

Research Article

Several Matrix Euclidean Norm Inequalities Involving Kantorovich Inequality

Litong Wang and Hu Yang

College of Mathematics and Physics, Chongqing University, Chongqing 400030, China

Correspondence should be addressed to Litong Wang, wanglt80@163.com

Received 21 April 2009; Accepted 4 August 2009

Recommended by Andrei Volodin

Kantorovich inequality is a very useful tool to study the inefficiency of the ordinary least-squares estimate with one regressor. When regressors are more than one statisticians have to extend it. Matrix, determinant, and trace versions of it have been presented in the literature. In this paper, we provide matrix Euclidean norm Kantorovich inequalities.

Copyright © 2009 L. Wang and H. Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Suppose that A is an $n \times n$ positive definite matrix and x is an $n \times 1$ real vector, then the well-known Kantorovich inequality can be expressed as

$$(x'Ax)(x'A^{-1}x) \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n} (x'x)^2, \quad (1.1)$$

where $\lambda_1 \geq \dots \geq \lambda_n > 0$ are the eigenvalues of A . It is a very useful tool to study the inefficiency of the ordinary least-squares estimate with one regressor in the linear model. Watson [1] introduced the ratio of the variance of the best linear unbiased estimator to the variance of the ordinary least-squares estimator. Such a lower bound of this ratio was provided by Kantorovich inequality (1.1); see, for example, [2, 3]. When regressors are more than one statisticians have to extend it. Marshall and Olkin [4] were the first to generalize Kantorovich inequality to matrices (see, e.g., [5])

$$X'A^{-1}X \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n} X'X(X'AX)^{-1}X'X, \quad (1.2)$$

where X is an $n \times p$ real matrix. If $X'X = I_p$, then (1.2) becomes

$$X'A^{-1}X \leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n} (X'AX)^{-1}. \quad (1.3)$$

Bloomfield and Watson [6] and Knott [7] simultaneously established the inequality

$$\det(X'AX) \det(X'A^{-1}X) \leq \prod_{i=1}^m \frac{(\lambda_i + \lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}}, \quad (1.4)$$

where X is an $n \times p$ real matrix such that $X'X = I_p$ and $m = \min\{p, n - p\}$. Yang [8] presented its trace version

$$\frac{\operatorname{tr}(X'AX)}{\operatorname{tr}(X'A^{-1}X)^{-1}} \leq \left(\frac{\sum_{i=1}^p (\lambda_i + \lambda_{n-i+1})}{2 \sum_{i=1}^p \sqrt{\lambda_i\lambda_{n-i+1}}} \right)^2, \quad (1.5)$$

where X is an $n \times p$ ($2p \leq n$) real matrix such that $X'X = I_p$.

To the best of our knowledge, there has not been any matrix Euclidean norm version of Kantorovich inequality yet. Our goal is to present its matrix Euclidean norm version.

This paper is arranged as follows. In Section 2, we will give some lemmas which are useful in the following section. In Section 3, some matrix inequalities are established by Kantorovich inequality or Pólya-Szegő inequality, which are referred to as the extensions of Kantorovich inequality as well and conclusions are given in Section 4.

2. Some Lemmas

We will start with some lemmas which are very useful in the following.

Definition 2.1. Let A be an $n \times n$ complex square matrix. A is called a normal matrix if $A^*A = AA^*$.

Lemma 2.2. Let A be an $n \times n$ complex square matrix and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A , then

$$\sum_{i=1}^n |\lambda_i|^2 \leq \|A\|^2, \quad (2.1)$$

where $\|A\|^2 = \operatorname{tr}(A^*A)$ denotes the squared Euclidean norm of A . The equality in (2.1) holds if and only if A is a normal matrix.

Proof. See [5]. □

Lemma 2.3 (Pólya-Szegö inequality). *There is*

$$\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2 \leq \frac{(m_1 m_2 + M_1 M_2)^2}{4m_1 m_2 M_1 M_2} \left(\sum_{i=1}^n a_i b_i \right)^2, \quad (2.2)$$

where $0 < m_1 \leq a_i \leq M_1$, $0 < m_2 \leq b_i \leq M_2$ ($i = 1, \dots, n$).

Moreover Greub and Rheinboldt [9] generalized Pólya-Szegö inequality to matrices.

Lemma 2.4 (Poincaré). *Let A be an $n \times n$ Hermitian matrix, and let U be an $n \times k$ column orthogonal and full rank matrix, that is $U^*U = I_k$, then one has*

$$\lambda_{n-k+i}(A) \leq \lambda_i(U^*AU) \leq \lambda_i(A), \quad i = 1, \dots, k. \quad (2.3)$$

Let $\phi = (\varphi_1, \dots, \varphi_n)$ be a unitary matrix, whose column vectors are eigenvectors corresponding to $\lambda_1(A) \geq \dots \geq \lambda_n(A)$, respectively. Assume $\phi_{(k)} = (\varphi_1, \dots, \varphi_k)$ and $\phi_{[k]} = (\varphi_{n-k+1}, \dots, \varphi_n)$, then $\lambda_i(U^*AU) = \lambda_i(A)$, $i = 1, \dots, k$, if and only if $U = \phi_{(k)}D$; while $\lambda_{n-k+i}(A) = \lambda_i(U^*AU)$, $i = 1, \dots, k$, if and only if $U = \phi_{[k]}D$, where D is a $k \times k$ unitary matrix.

Proof. See [5]. □

3. Main Results

Theorem 3.1. *Let A and B be $n \times n$ nonnegative definite Hermitian matrices with $\text{rank}(A) = \text{rank}(B)$, and let X be an $n \times p$ complex matrix satisfying $X^*X = I_p$. Then one has*

$$\|X^*AX\| \|X^*BX\| \leq \frac{1}{2} \left(\sqrt{\frac{\lambda_1 \mu_1}{\lambda_p \mu_p}} + \sqrt{\frac{\lambda_p \mu_p}{\lambda_1 \mu_1}} \right) \sum_{i=1}^p \lambda_i \mu_{p-i+1}, \quad (3.1)$$

where $\lambda_1 \geq \dots \geq \lambda_g > 0$ ($p \leq g \leq n$) and $\mu_1 \geq \dots \geq \mu_g > 0$ are eigenvalues of matrices A and B , respectively.

Proof. We easily get that X^*AX is a Hermitian matrix since A is a Hermitian matrix. Hence X^*AX is a normal matrix and then we can derive from Lemma 2.2 that

$$\|X^*AX\|^2 = \sum_{i=1}^p \lambda_i^2(X^*AX). \quad (3.2)$$

By Lemma 2.4 we get that

$$\sum_{i=1}^p \lambda_i^2(X^*AX) \leq \sum_{i=1}^p \lambda_i^2(A) = \sum_{i=1}^p \lambda_i^2. \quad (3.3)$$

Similarly,

$$\|X^*BX\|^2 \leq \sum_{i=1}^p \mu_i^2. \quad (3.4)$$

Note that

$$\|X^*AX\|^2 \|X^*BX\|^2 \leq \sum_{i=1}^p \lambda_i^2 \sum_{i=1}^p \mu_i^2. \quad (3.5)$$

The latter expression of (3.5) may be expressed as

$$\sum_{i=1}^p \lambda_i^2 \sum_{i=1}^p \mu_i^2 = \sum_{k=1}^p \lambda_{i_k}^2 \sum_{k=1}^p \mu_{i_k}^2, \quad (3.6)$$

where i_1, i_2, \dots, i_p is an arbitrary permutation of $1, 2, \dots, p$. Clearly, $\lambda_1 = \max_k \{\lambda_{i_k}\}$, $\lambda_p = \min_k \{\lambda_{i_k}\}$, $\mu_1 = \max_k \{\mu_{i_k}\}$ and $\mu_p = \min_k \{\mu_{i_k}\}$. Therefore, let $M_1 = \lambda_1$, $m_1 = \lambda_p$, $M_2 = \mu_1$, and $m_2 = \mu_p$, then we can derive from Pólya-Szegő inequality that

$$\sum_{i=1}^p \lambda_i^2 \sum_{i=1}^p \mu_i^2 = \sum_{k=1}^p \lambda_{i_k}^2 \sum_{k=1}^p \mu_{i_k}^2 \leq \frac{(\lambda_1 \mu_1 + \lambda_p \mu_p)^2}{4\lambda_1 \mu_1 \lambda_p \mu_p} \left(\sum_{k=1}^p \lambda_{i_k} \mu_{i_k} \right)^2. \quad (3.7)$$

Since inequality (3.7) holds for any permutation of $1, \dots, p$, thus we find

$$\sum_{i=1}^p \lambda_i^2 \sum_{i=1}^p \mu_i^2 \leq \frac{(\lambda_1 \mu_1 + \lambda_p \mu_p)^2}{4\lambda_1 \mu_1 \lambda_p \mu_p} \min_{i_k} \left\{ \left(\sum_{k=1}^p \lambda_{i_k} \mu_{i_k} \right)^2 \right\}. \quad (3.8)$$

In the following, the remaining problem is to choose a proper permutation of $1, 2, \dots, p$ to minimize

$$\sum_{k=1}^p \lambda_{i_k} \mu_{i_k}. \quad (3.9)$$

This may be solved by a nontrivial but elementary combinatorial argument, thus we find

$$\min_{i_k} \sum_{k=1}^p \lambda_{i_k} \mu_{i_k} = \sum_{i=1}^p \lambda_i \mu_{p-i+1}. \quad (3.10)$$

Then

$$\begin{aligned} \|X^*AX\| \|X^*BX\| &\leq \left[\frac{(\lambda_1\mu_1 + \lambda_p\mu_p)^2}{4\lambda_1\mu_1\lambda_p\mu_p} \left(\sum_{i=1}^p \lambda_i\mu_{p-i+1} \right)^2 \right]^{1/2} \\ &= \frac{1}{2} \left(\sqrt{\frac{\lambda_1\mu_1}{\lambda_p\mu_p}} + \sqrt{\frac{\lambda_p\mu_p}{\lambda_1\mu_1}} \right) \sum_{i=1}^p \lambda_i\mu_{p-i+1}. \end{aligned} \tag{3.11}$$

□

Remark 3.2. When A is positive definite Hermitian matrix and $B = A^{-1}$, inequality (3.1) plays an important role in the linear model $\{y, X\beta, A\}$. The covariance matrices of the ordinary least-squares estimator and the best linear unbiased estimator are given in this model

$$\begin{aligned} \text{cov}[\text{OLSE}(X\beta)] &= X(X'X)^{-1}X'AX(X'X)^{-1}X', \\ \text{cov}[\text{BLUE}(X\beta)] &= X(X'A^{-1}X)^{-1}X'. \end{aligned} \tag{3.12}$$

Applying inequality (3.1), we can establish a lower bound of the inefficiency of least-squares estimator

$$\frac{\|\text{cov}[\text{BLUE}(X\beta)]\|}{\|\text{cov}[\text{OLSE}(X\beta)]\|} \geq \frac{2p\sqrt{\lambda_1\lambda_p\lambda_{n-p+1}\lambda_n}}{(\lambda_1\lambda_{n-p+1} + \lambda_n\lambda_p) \sum_{i=1}^p (\lambda_i/\lambda_{n-p+i})}. \tag{3.13}$$

See also [10].

In Theorem 3.1, we need the assumption that $X^*X = I_p$. However, we should also point out that the matrix X may not meet such an assumption in practice. Therefore, we relax this assumption in the following but the results are weakened.

Theorem 3.3. *Let A and B be $n \times n$ nonnegative definite Hermitian matrices with $\text{rank}(A) = \text{rank}(B)$, and let X be an $n \times p$ complex matrix, then one has*

$$\|X^*AX\| \|X^*BX\| \leq \frac{1}{2} \left(\sqrt{\frac{\lambda_1\mu_1}{\lambda_g\mu_g}} + \sqrt{\frac{\lambda_g\mu_g}{\lambda_1\mu_1}} \right) \sum_{i=1}^g \lambda_i\mu_{g-i+1} \|X^*X\|^2, \tag{3.14}$$

where $\lambda_1 \geq \dots \geq \lambda_g > 0$ ($g \leq n$) and $\mu_1 \geq \dots \geq \mu_g > 0$ are eigenvalues of matrices A and B , respectively.

Proof. If $X = 0$, the result obviously holds. Next set $X \neq 0$. Let the spectral decomposition of A be $A = Q^* \Lambda Q$, where Q is an orthogonal matrix and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_g, 0, \dots, 0)$. Let $T = QX$, then

$$\begin{aligned} \|TT^*\| &= \|QXX^*Q^*\| = \|XX^*\| = \|X^*X\|, \\ \|X^*AX\| &= \|T^*\Lambda T\| = [\text{tr}(T^*\Lambda TT^*\Lambda T)]^{1/2} = \left[\sum_{i=1}^p \lambda_i (T^*\Lambda TT^*\Lambda T) \right]^{1/2} \\ &= \left[\sum_{i=1}^n \lambda_i (\Lambda TT^*\Lambda TT^*) \right]^{1/2} \leq \|\Lambda TT^*\| \leq \|\Lambda\| \|TT^*\|. \end{aligned} \quad (3.15)$$

We can derive from (3.15) that

$$\|X^*AX\| \leq \|\Lambda\| \|X^*X\|. \quad (3.16)$$

Similarly,

$$\|X^*BX\| \leq \|\Delta\| \|X^*X\|, \quad (3.17)$$

where $\Delta = \text{diag}(\mu_1, \dots, \mu_g, 0, \dots, 0)$. We thus have

$$\frac{\|X^*AX\| \|X^*BX\|}{\|X^*X\|^2} \leq \|\Lambda\| \|\Delta\| = \sum_{i=1}^g \lambda_i^2 \sum_{i=1}^g \mu_i^2. \quad (3.18)$$

According to the proof of Theorem 3.1 we can get that

$$\frac{\|X^*AX\| \|X^*BX\|}{\|X^*X\|^2} \leq \frac{1}{2} \left(\sqrt{\frac{\lambda_1 \mu_1}{\lambda_g \mu_g}} + \sqrt{\frac{\lambda_g \mu_g}{\lambda_1 \mu_1}} \right) \sum_{i=1}^g \lambda_i \mu_{g-i+1} \|X^*X\|^2. \quad (3.19)$$

The proof is completed. \square

Corollary 3.4. Let A be an $n \times n$ positive definite Hermitian matrix with eigenvalues $\lambda_1 \geq \dots \geq \lambda_n > 0$, and let X be an arbitrary $n \times p$ complex matrix, then one has

$$\|X^*AXX^*A^{-1}X\| \leq \frac{n}{2} \left(\frac{\lambda_1}{\lambda_n} + \frac{\lambda_n}{\lambda_1} \right) \|X^*X\|^2. \quad (3.20)$$

Proof. It is very easy to prove therefore we omit the proof. \square

Theorem 3.5. Let A and B be $n \times n$ positive definite Hermitian matrices with $AB = BA$, $\lambda_1 \geq \dots \geq \lambda_n > 0$ and $\mu_1 \geq \dots \geq \mu_n > 0$ be the eigenvalues of A and B , respectively, and let X be an arbitrary $n \times p$ complex matrix. Then,

$$\|X^* A^2 X X^* B^2 X\| \leq \frac{n}{2} \left(\frac{\lambda_1 \mu_1}{\lambda_n \mu_n} + \frac{\lambda_n \mu_n}{\lambda_1 \mu_1} \right) \|X^* A B X\|^2. \quad (3.21)$$

Proof. If $X = 0$, the result obviously holds. Next set $X \neq 0$. Since $AB = BA$, there exists a unitary matrix V such that $A = V \Delta V^*$ and $B = V M V^*$, where $\Delta = \text{diag}(\lambda_1, \dots, \lambda_n)$, $M = \text{diag}(\mu_{i_1}, \dots, \mu_{i_n})$.

Define $Z = (\Delta M)^{1/2} V^* X$, $C = \Delta M^{-1} = \text{diag}(\lambda_1 / \mu_{i_1}, \dots, \lambda_n / \mu_{i_n})$. By Corollary 3.4, we can get that

$$\frac{\|X^* A^2 X X^* B^2 X\|}{\|X^* A B X\|^2} = \frac{\|Z^* C Z Z^* C^{-1} Z\|}{\|Z^* Z\|^2} \leq \frac{n}{2} \left(\frac{\delta_1}{\delta_n} + \frac{\delta_n}{\delta_1} \right), \quad (3.22)$$

where $\delta_1 = \max_k \{\lambda_k / \mu_{i_k}\}$, $\delta_n = \min_k \{\lambda_k / \mu_{i_k}\}$. The right-hand side of (3.22) may be denoted by d , then

$$d = \frac{n}{2} \left(\frac{\delta_1}{\delta_n} + \frac{\delta_n}{\delta_1} \right) = \frac{n}{2} \left(\frac{\delta_1}{\delta_n} + \frac{1}{\delta_1 / \delta_n} \right). \quad (3.23)$$

It is easy to prove that d is a monotone increasing function of δ_1 / δ_n on interval $[1, \infty)$. Write $\alpha_1 = \mu_1 / \lambda_n$, $\alpha_n = \mu_n / \lambda_1$, then we have $\alpha_1 / \alpha_n \geq \delta_1 / \delta_n$. From the definitions of δ_1 and δ_n , we thus have

$$d \leq \frac{n}{2} \left(\frac{\lambda_1 \mu_1}{\lambda_n \mu_n} + \frac{\lambda_n \mu_n}{\lambda_1 \mu_1} \right). \quad (3.24)$$

This completes the proof. \square

Corollary 3.6. If A and B are positive semidefinite Hermitian matrices with eigenvalues $\lambda_1 \geq \dots \geq \lambda_g > 0$ and $\mu_1 \geq \dots \geq \mu_g > 0$, respectively, inequality (3.21) becomes

$$\|X^* A^2 X X^* B^2 X\| \leq \frac{g}{2} \left(\frac{\lambda_1 \mu_1}{\lambda_g \mu_g} + \frac{\lambda_g \mu_g}{\lambda_1 \mu_1} \right) \|X^* A B X\|^2. \quad (3.25)$$

Theorem 3.7. Let A and B be an $n \times n$ positive semidefinite Hermitian matrices with $\text{rank}(A) = \text{rank}(B)$, $\lambda_1 \geq \dots \geq \lambda_g > 0$ and $\mu_1 \geq \dots \geq \mu_g > 0$ be the eigenvalues A and B , respectively, and let X, Y be $n \times p$, $n \times q$ complex matrices with $\text{rank}(X) = \text{rank}(Y)$. Then

$$\|X^* A Y\| \|Y^* B X\| \leq \frac{1}{2} \left(\sqrt{\frac{\lambda_1 \mu_1}{\lambda_g \mu_g}} + \sqrt{\frac{\lambda_g \mu_g}{\lambda_1 \mu_1}} \right) \sum_{i=1}^g \lambda_i \mu_{g-i+1} \|X^* X\| \|Y^* Y\|. \quad (3.26)$$

Proof. Note that

$$\begin{aligned}\|X^*AY\|^2 &= \text{tr}(Y^*AXX^*AY) = \text{tr}(YY^*AXX^*A) \leq \lambda_1(A)\text{tr}(YY^*AXX^*) \\ &= \lambda_1(A)\text{tr}(XX^*YY^*A) \leq \lambda_1^2(A)\text{tr}(XX^*YY^*) \\ &= \lambda_1(A^2)\text{tr}(XX^*YY^*) \leq \text{tr}(A^2)\text{tr}(XX^*YY^*).\end{aligned}\quad (3.27)$$

Similarly,

$$\|Y^*BX\|^2 \leq \text{tr}(B^2)\text{tr}(XX^*YY^*). \quad (3.28)$$

Using the abbreviations $S = XX^*$, $T = YY^*$. Clearly, $S \geq 0$, $T \geq 0$. Let $a_1 \geq \dots \geq a_s$, $b_1 \geq \dots \geq b_s$ be the eigenvalues of S , T , respectively. Applying Hölder inequality, we can derive that

$$\text{tr}(ST) \leq \sum_{i=1}^s a_i b_i \leq \left(\sum_{i=1}^s a_i^2 \right)^{1/2} \left(\sum_{i=1}^s b_i^2 \right)^{1/2} = \left(\text{tr}(S^2) \right)^{1/2} \left(\text{tr}(T^2) \right)^{1/2} = \|S\| \|T\|. \quad (3.29)$$

Thus

$$\frac{\|X^*AY\| \|Y^*BX\|}{\|X^*X\| \|Y^*Y\|} \leq \|A\| \|B\| \leq \frac{1}{2} \left(\sqrt{\frac{\lambda_1 \mu_1}{\lambda_g \mu_g}} + \sqrt{\frac{\lambda_g \mu_g}{\lambda_1 \mu_1}} \right) \sum_{i=1}^g \lambda_i \mu_{g-i+1}. \quad (3.30)$$

This completes the proof. \square

Corollary 3.8. When $B = A^{-1}$, inequality (3.26) becomes

$$\|X^*AY\| \|Y^*A^{-1}X\| \leq \frac{n}{2} \left(\frac{\lambda_1}{\lambda_n} + \frac{\lambda_n}{\lambda_1} \right) \|X^*X\| \|Y^*Y\|. \quad (3.31)$$

4. Conclusions

The study of the inefficiency of the ordinary least-squares estimator in the linear model requires a lower bound for the efficiency defined as the ratio of the variance or covariance of the best linear unbiased estimator to the variance or covariance of the ordinary least-squares estimator. Such a bound can be given by Kantorovich inequality or its extensions. Matrix, determinant, and trace versions of it have been presented in the literature. In this paper, we present its matrix Euclidean norm version.

Acknowledgment

The authors thank very much the associate editors and reviewers for their insightful comments and kind suggestions that lead to improving the presentation.

References

- [1] G. S. Watson, *Serial correlation in regression analysis*, Ph.D. thesis, Department of Experimental Statistics, North Carolina State College, Raleigh, NC, USA, 1951.
- [2] G. S. Watson, G. Alpargu, and G. P. H. Styan, "Some comments on six inequalities associated with the inefficiency of ordinary least squares with one regressor," *Linear Algebra and Its Applications*, vol. 264, pp. 13–54, 1997.
- [3] S. W. Drury, S. Liu, C.-Y. Lu, S. Puntanen, and G. P. H. Styan, "Some comments on several matrix inequalities with applications to canonical correlations: historical background and recent developments," *Sankhya*, vol. 64, no. 2, pp. 453–507, 2002.
- [4] A. W. Marshall and I. Olkin, "Multivariate distributions generated from mixtures of convolution and product families," in *Topics in Statistical Dependence (Somerset, PA, 1987)*, vol. 16 of *Ims Lecture Notes-Monograph Series*, pp. 371–393, Institute of Mathematical Statistics, Hayward, Calif, USA, 1990.
- [5] S. G. Wang and Z. Z. Jia, *Matrix Inequalities*, Science Press, Beijing, China, 1994.
- [6] P. Bloomfield and G. S. Watson, "The inefficiency of least squares," *Biometrika*, vol. 62, pp. 121–128, 1975.
- [7] M. Knott, "On the minimum efficiency of least squares," *Biometrika*, vol. 62, pp. 129–132, 1975.
- [8] H. Yang, "Extensions of the Kantorovich inequality and the error ratio efficiency of the mean square," *Mathematica Applicata*, vol. 1, no. 4, pp. 85–90, 1988 (Chinese).
- [9] W. Greub and W. Rheinboldt, "On a generalization of an inequality of L. V. Kantorovich," *Proceedings of the American Mathematical Society*, vol. 10, pp. 407–415, 1959.
- [10] H. Yang and L. Wang, "An alternative form of the Watson efficiency," *Journal of Statistical Planning and Inference*, vol. 139, no. 8, pp. 2767–2774, 2009.