1} is pd, X be $p \times r$ matrix of rank r and $\lambda_1 \geq \cdots \lambda_p$ be the eigenvalues of BC^{-1}. Then

$$(i). \quad \frac{|X'B^2X||X'C^2X|}{|X'BCX|^2} \leq \prod_{i=1}^{m} \frac{(\lambda_i + \lambda_{p-i+1})^2}{4\lambda_i \lambda_{p-i+1}} \tag{5.10}$$

$$\begin{aligned}(ii). \quad &\text{trace}\left(X'B^2X(X'BCX)^{-1}X'C^2X(X'BCX)^{-1}\right) \\ &\leq \sum_{i=1}^{m} \frac{(\lambda_i + \lambda_{p-i+1})^2}{4\lambda_i \lambda_{p-i+1}} + (p-m)\end{aligned} \tag{5.11}$$

where $m = \min(r, p-r)$. The result (5.10) generalizes the inequality of Greub and Rheinboldt (1959) proved for the special case of $r = 1$.

The inequalities (5.8)-(5.11) are referred to as Khatri-Rao inequalities in Drury, Liu, Lu, Puntanen and Styan (2000).

6 Antieigenvalues of A with respect to B

Let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of an nnd $p \times p$ matrix A with respect to a pd matrix B, i.e., solutions of the matrix equation

$$|A - \lambda B| = 0. \tag{6.1}$$

Corresponding to a root λ_i, there is an eigenvector x_i such that

$$Ax_i = \lambda_i Bx_i.$$

The cosine of the angle between Ax_i and Bx_i, $\cos(Ax_i, Bx_i) = 1$. We define the first antieigenvalue as

$$\begin{aligned}\min_x [\cos(Ax, Bx)] &= \min_x \frac{x'ABx}{\sqrt{(x'A^2x)(x'B^2x)}} \\ =&\min_y \frac{(y'Cy)}{\sqrt{(y'y)(y'CC'y)}}, \ C = B^{-1}A. \end{aligned} \tag{6.2}$$

Case 1. C is symmetric.

The ratio (6.2) reduces to

$$\frac{y'Cy}{\sqrt{(y'y)(y'C^2y)}}$$

and the minimum value is, by Kantorovich inequality,

$$\mu_1 = \frac{2\sqrt{\lambda_1\lambda_p}}{\lambda_1+\lambda_2}.$$

Case 2. C is not symmetric.

There is no closed form solution. The vector y at which (6.2) takes a stationary value is a solution of the equation

$$\begin{gathered}(C+C')u = \alpha y + \beta CC'y \\ \alpha = \frac{y'(C+C')y}{y'y}, \ \beta = \frac{y'(C+C')y}{y'CC'y}\end{gathered} \tag{6.3}$$

which can be written in another form

$$\begin{aligned}(C+C')y &= \alpha(I+\nu CC')y \\ \nu &= y'y/y'CC'y\end{aligned} \tag{6.4}$$

and also as

$$\begin{aligned}(C+C')y &= \beta(CC'+\nu I)y \\ \nu &= y'CC'y/y'y.\end{aligned} \tag{6.5}$$

In (6.5), $\nu \in [\nu_p, \nu_1]$, where ν_1 and ν_p are the largest and smallest eigenvalues of CC'. For any given ν, we have from first equation in (6.5) p roots and corresponding eigenvectors

$$\beta_i(\nu), y_i(\nu),\ i = 1, \dots, p \tag{6.6}$$

which are continuous functions of ν [Kato (1980)]. Define the function

$$g_i(\nu) = \frac{y_i'(\nu)CC'y_i(\nu)}{y_i'(\nu)y_i(\nu)},\ i = 1, \dots, p. \tag{6.7}$$

We have to pick up those values of (i, ν) for which $g_i(\nu) = 0$. Let ν_i be the value at which $g_i(\nu_i) = 0$. Compute

$$\beta_i(\nu_i) = \frac{y_i'(\nu_i)(C + C')y_i(\nu_i)}{y_i'(\nu_i)(CC' + \nu_i I)y_i(\nu_i)}. \tag{6.8}$$

The stationary values of the function

$$\cos^2\theta = (y'Cy)^2/(y'y)(y'CC'y)$$

are provided by

$$\beta_i^2\nu_i,\ i = 1, \dots, p \tag{6.9}$$

from which the minimum or maximum value can be chosen.

7 Homologous canonical correlations

Let u and v be p-vector random variables with the joint covariance matrix

$$\begin{pmatrix} A & C \\ C' & B \end{pmatrix}. \tag{7.1}$$

For instance u may be p measurements on a parent and v the corresponding measurements taken on an offspring. In the theory of canonical correlations, linear functions $a'u$ and $b'v$, where a and b are different, are sought to maximize the correlation between $a'u$ and $b'v$. We raise the question: for what linear function of the measurements the parent offspring correlation is a maximum. We consider linear functions $x'u$ and $x'v$ and compute their correlation coefficient

$$\rho = \frac{x'Cx}{\sqrt{(x'Ax)(x'Bx)}}. \tag{7.2}$$

The values of (7.2) at stationary points have been termed as homologous canonical correlations by Rao and Rao (1987).

To obtain the stationary values of (7.2), we equate the derivative of (7.2) with respect to x to zero [Rao (1973), p.72]. This yields the equations

$$\begin{aligned}\lambda Ax + \mu Bx &= 2Cx \\ \lambda x'Ax &= x'Cx\end{aligned} \tag{7.3}$$

introducing two additional constants λ and μ, which is equivalent to the equations

$$\begin{aligned}\lambda(A + \nu B)x &= 2Cx \\ x'Ax &= \nu x'Bx\end{aligned} \tag{7.4}$$

introducing two additional variables λ and ν.

Since A and B are pd matrices, there exists a nonsingular transformation S such that $A = S\Delta S'$ and $B = SS'$ where Δ is a diagonal matrix [Rao (1973), p.41]. Then writing $y = Sx, W = S^{-1}C(S^{-1})'$, the equations (7.4) assume the simpler form

$$\begin{aligned}\lambda(\Delta + \nu I)y &= 2Cy \\ y'\Delta y &= \nu y'y.\end{aligned} \tag{7.5}$$

If $\delta_1, \dots, \delta_p$ are the diagonal elements of Δ and $y_1, \dots, y_p$ are the components of y, we obtain the equations for $y_1, \dots, y_p$ from (7.5) as

$$2y'y\,[(e_i'Cy)y_1 - (e_1'Cy)y_i] = y_1 y_i(\delta_i - \delta_1)y'Cy,\ i = 1, \dots, p \tag{7.6}$$

where e_i is the elementary vector with unity as the i-th component and zeros elsewhere. In (7.6). we have $(p-1)$ quartic equations in $(p-1)$ ratios $(x_2/x_1), \dots, (x_p/x_1)$. The solution of these equations poses a complicated computational problem except in the case of $p = 2$ when we have one quartic equation as observed by Kouvaritakis and Cameron (1980).

Rao and Rao (1987) discuss a general computational method for obtaining the stationary values of (7.2) using the equation (7.4)

$$\begin{aligned}2Cx &= \lambda(A + \nu B)x \\ x'Ax &= \nu x'Bx.\end{aligned} \tag{7.7}$$

Observe that $\nu\varepsilon[\nu_p,\nu_1]$ where $\nu_1 \geq \ldots \geq \nu_p$ are the eigenvalues of A with respect to B, i.e., the roots of $|A-\nu B|=0$. For any given $\nu\varepsilon[\nu_p,\nu_1]$, the first equation in (7.7) provides p eigenvalues

$$\lambda_1(\nu) \geq \ldots \geq \lambda_p(\nu) \tag{7.8}$$

of $2C$ with respect to $(A+\nu B)$ and p associated eigenvectors

$$x_1(\nu), \ldots, x_p(\nu). \tag{7.9}$$

The pair $(\nu, x_i(\nu))$ will be a solution of (7.7) if and only if

$$\nu = x_i'(\nu)Ax_i(\nu)/x_i'(\nu)Bu_i(\nu). \tag{7.10}$$

The method is illustrated with an example in Rao and Rao (1987). It would be of interest to find a suitable algorithm to solve the equations (7.7).

The special case of (7.2) with $C=I$ (which reduces to the product of two Rayleigh quotients) originally arose in attempts to design control systems with minimum norm feedback matrices [Kouvaritakis and Cameron (1980), Cameron and Kouvaritakis (1980)] and also in the study of the stability of multivariable nonlinear feedback systems [Cameron (1983)].

8 Additional remarks

Ando (2000a,b,c) extended the Kantorovich and Shisha-Mond inequalities to a "compression" $\Phi_C(A)$ of a unital positive map $\Phi(A)$ from $p\times p$ matrices to $r\times r$ matrices. The map Φ between C-algebras is said to be unital positive if it is unit-preserving and positivity-preserving, respectively. For a pd matrix A, he showed that

$$\left|\{\Phi_C(A)\}^{-1}\Phi_C(A^2)\{\Phi_C(A)\}^{-1}\right| = \frac{|\Phi_C(A^2)|}{|\Phi_C(A)|^2} \leq \prod_{i=1}^{m} \frac{(\lambda_i+\lambda_{p-i+1})^2}{4\lambda_i\lambda_{p-i+1}} \tag{8.1}$$

$$\operatorname{trace}\left(\Phi_C(A)-\{\Phi_C(A^{-1})\}^{-1}\right) \leq \sum_{i=1}^{m}\left(\sqrt{\lambda_i}-\sqrt{\lambda_{p-i+1}}\right)^2 \tag{8.2}$$

where $m=\min(r,n-r)$. He also established a series of majorization results:

$$\{\xi_i\}_{i=1}^m \prec_w \left\{\frac{1}{4}(\lambda_i-\lambda_{p-i+1})^2\right\}_{i=1}^m \tag{8.3}$$

where $\xi_1, \ldots, \xi_m$ are the eigenvalues of $X'A^2X - (X'AX)^2$, and

$$\{\log \eta_i\}_{i=1}^m \prec_w \left\{ \log \left[\frac{(\lambda_i + \lambda_{p-i+1})^2}{4\lambda_i \lambda_{p-i+1}} \right] \right\}_{i=1}^m \tag{8.4}$$

where η_i are eigenvalues of

$$\{\Phi_C(A)\}^{-1} \Phi_C(A^2) \{\Phi_C(A)\}^{-1}. \tag{8.5}$$

Ando mentions that Bloomfield-Watson-Knott, Khatri-Rao and Rao inequalities are three "most essential" results for deriving his majorization results.

References

Alparger, G. (1996). The Kantorovich inequality, with some extensions and with some statistical applications. Thesis submitted to McGill University, Montreal, for the degree of Master of Science.

Ando, T. (2000a). Majorization of eigenvalues. Lecture notes for lecture presented at the Special Conference on Functional Analysis and Linear Algebra, Indian Statistical Institute - Delhi Centre, January 3-7, 2000, preprint.

Ando, T. (2000b).Bloomfield-Watson type inequalities for eigenvalues. Lecture notes for lecture presented at the International Conference on Mathematical Analysis and its Applications (ICMAA-2000), National Sun Yat-sen University, Taiwan, January 17-21, 2000, preprint.

Ando, T. (2000c). From determinant and trace inequalities to majorization of eigenvalues. Lecture notes for invited lecture to be presented at the Ninth International Workshop on matrices and Statistics, in Celebration of C.R. Rao's 80th Birthday (Hyderabad, India, December 9-13, 2000), preprint.

Bartman, F.C. and Bloomfield, P. (1981). Inefficiency and correlation, *Biometrika*, **68**, 67-71.

Bloomfield, P. and Watson, G.S. (1975). The inefficiency of least squares, *Biometrika*, **62**, 121-128.

Cameron, R.G. and Kouvaritakis, B. (1980). Minimizing the norm of output feedback controllers used in pole placement, *Int. J. Control*, **32**, 759-770.

Cameron, R.G. (1983). Minimizing the product of two Raleigh quotients, *Linear and Multilinear Algebra*, **13**, 177-178.

Davis, C. and Schneider, H. (1996). Helmut Wielandt: Mathematishe Werke/Mathematical Works - Volume 2: *Linear Algebra and Analysis*, Watter de Gruyter, Berlin, 156-162.

Drury, S.W., Liu, S., Lu, C., Puntanen, S. and Styan, G.H. (2000). Some comments on several matrix inequalities with applications to canonical correlations: Historical background and recent developments, Report A 332, Department of mathematics, Statistics and Philosophy, University of Tampere, Tampere, Finland.

Eaton, M.L. (1976). A maximization problem and its application to canonical correlation, *J. Multivariate Anal.*, **6**, 422-425.

Frucht Werthrimer, Robert. (1943). Sobre algunas desigulddades: Observacion relativa a la solución del Problem No.21, indicada por el Ing. Ernesto M. Saleme (in Spanish), 170 Mathemastical Notoe-Boletin del Instituto de Mathemática (Rosaria) **3**, 41-46.

Greub, W. and Rheinboldt, W. (1959). On a generalization of an inequality of L.V. Kantorovich, *Proc. Math. Soc.*, **10**, 407-415.

Gustafson, K.E. (1968). The angle of an operator and positive operator products, *Bull. AMS*, **74**, 488-492.

Gustafson, K.E. (1999). The geometrical meaning of Kantorovich-Wielandt inequalities, *Linear Algebra and Appl.*, **296**, 143-151.

Gustafson, K.E. (2000a). An extended operator trigonometry, *Linear Algebra Appl.*, **319**, 117-135.

Gustafson, K. (2000b). Quantum trigonometry, infinite dimensional analysis, *Quantum Probability and Related Topics*, **3**, 33-52.

Gustafson, K. (2002). Operator trigonometry of statistics and economics, *Linear Algebra Appl.*, **354**, 141-158.

Gustafson, K.E. and Rao, O.K.M. (1997). *Numerical Range: The Field of Values of Linear Operators and Matrices*, Springer, Berlin.

Kantorovich, L. (1952). *Functional Analysis and Applied Mathematics* (translated from the Russian Edition (1948)). Report No.1509, Washington, D.C.: National Bureau of Standards.

Kato, T. (1980). *Perturbation Theory*, (Second Edition, Chapter 2), Springer Verlag.

Khatri, C.G. (1978). Some optimization problems with applications to canonical correlations and sphericity tests, *J. Multivariate Analysis*, **8**, 453-467. Correction (1982), *J. Multivariate Analysis*, **12**, 612.

Khatri, C.G. and Rao, C.R. (1981). Some extensions of Kantorovich inequality and statistical applications, *J. Multivariate Analysis*, **11**, 499-505.

Khatri, C.G. and Rao, C.R. (1982). Some generalizations of Kantorovich inequality, *Sankhyā* A, **44**, 91-102.

Khattree, R. (2001). On the calculation of antieigenvalues and antiengenvectors, *J. Interdisciplinary Mathematics*, **4**, 195-199.

Khattree, R. (2002). Generalized antieigenvalues of order r, *American J. Math. Management Sciences* (to appear).

Khattree, R. (2003). Antieigenvalues and antieigenvectors in statistics, *J. Stat. Planning and Inference*, **114**, 131-144.

Knott, M. (1975). On the minimum efficiency of least squares, *Biometrika*, **62**, 129-132.

Kouvaritakis, B. and Cameron, R.G. (1980). Pole placement with minimized norm controllers, *Proc. IEEE*, **127**, 32-36.

Krein, M.G. (1969). Angular localization of spectrum of a multiplication integral in Hilbert space, *Func. Anal. and Applic.*, **3**, 89-90.

Rao, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals, In *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol.1, University of California Press, 355-372.

Rao, C.R. (1973). *Linear Statistical Inference and its Applications* (Second Edition), John Wiley, New York.

Rao, C.R. (1985). The inefficiency of least squares, *Linear Algebra and Appl.*, **70**, 249-255.

Rao, C.R. and Rao, C. Veerendra. (1987). Stationary values of two Rayleigh quotients: Homologous canonical correlations, *Sankhyā*, **49** B, 113-125.

Rao, C.R. and Rao, M.B. *Matrix Algebra and its Applications to Statistics and Econometrics*, World Scientific, Singapore.

Shisha, O. and Mond, B. (1967). Bounds on difference of means, In *Inequalities: Proceedings of a Symposium held at Wright-Patterson Air Force Base*, Ohio, August 19-27, 1965, Academic Press, New York, 293-308.

Strang, W.G. (1960). On Kantorovich inequality, *Proc. Amer. Math. Soc.*, **11**, 468.

Styan, G.P.H. (1983). On some inequalities associated with ordinary least squares and the Kantorovich inequality, In *Festschrift for Eino Haikala on his Seventieth birthday*, Univ. of Tampere, 158-166.

Venebles, W. (1976). Some implications of the union intersection principle for tests of sphericity, *J. Multivariate Anal.*, **6**, 185-190.

Wielandt, H. (1953). Inclusion theorems for eigenvalues, In *Simultaneous Linear Equations and the Determination of Eigenvalues* (Eds. L.J. Paige and O. Taussy), National Bureau of Standards Applied Mathematics Series, Vol 29, 75-78.