RESEARCH **Open Access**

The efficiency comparisons between OLSE and BLUE in a singular linear model

Litong Wang[1*] and Guobing Pan[2]

*Correspondence:
wanglt80@163.com
[1]Department of Applied
Mathematics, Zhejiang University of
Technology, Hangzhou, China
Full list of author information is
available at the end of the article

Abstract

This paper is mainly concerned with the efficiency comparison between OLSE and BLUE in a singular linear model. We define the efficiencies between OLSE and BLUE by means of the matrix Euclidean norm and prove a matrix Euclidean norm version of the Kantorovich inequality to limit upper or lower bounds of these efficiencies. It relaxes the assumptions that the covariance matrix is positive definite and the design matrix has full column rank.

MSC: Primary 62J05; secondary 62H05; 62H20

Keywords: Euclidean norm; least squares; singular linear model; Watson efficiency

1 Introduction

Inequalities are studied and utilized widely in many fields such as in matrix theory, statistics and so on. In statistics, they are often used to make efficiency comparisons between two estimators. For example, Wang and Shao [1] have discussed the efficiency comparisons between the ordinary least squares estimator (OLSE) and the best linear unbiased estimator (BLUE) in linear models. In this paper, our goal is to make the comparison of efficiencies between OLSE and BLUE in a singular linear model by using matrix norm versions of the Kantorovich inequality involving a nonnegative definite matrix.

Consider the following linear regression model:

$$y = X\beta + \varepsilon, \tag{1.1}$$

where $y \in R^n$ is the vector of n observations, $X \in R^{n \times p}$ is the known design matrix, $\beta \in R^P$ is the unknown vector of regression coefficients and $\varepsilon \in R^n$ is the error vector with mean vector zero and the covariance matrix Σ.

When X has full column rank and Σ is assumed to be positive definite, it is well known that the best linear unbiased estimator (BLUE) of β can be expressed as

$$\widetilde{\beta} = \left(X'\Sigma^{-1}X\right)^{-1}X'\Sigma^{-1}y \tag{1.2}$$

and the ordinary least squares estimator (OLSE) of β is given by

$$\widehat{\beta} = \left(X'X\right)^{-1}X'y. \tag{1.3}$$

According to the Löwner ordering, we can easily compute from (1.2) and (1.3) that $\text{cov}(\widehat{\beta}) - \text{cov}(\widetilde{\beta}) \geq 0$, which is nonnegative definite. Since there is no unique way to measure how

'bad' the OLSE can be with respect to the BLUE, various criteria have been considered in the literature; see, *e.g.*, [2–6]. Among these criteria, the frequently used measure is the Watson efficiency [2] defined as follows:

$$\phi_1 = \frac{|\operatorname{cov}(\widetilde{\beta})|}{|\operatorname{cov}(\widehat{\beta})|} = \frac{|X'X|^2}{|X'\Sigma X| \cdot |X'\Sigma^{-1}X|}, \tag{1.4}$$

where $|\cdot|$ indicates the determinant of the matrix concerned. The lower bound is provided by the Bloomfild-Watson-Knott inequality; see, *e.g.*, [7, 8]. However, Yang and Wang [9] have shown that such a criterion is not always so satisfactory and provided an alternative form defined as the ratio of the Euclidean norms (or Frobenius norms) of the corresponding covariance matrices:

$$\phi_2 = \frac{\|\operatorname{cov}(\widetilde{\beta})\|}{\|\operatorname{cov}(\widehat{\beta})\|} = \frac{\|(X'\Sigma^{-1}X)^{-1}\|}{\|(X'X)^{-1}X'\Sigma X(X'X)^{-1}\|}. \tag{1.5}$$

Many authors assume that the covariance matrix is nonsingular in their analysis of this classic linear model. But the number of characteristics that could be included in the model may be clearly limited by this assumption of nonsingularity. A few authors relax the condition of nonsingularity and consider a singular linear model. For example, Liski *et al.* [10] and Liu [4] make efficiency comparisons between the OLSE and BLUE in a singular linear model. In the present paper, the singular linear model is further studied.

The Watson efficiency ϕ_1 has been generalized to a weakly singular model; see, *e.g.*, [11]. For a general case of the underlying singular linear model, it is not interesting because the denominator reduces to zero. In order to relax assumptions on the rank of X and Σ, we mainly discuss its alternative form based on the Euclidean norm [9, 12].

We hereinafter introduce some useful notations. Let the symbols A', A^-, A^+, $\mathscr{R}(A)$, $\mathscr{R}(A)^\perp$ and $\operatorname{rk}(A)$ stand for the transpose, a generalized inverse, the Moore-Penrose inverse, the column space, the orthogonal complement of the column space and the rank of the matrix A, respectively. Moreover, write $P_A = AA^+ = A(A'A)^+A'$ and $M_A = I - P_A$, in particular, $H = P_X$, $M = I - H$. $\lambda_i(A)$ denotes the ith largest eigenvalue of the matrix A.

2 A new Kantorovich-type inequality

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

Lemma 2.1 *Let A be an $n \times n$ complex matrix and $\lambda_1, \ldots, \lambda_n$ be eigenvalues of A. Then we have*

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq \|A\|^2$$

and the equality holds if and only if A is a regular matrix.

Proof The proof is very easy, we therefore omit it here. □

Lemma 2.2 *Let A be an $n \times n$ positive semidefinite Hermitian matrix and U be an orthogonal projection matrix with $\operatorname{rk}(U) = k$. Then we have*

$$\lambda_{n-k+i}(A) \leq \lambda_i(AU) \leq \lambda_i(A), \quad i = 1, \ldots, k.$$

The proof can be found in [13]. See also [14].

Lemma 2.3 *The Pólya and Szegö inequality*

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

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

Lemma 2.4 *Let A and B be two $n \times n$ positive semidefinite Hermitian matrices, and U be an $n \times k$ matrix, $\mathcal{R}(A) \subset \mathcal{R}(B)$, $\mathrm{rk}(B) = q$, $\mathrm{rk}(BU) = t$, then*

$$\lambda_{q-t+i}(B^- A) \leq \lambda_i\left((U^* BU)^- U^* AU\right) \leq \lambda_i(B^- A), \quad i = 1, \ldots, t.$$

The proof can be found in [15]. See also [14].

Theorem 2.5 *Let A be an $n \times n$ positive semidefinite Hermitian matrix and $\lambda_1 \geq \cdots \geq \lambda_s > 0$ ($s \leq n$) be the ordered eigenvalues of A, and let U be an $n \times p$ complex matrix such that $U^* U = I_p$. If $p \leq s$, we then have*

$$\frac{\|(U' A^+ U)^{-1}\|}{\|U' AU\|} \geq \frac{2p\sqrt{\lambda_1 \lambda_p \lambda_{s-p+1} \lambda_s}}{(\lambda_1 \lambda_{s-p+1} + \lambda_s \lambda_p)\sum_{i=1}^{p} \frac{\lambda_i}{\lambda_{s-p+i}}}.$$

Proof The proof is similar to Theorem 1 in [9], we therefore omit it here. □

3 The comparison of efficiencies

The Watson efficiency [2, 16] and its decompositions [11] are usually used to measure the efficiency of the ordinary least squares. However, Yang and Wang [9] show that such a criterion does not always work well in some cases and propose an alternative form

$$\rho = \frac{\|\mathrm{cov}(X\widetilde{\beta})\|}{\|\mathrm{cov}(X\widehat{\beta})\|} = \frac{\|X(X'\Sigma^{-1}X)^{-1}X'\|}{\|X(X'X)^{-1}X'\Sigma X(X'X)^{-1}X'\|}. \tag{3.1}$$

The above formula and its lower bound both require the covariance matrix Σ to be positive definite and the design matrix to have full column rank. This assumption limits clearly the number of characters which may be included in the model. We here generalize this formula to the situation where the matrices X and Σ can be of arbitrary rank.

In the following, we divide singular linear models into three categories in accordance with the assumptions on X and Σ. These categories are as follows:

(1) $\mathcal{R}(X) \subset \mathcal{R}(\Sigma)$, $\mathrm{rk}(X) = p$, Σ is possibly singular;

(2) $\mathcal{R}(X) \subset \mathcal{R}(\Sigma)$, $\mathrm{rk}(X) < p$, Σ is possibly singular;

(3) Σ is possibly singular.

From now on, we always assume $\mathrm{rk}(\Sigma) = s$ ($s < n$). Then any given singular linear model can be uniquely classified into i ($i = 1, 2, 3$). Many authors have contributed to the theory

in the literature; see, *e.g.*, [17, 18]. The general representations for the BLUE of $X\beta$ and their covariance matrices can be given respectively by

$$X\widetilde{\beta} = Hy - H\Sigma M(M\Sigma M)^- My = X(X'W^-X)^- X'W^- y, \tag{3.2}$$

$$\text{cov}(X\widetilde{\beta}) = H\Sigma H - H\Sigma M(M\Sigma M)^- M\Sigma H = X(X'W^-X)^- X' - XUX', \tag{3.3}$$

where $W = \Sigma + XUX'$ and here $U \geq 0$ is an arbitrary matrix such that $\mathcal{R}(W) = \mathcal{R}(X : \Sigma)$. In particular, Σ^- can play the same role as Σ^{-1} does when Σ is nonsingular as long as $\mathcal{R}(X) \subset \mathcal{R}(\Sigma)$. That is to say, W^- can be replaced by Σ^- in this case. We then have

$$\text{cov}(X\widetilde{\beta}) = X(X'\Sigma^-X)^- X'. \tag{3.4}$$

The covariance matrix of the well-known OLSE of $X\beta$ is given by

$$\text{cov}(X\widehat{\beta}) = X(X'X)^- X'\Sigma X(X'X)^- X' = H\Sigma H. \tag{3.5}$$

In the following, we make efficiency comparisons between the OLSE and BLUE in a singular model according to the above category.

Firstly, we will discuss the category (1). The matrix product $X'\Sigma^-X$ is invariant for all the choices of generalized inverse Σ^- because of the column space inclusion $\mathcal{R}(X) \subset \mathcal{R}(\Sigma)$. Applying the rank rule of the matrix product [19], we can get that

$$\text{rk}(X'\Sigma X) = \text{rk}(\Sigma X) = \text{rk}(X) - \dim \mathcal{R}(X) \cap \mathcal{R}(\Sigma)^{\perp} = p. \tag{3.6}$$

Note that $\mathcal{R}(\Sigma) = \mathcal{R}(\Sigma^+)$, and then we can conclude that $X'\Sigma X$ and $X'\Sigma^+X$ are both nonsingular. In the literature, such a model is often regarded as a weakly singular model or the Zyskind-Martin model [20]. In this model, the relative efficiency ρ becomes

$$\rho_1 = \frac{\|\text{cov}(X\widetilde{\beta})\|}{\|\text{cov}(X\widehat{\beta})\|} = \frac{\|X(X'\Sigma^+X)^{-1}X'\|}{\|X(X'X)^{-1}X'\Sigma X(X'X)^{-1}X'\|}. \tag{3.7}$$

It is easy to prove that $\rho_1 \leq 1$. The following theorem gives its lower bound.

Theorem 3.1 *In the linear regression model* (1.1), *let* $\lambda_1 \geq \cdots \geq \lambda_s$ ($s < n$) *be the ordered eigenvalues of* Σ *and* X *be an* $n \times p$ *design matrix with* $\text{rk}(X) = p$, $\mathcal{R}(X) \subset \mathcal{R}(\Sigma)$. *Then we have*

$$\rho_1 \geq \frac{2p\sqrt{\lambda_1\lambda_p\lambda_{s-p+1}\lambda_s}}{(\lambda_1\lambda_{s-p+1} + \lambda_s\lambda_p) \sum_{i=1}^p \frac{\lambda_i}{\lambda_{s-p+i}}}.$$

Proof We can firstly compute that

$$\rho_1^2 = \frac{\text{tr}[(X'X)^{-\frac{1}{2}}X'\Sigma^+X(X'X)^{-1}X'\Sigma^+X(X'X)^{-\frac{1}{2}}]^{-1}}{\text{tr}(X'X)^{-\frac{1}{2}}X'\Sigma X(X'X)^{-1}X'\Sigma X(X'X)^{-\frac{1}{2}}}.$$

There exists some orthogonal matrix P such that $\Sigma = P\Lambda P'$, so $\Sigma^+ = P\Lambda^+ P'$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_s, 0, \ldots, 0)$. Let $U = P'X(X'X)^{-\frac{1}{2}}$, and then we have that $U'U = I_p$ and

$$\rho_1^2 = \frac{\mathrm{tr}(U'\Lambda^+ UU'\Lambda^+ U)^{-1}}{\mathrm{tr}\,U'\Lambda UU'\Lambda U} = \frac{\|(U'\Lambda^+ U)^{-1}\|^2}{\|U'\Lambda U\|^2}.$$

Using Theorem 2.5, the result in Theorem 3.1 can be established. □

Secondly, we will consider the category (2). Let $\mathrm{rk}(X) = r$ $(r < p)$. Using equation (3.6), we can get that $\mathrm{rk}(X'\Sigma X) = \mathrm{rk}(X'\Sigma^+ X) = r$. Analogically, the relative efficiency ρ becomes

$$\rho_2 = \frac{\|\mathrm{cov}(X\widetilde{\beta})\|}{\|\mathrm{cov}(X\widehat{\beta})\|} = \frac{\|X(X'\Sigma^+ X)^+ X'\|}{\|X(X'X)^+ X'\Sigma X(X'X)^+ X'\|}. \tag{3.8}$$

It is easy to prove that $\rho_2 \leq 1$. The following theorem gives its lower bound.

Theorem 3.2 *In the linear regression model* (1.1), *let* $\lambda_1 \geq \cdots \geq \lambda_s$ $(s < n)$ *be the ordered eigenvalues of* Σ *and* X *be an* $n \times p$ *design matrix with* $\mathrm{rk}(X) = r$ $(r < p)$, $\mathcal{R}(X) \subset \mathcal{R}(\Sigma)$. *We then have*

$$\rho_2 \geq \frac{2r\sqrt{\lambda_1 \lambda_r \lambda_{s-r+1} \lambda_s}}{(\lambda_1 \lambda_{s-r+1} + \lambda_s \lambda_r) \sum_{i=1}^{r} \frac{\lambda_i}{\lambda_{s-r+i}}}.$$

Proof It is easy to prove that

$$X(X'\Sigma X)^+ X' = (H\Sigma^+ H)^+.$$

Then the proof is similar to Theorem 3.1, therefore we omit it here. □

Finally, we take into account the category (3). Owing to (3.3), we may write

$$\mathrm{cov}(X\widetilde{\beta}) = \mathrm{cov}(X\widehat{\beta}) - H\Sigma M(M\Sigma M)^- M\Sigma H.$$

Therefore, we define

$$\rho_3 = \frac{\|\mathrm{cov}(X\widehat{\beta}) - \mathrm{cov}(X\widetilde{\beta})\|}{\|\mathrm{cov}(X\widehat{\beta})\|} = \frac{\|H\Sigma M(M\Sigma M)^- M\Sigma H\|}{\|H\Sigma H\|}. \tag{3.9}$$

Let $\dim \mathcal{R}(X) \cap \mathcal{R}(\Sigma) = g$ $(g \geq 0)$ and $\mathrm{rk}(X) = r$ $(r \leq p)$. Note that

$$\begin{aligned}
\mathrm{rk}(X'\Sigma X) &= \mathrm{rk}(X'X(X'X)^+ X'\Sigma X(X'X)^+ X'X) \\
&\leq \mathrm{rk}(X(X'X)^+ X'\Sigma X(X'X)^+ X') \leq \mathrm{rk}(X'\Sigma X).
\end{aligned}$$

Due to equation (3.6), we have

$$\mathrm{rk}(H\Sigma H) = \mathrm{rk}(X) - \dim \mathcal{R}(X) \cap \mathcal{R}(\Sigma)^\perp. \tag{3.10}$$

As a result from

$$\mathcal{R}(X) = \mathcal{R}(X) \cap \big(\mathcal{R}(\Sigma) \oplus \mathcal{R}(\Sigma)^{\perp}\big) = \big(\mathcal{R}(X) \cap \mathcal{R}(\Sigma)\big) \oplus \big(\mathcal{R}(X) \cap \mathcal{R}(\Sigma)^{\perp}\big),$$

we then have

$$\mathrm{rk}(H\Sigma H) = r - (r - g) = g. \tag{3.11}$$

Similarly, we can obtain that

$$\mathrm{rk}(M\Sigma M) = \mathrm{rk}(\Sigma M) = \mathrm{rk}(M) - \dim \mathcal{R}(M) \cap \mathcal{R}(\Sigma)^{\perp} = \dim \mathcal{R}(M) \cap \mathcal{R}(\Sigma). \tag{3.12}$$

In view of $\mathcal{R}(M) = \mathcal{R}(X)^{\perp}$ and

$$\mathcal{R}(\Sigma) = \mathcal{R}(\Sigma) \cap \big(\mathcal{R}(X) \oplus \mathcal{R}(X)^{\perp}\big) = \big(\mathcal{R}(\Sigma) \cap \mathcal{R}(X)\big) \oplus \big(\mathcal{R}(\Sigma) \cap \mathcal{R}(X)^{\perp}\big),$$

we can get that

$$\mathrm{rk}(M\Sigma M) = \mathrm{rk}(\Sigma M) = s - g. \tag{3.13}$$

Theorem 3.3 *In the linear regression model* (1.1), *let* $\lambda_1 \geq \cdots \geq \lambda_s$ $(s < n)$ *be the ordered eigenvalues of* Σ *and* X *be an* $n \times p$ *design matrix with* $\mathrm{rk}(X) = r$ $(r \leq p)$, $\dim \mathcal{R}(X) \cap \mathcal{R}(\Sigma) = g$, $\mathrm{rk}(H\Sigma M(M\Sigma M)^- M\Sigma H) = h$, *we then have*

$$\rho_3 \leq \begin{cases} \frac{1}{2(s+r-n)}\big(\sqrt{\frac{\lambda_1\lambda_{n-r+1}}{\lambda_s\lambda_{s+r-n}}} + \sqrt{\frac{\lambda_s\lambda_{s+r-n}}{\lambda_1\lambda_{n-r+1}}}\big)\big(\sum_{i=1}^{s+r-n} \frac{\lambda_i}{\lambda_{n-r+i}}\big), & \text{if } h \leq s+r-n, \\ \frac{1}{2(s+r-n)}\big(\sqrt{\frac{\lambda_1\lambda_{s-h+1}}{\lambda_s\lambda_h}} + \sqrt{\frac{\lambda_s\lambda_h}{\lambda_1\lambda_{s-h+1}}}\big)\big(\sum_{i=1}^{h} \frac{\lambda_i}{\lambda_{s-h+i}}\big), & \text{if } h > s+r-n. \end{cases}$$

Proof For convenience, let $a = \|H\Sigma M(M\Sigma M)^- M\Sigma H\|$ and $b = \|H\Sigma H\|$. Then $H\Sigma M \times (M\Sigma M)^- M\Sigma H$ is invariant for all the choices of generalized inverses $(M\Sigma M)^-$. From Lemma 2.1, we can easily get that

$$a^2 = \sum_{i=1}^{h} \lambda_i^2\big(H\Sigma M(M\Sigma M)^- M\Sigma H\big) = \sum_{i=1}^{h} \lambda_i^2\big((M\Sigma M)^- M\Sigma H\Sigma M\big).$$

Obviously, $h = \mathrm{rk}(H\Sigma M(M\Sigma M)^- M\Sigma H) \leq \mathrm{rk}(\Sigma H) = g$ and $h \leq \mathrm{rk}(\Sigma M) = s - g$. Since Σ and $\Sigma H\Sigma$ are positive semidefinite matrices and $\mathcal{R}(\Sigma H\Sigma) \subset \mathcal{R}(\Sigma)$, we can derive from Lemma 2.4 that

$$a^2 \leq \sum_{i=1}^{h} \lambda_i^2\big(\Sigma^- \Sigma H\Sigma\big) = \sum_{i=1}^{h} \lambda_i^2\big(\Sigma \Sigma^- \Sigma H\big) = \sum_{i=1}^{h} \lambda_i^2(\Sigma H).$$

Here H is an orthogonal projection matrix, and then we obtain from Lemma 2.2 that

$$a^2 \leq \sum_{i=1}^{h} \lambda_i^2(\Sigma) \leq \sum_{i=1}^{h} \lambda_i^2.$$

Furthermore, since $H\Sigma H$ is a Hermitian matrix, by Lemma 2.1, we can get that

$$b^2 = \sum_{i=1}^{g} \lambda_i^2(H\Sigma H) = \sum_{i=1}^{g} \lambda_i^2(\Sigma H).$$

The Sylvester theorem (see, *e.g.*, [14]) shows that $n - r + g > s$. Analogically, we can get that

$$b^2 \geq \sum_{i=1}^{s+r-n} \lambda_{n-r+i}^2.$$

Applying the well-known arithmetic-harmonic mean inequality, we have

$$\frac{1}{b^2} \leq \frac{1}{(s+r-n)^2} \sum_{i=1}^{s+r-n} \lambda_{n-r+i}^{-2}.$$

Firstly, we suppose that $h \leq s + r - n$. We can then compute that

$$\rho_3^2 = \frac{a^2}{b^2} \leq \frac{1}{(s+r-n)^2} \sum_{i=1}^{s+r-n} \lambda_i^2 \sum_{i=1}^{s+r-n} \lambda_{n-r+i}^{-2}.$$

By the Pólya and Szegö inequality and a nontrivial but elementary combinational argument, we can establish the first inequality. In fact, the second inequality is similar. □

4 Conclusions

In this article, we use several new matrix norm versions of the Kantorovich inequality involving a nonnegative definite matrix to make the comparison of efficiencies between OLSE and BLUE in a singular linear model. The singular linear model is divided into three categories in accordance with the assumptions on the ranks of X and Σ. We introduce some new relative efficiency criteria and their lower or upper bounds are given based on matrix norm inequalities in Theorem 3.1, Theorem 3.2 and Theorem 3.3.

Author details
^{1}Department of Applied Mathematics, Zhejiang University of Technology, Hangzhou, China. ^{2}College of Mechanical Electrical Engineering, Zhejiang University of Technology, Hangzhou, China.

References
1. Wang, SG, Shao, J: Constrained Kantorovich inequality and relative efficiency of least squares. J. Multivar. Anal. **42**, 284-298 (1992)
2. Watson, GS: Serial correlation in regression analysis. PhD thesis, Department of Experimental Statistics, North Carolina State College, Raleigh (1951)
3. Khatri, CG, Rao, CR: Some generalizations of the Kantorovich inequality. Sankhya **44**, 91-102 (1982)

4. Liu, SZ: Efficiency comparisons between the OLSE and the BLUE in a singular linear model. J. Stat. Plan. Inference **84**, 191-200 (2000)
5. Rao, CR: The inefficiency of least squares: extension of the Kantorovich inequality. Linear Algebra Appl. **70**, 249-255 (1985)
6. Yang, H: Extensions of the Kantorovich inequality and the error ratio efficiency of the mean square. Math. Appl. **4**, 85-90 (1988)
7. Bloomfield, P, Watson, GS: The inefficiency of least squares. Biometrika **62**, 121-128 (1975)
8. Knott, M: On the minimum efficiency of the least square. Biometrika **62**, 129-132 (1975)
9. Yang, H, Wang, LT: An alternative form of the Watson efficiency. J. Stat. Plan. Inference **139**, 2767-2774 (2009)
10. Liski, EP, Puntanen, S, Wang, SG: Bounds for the trace of the difference of the covariance matrices of the OLSE and BLUE. Linear Algebra Appl. **176**, 121-130 (1992)
11. Chu, K, Isotalo, J, Puntanen, S, Styan, GPH: On decomposing the Watson efficiency of ordinary least squares in a partitioned weakly singular linear model. Sankhya **66**, 634-651 (2004)
12. Wang, LT, Yang, H: Several matrix Euclidean norm inequalities involving Kantorovich inequality. J. Inequal. Appl. **2009**, Article ID 291984 (2009)
13. Poincare, H: Sur les equation aux derives partielles de la physique mathematique. Am. J. Math. **12**, 211-294 (1890)
14. Wang, SG, Jia, ZZ: Matrix Inequality. Anhui Education Press, Hefei (1994)
15. Scott, AJ, Styan, GPH: On a separation theorem for generalized eigenvalues and a problem in the analysis of sample surveys. Linear Algebra Appl. **70**, 209-224 (1985)
16. Chu, K, Isotalo, J, Puntanen, S, Styan, GPH: The efficiency factorization multiplier for the Watson efficiency in partitioned linear models: some examples and a literature review. J. Stat. Plan. Inference **137**, 3336-3351 (2007)
17. Christensen, R: Plane Answers to Complex Questions: The Theory of Linear Models. Springer, New York (1987)
18. Groß, J: The general Gauss-Markov model with possibly singular dispersion matrix. Stat. Pap. **45**, 311-336 (2004)
19. Meyer, CD: Matrix Analysis and Applied Linear Algebra. Cambridge University Press, Cambridge (2001)
20. Zyskind, G, Martin, FB: On best linear estimation and a general Gauss-Markov theorem in linear models with arbitrary nonnegative covariance structure. SIAM J. Appl. Math. **17**, 1190-1202 (1969)