Brazilian Journal of Probability and Statistics
2012, Vol. 26, No. 2, 113–122
DOI: 10.1214/10-BJPS123

A simple formula for asymptotic distributional risk of some estimators

Sévérien Nkurunziza

University of Windsor

Abstract. In this article, we are interested in deriving the asymptotic distributional risk function of a class of estimator concerning the mean parameter matrix of matrices variate random sample. The proposed result is useful in decision theory, more precisely in risk analysis of a class of some robust estimators such as Stein-rule types estimators.

1 Introduction

In this article, we are interested in establishing the asymptotic distributional risk function of a class of estimators concerning the common mean parameter matrix $\boldsymbol{\theta}$ of n identically distributed random variate matrices of order $q \times k$. Namely, let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ be ergodic and strictly stationary random matrices with mean $\boldsymbol{\theta}$ and covariance–variance $\boldsymbol{\Omega} \otimes \boldsymbol{\Lambda}$ where $\boldsymbol{\Omega}$, $\boldsymbol{\Lambda}$ are respectively $k \times k$ and $q \times q$ positive definite matrices. Here, $A \otimes B$ stands for the Kronecker product of the matrices A and B. In the sequel, and for the sake of simplicity, we assume that the random matrices $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are mutually independent and identically distributed. However, the proposed result holds for an ergodic and strictly stationary process that satisfies the conditions for the central limit theorem for stationary processes (see, e.g., Eagleson, 1975, and references therein).

As mentioned above, the main contribution is to establish the asymptotic distributional risk function of a class of estimators of $\boldsymbol{\theta}$. Also, as intermediate step, we extend Theorem 1 and Theorem 2 in Bock and Judge (1978). Briefly, these theorems are useful in evaluating the asymptotic efficiency of a large class of Shrinkage-type estimators. To set up notation, let $\mathbf{A}$ be a matrix and let $\|\mathbf{A}\|^2_{\Xi_1,\Xi_2} = \operatorname{trace}(\mathbf{A}'\Xi_1\mathbf{A}\Xi_2)$ with Ξ_1 a known nonnegative definite matrix, and Ξ_2 a known positive definite matrix. Further, let h be known Borel measurable and real-valued integrable function, let $\mathbf{F}$ and $\mathbf{L}$ be respectively $p \times q$ and $k \times m$-known matrix full rank with $p < q$ and $p \leq m \leq k$, and let $\mathbf{d}$ be a $p \times m$-known matrix. We consider the following class of estimator

$$\widehat{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}} + h\bigl(\|(\overline{\mathbf{X}} - \widetilde{\boldsymbol{\theta}})\mathbf{L}\|^2_{\Xi_1,\Xi_2}\bigr)(\overline{\mathbf{X}} - \widetilde{\boldsymbol{\theta}}), \tag{1.1}$$

Key words and phrases. Estimator, matrix parameter, matrix variate normal, risk function.
Received October 2009; accepted April 2010.

with $\overline{\mathbf{X}}$ the sample mean, $\Xi_1 = \mathbf{F}'(\mathbf{F}\widehat{\boldsymbol{\Lambda}}\mathbf{F}')^{-1}\mathbf{F}$, $\Xi_2 = (\mathbf{L}'\widehat{\boldsymbol{\Omega}}\mathbf{L})^{-1}$ where $\widehat{\boldsymbol{\Lambda}}$ and $\widehat{\boldsymbol{\Omega}}$ are strongly consistent estimators of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$ respectively. For example, one can take

$$\widehat{\boldsymbol{\Lambda}} = n^{-1}\sum_{i=1}^{n}(\mathbf{X}-\overline{\mathbf{X}})(\mathbf{X}-\overline{\mathbf{X}})', \quad \text{and} \quad \widehat{\boldsymbol{\Omega}} = n^{-1}\sum_{i=1}^{n}(\mathbf{X}-\overline{\mathbf{X}})'(\mathbf{X}-\overline{\mathbf{X}}).$$

Further,

$$\widetilde{\boldsymbol{\theta}} = \overline{\mathbf{X}} - \widehat{\boldsymbol{\Lambda}}\mathbf{F}'(\mathbf{F}\widehat{\boldsymbol{\Lambda}}\mathbf{F}')^{-1}(\mathbf{F}\overline{\mathbf{X}}\mathbf{L}-\mathbf{d})(\mathbf{L}'\widehat{\boldsymbol{\Omega}}\mathbf{L})^{-1}\mathbf{L}'\widehat{\boldsymbol{\Omega}}. \tag{1.2}$$

It is noticed that $\overline{\mathbf{X}}$ is the unrestricted least-square estimator (LSE) of $\boldsymbol{\theta}$ while $\widetilde{\boldsymbol{\theta}}$ corresponds to the restricted LSE of $\boldsymbol{\theta}$ with respect to the restriction

$$\mathbf{F}\boldsymbol{\theta}\mathbf{L} = \mathbf{d}, \tag{1.3}$$

with $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$ replaced by their corresponding strongly consistent estimators. The restriction in (1.3) is useful for example in variable selection and model assessment as well as in profile analysis. For more details, we refer to Izenman (2008, Chapter 6).

Further, $\widehat{\boldsymbol{\theta}}$ can be seen as a shrinkage-pretest type estimator of $\boldsymbol{\theta}$ that combines the sample information and the nonsample information from uncertain restriction in (1.3). In case where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are p-column random vectors, it has been shown that if $p \geq 3$, for some h, the estimator $\widehat{\boldsymbol{\theta}}$ outperforms over the unrestricted LSE $\overline{\mathbf{X}}$. In such an investigation, the derivation of the asymptotic distributional risk function of the estimator $\widehat{\boldsymbol{\theta}}$ is required, but this is not straightforward for the case as studied here where $\widehat{\boldsymbol{\theta}}$ is a $q \times k$-matrix estimator.

In this paper, we consider the loss function $l(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}; \mathbf{W}) = n\,\mathrm{trace}[\mathbf{L}'(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\mathbf{W}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})\mathbf{L}]$, for a nonnegative definite matrix $\mathbf{W}$, and thus, we consider to compute risk function of $\widehat{\boldsymbol{\theta}}$,

$$\mathrm{R}_n(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = n\mathrm{E}\big[\mathrm{trace}\big(\mathbf{L}'(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\mathbf{W}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})\mathbf{L}\big)\big].$$

Since we do not assume any specific distribution population, $\mathrm{R}_n(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta})$ cannot be explicitly determined. Thus, we consider a sequence of local alternative restrictions

$$\mathbf{F}\boldsymbol{\theta}\mathbf{L} = \mathbf{d} + \delta/\sqrt{n}, \qquad n = 1, 2, 3, \ldots \tag{1.4}$$

with $\boldsymbol{\delta}$ a nonzero $p \times m$-matrix. Also, we assume that $\|\boldsymbol{\delta}\| < \infty$. Further, we consider to derive the asymptotic distributional risk (ADR) of $\widehat{\boldsymbol{\theta}}$ as introduced by Ahmed (2001), Ahmed and Saleh (1999) among others. As in the quoted articles, (ADR) of $\widehat{\boldsymbol{\theta}}$ is defined as $\mathrm{ADR}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathrm{E}[\mathrm{trace}(\boldsymbol{\rho}_0'\mathbf{W}\boldsymbol{\rho}_0)]$, with $\boldsymbol{\rho}_0$ the limit in distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})\mathbf{L}$. Under some regularities conditions, $\boldsymbol{\rho}_0$ is also the limit in mean-square sense and then, we have $\lim_{n\to\infty}\mathrm{R}_n(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathrm{ADR}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta})$. It should be noticed that such regularities conditions hold for example if the random sample is from multivariate normal population.

The reminder of this paper is organized as follows. Section 2 presents a proposition and two theorems which are used in deriving the asymptotic distributional risk function. Section 3 gives the main contribution of the paper, that is, the asymptotic distributional risk function formula. Finally, Section 4 is the Conclusion, and for the convenience of the reader, technical details are given in the Appendix.

2 Preliminary results

In this section, we present some intermediate results which are useful in deriving the asymptotic distributional risk function, that is the main contribution of the paper. To set up some notation, note that for a $m \times p$-matrix $\mathbf{A}$, one can write $\mathbf{A} = (A_1, A_2, \dots, A_p)$, $A_j \in \mathbb{R}^p$, $j = 1, 2, \dots, p$, where $\mathbb{R}^p$ denotes the p dimensional real space. Further, let $\text{Vec}(\mathbf{A})$ denote the np column vector obtained by stacking together the columns of $\mathbf{A}$ one underneath the other, that is, $\text{Vec}(\mathbf{A}) = (A_1', A_2', \dots, A_p')'$. It is noticed, some authors define $\text{Vec}(\mathbf{A})$ as np column vector obtained by stacking together the rows of $\mathbf{A}$ one underneath the other. Both concepts are equivalent through transposition. Indeed, $\text{Vec}(\mathbf{A})$ obtained that way corresponds to $\text{Vec}(\mathbf{A}')$ used in this paper. Further, to simplify the notation, let $\boldsymbol{\varrho}_n = \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})\mathbf{L}$, let $\boldsymbol{\xi}_n = \sqrt{n}(\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}})\mathbf{L}$ and let $\boldsymbol{\zeta}_n = \sqrt{n}(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta})\mathbf{L}$. Further, let $\mathbf{J} = \boldsymbol{\Lambda}\mathbf{F}'(\mathbf{F}\boldsymbol{\Lambda}\mathbf{F}')^{-1}$, let $\boldsymbol{\delta}^* = \mathbf{J}\boldsymbol{\delta}$ and let $\boldsymbol{\Sigma} = \mathbf{J}\mathbf{F}\boldsymbol{\Lambda}$. The following proposition plays a central role in establishing the main result. Indeed, even though the original measurements are not normals, the following proposition allows us to apply some properties of matrices normal variate.

Proposition 2.1. *Let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ be iid random matrices with $\mathrm{E}(\mathbf{X}_1) = \boldsymbol{\theta}$ and $\text{Var}(\text{Vec}(\mathbf{X}_1)) = \boldsymbol{\Omega} \otimes \boldsymbol{\Lambda}$ with $\boldsymbol{\Omega}$, $\boldsymbol{\Lambda}$ positive definite matrices. Then, under local alternative in (1.4), we have*

$$\begin{pmatrix} \boldsymbol{\varrho}_n \\ \boldsymbol{\xi}_n \end{pmatrix} \xrightarrow[n\to\infty]{\mathcal{L}} \begin{pmatrix} \boldsymbol{\varrho} \\ \boldsymbol{\xi} \end{pmatrix} \sim \mathcal{N}_{2k\times 2m}\left(\begin{pmatrix} \mathbf{0} \\ \boldsymbol{\delta}^* \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Lambda} \otimes (\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) & \boldsymbol{\Sigma} \otimes (\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) \\ \boldsymbol{\Sigma} \otimes (\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) & \boldsymbol{\Sigma} \otimes (\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) \end{pmatrix}\right).$$

Furthermore,

$$\begin{pmatrix} \boldsymbol{\xi}_n \\ \boldsymbol{\zeta}_n \end{pmatrix} \xrightarrow[n\to\infty]{\mathcal{L}} \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{pmatrix} \sim \mathcal{N}_{2k\times 2m}\left(\begin{pmatrix} \boldsymbol{\delta}^* \\ -\boldsymbol{\delta}^* \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} \otimes \mathbf{L}'\boldsymbol{\Omega}\mathbf{L} & \mathbf{0} \\ \mathbf{0} & (\boldsymbol{\Lambda} - \boldsymbol{\Sigma}) \otimes \mathbf{L}'\boldsymbol{\Omega}\mathbf{L} \end{pmatrix}\right).$$

The proof is outlined in the Appendix.

Further, below we present two theorems which extend Theorem 1 and Theorem 2 in Bock and Judge (1978).

Theorem 2.1. *Let*

$$(\mathbf{X}', \mathbf{Y}')' \sim \mathcal{N}_{2q\times 2k}\left((\mathbf{M}_1, \mathbf{M}_2), \begin{pmatrix} \boldsymbol{\Upsilon}_{11} \otimes \boldsymbol{\Lambda}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Upsilon}_{22} \otimes \boldsymbol{\Lambda}_{22} \end{pmatrix}\right),$$

where $\boldsymbol{\Lambda}_{11}$ *is a positive definite matrix, and* $\boldsymbol{\Upsilon}_{11}$, $\boldsymbol{\Upsilon}_{11}$, $\boldsymbol{\Lambda}_{22}$ *are nonnegative definite matrices with rank* $p \leq k$. *Also, let* $\boldsymbol{\Xi}$ *be a symmetric and positive definite matrix which satisfies the two following conditions*:

(i) $\boldsymbol{\Upsilon}_{11}\boldsymbol{\Xi}$ *is idempotent matrix*;
(ii) $\boldsymbol{\Xi}\boldsymbol{\Upsilon}_{11}\boldsymbol{\Xi}\mathbf{M}_1 = \boldsymbol{\Xi}\mathbf{M}_1$.

Then, for any h Borel measurable and real-valued integrable function, and any nonnegative definite matrix $\mathbf{A}$, *we have*

$$\begin{aligned}&\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}_{11}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}_{11}\boldsymbol{\Xi}\mathbf{X}))\mathbf{Y}'\mathbf{A}\mathbf{X}]\\&\quad = \mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\boldsymbol{\Lambda}_{11}^{-1}\mathbf{M}_1'\boldsymbol{\Xi}\boldsymbol{\Upsilon}_{11}\boldsymbol{\Xi}\mathbf{M}_1)))]\mathbf{M}_2'\mathbf{A}\mathbf{M}_1.\end{aligned}$$

Proof. Since $\mathbf{X}$ and $\mathbf{Y}$ are independent, we have

$$\begin{aligned}&\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}_{11}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}_{11}\boldsymbol{\Xi}\mathbf{X}))\mathbf{Y}'\mathbf{A}\mathbf{X}]\\&\quad = (\mathrm{E}(\mathbf{Y}))'\mathbf{A}\,\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}_{11}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}_{11}\boldsymbol{\Xi}\mathbf{X}))\mathbf{X}],\end{aligned}$$

and then, the rest of the proof follows directly from Theorem A.1 given in the Appendix. □

Also, the following theorem is the crucial role in deriving the $\mathrm{ADR}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta})$.

Theorem 2.2. *Let* $\mathbf{X} \sim \mathcal{N}_{q\times k}(\mathbf{M}, \boldsymbol{\Upsilon}\otimes\boldsymbol{\Lambda})$ *where* $\boldsymbol{\Lambda}$ *is a positive definite matrix, and* $\boldsymbol{\Upsilon}$ *is a nonnegative definite matrix with rank* $p \leq k$. *Also, let* $\mathbf{A}$ *and* $\boldsymbol{\Xi}$ *be positive definite symmetric matrices and assume that* $\boldsymbol{\Xi}$ *satisfies the two following conditions*:

(i) $\boldsymbol{\Upsilon}\boldsymbol{\Xi}$ *is idempotent matrix*;
(ii) $\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{M} = \boldsymbol{\Xi}\mathbf{M}$.

Then, for any h Borel measurable and real-valued integrable function, we have

$$\begin{aligned}&\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))\,\mathrm{trace}(\mathbf{X}'\mathbf{A}\mathbf{X})]\\&\quad = \mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{M}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{M})))]\,\mathrm{trace}(\mathbf{A}\boldsymbol{\Upsilon})\,\mathrm{trace}(\boldsymbol{\Lambda})\\&\qquad + \mathrm{E}[h(\chi^2_{pq+4}(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{M}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{M})))]\,\mathrm{trace}(\mathbf{M}'\mathbf{A}\mathbf{M}).\end{aligned}$$

The proof of this theorem is given in the Appendix. Note that the theorem extends Theorem 2 given in Judge and Bock (1978) that becomes a particular case by taking $k = 1$, $\boldsymbol{\Upsilon} = \boldsymbol{\Xi} = 1$, $\boldsymbol{\Lambda} = \mathbf{I}_q$. Using Proposition 2.1 along with Theorems 2.1 and 2.2, we establish the asymptotic distributional risk function of the estimator $\widehat{\boldsymbol{\theta}}$, that is given in the following section.

3 Main result

Theorem 3.1. *Assume that Proposition 2.1 holds. Then, the asymptotic distributional risk function of* $\widehat{\boldsymbol{\theta}}$ *is*

$$\begin{aligned}
&\mathrm{ADR}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}; \mathbf{W}) \\
&\quad = \mathrm{trace}(\mathbf{W}(\boldsymbol{\Lambda} - \Sigma))\,\mathrm{trace}(\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) + \mathrm{trace}(\boldsymbol{\delta}^{*\prime}\mathbf{W}\boldsymbol{\delta}^{*}) \\
&\qquad - 2\mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}((\mathbf{L}'\boldsymbol{\Omega}\mathbf{L})^{-1}\boldsymbol{\delta}^{*\prime}\boldsymbol{\Xi}_1\boldsymbol{\delta}^{*})))]\,\mathrm{trace}(\boldsymbol{\delta}^{*\prime}\mathbf{W}\boldsymbol{\delta}^{*}) \\
&\qquad + \mathrm{E}[h^2(\chi^2_{pq+2}(\mathrm{trace}((\mathbf{L}'\boldsymbol{\Omega}\mathbf{L})^{-1}\boldsymbol{\delta}^{*\prime}\boldsymbol{\Xi}_1\boldsymbol{\delta}^{*})))]\,\mathrm{trace}(\mathbf{W}\boldsymbol{\Sigma})\,\mathrm{trace}(\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) \\
&\qquad + \mathrm{E}[h^2(\chi^2_{pq+4}(\mathrm{trace}((\mathbf{L}'\boldsymbol{\Omega}\mathbf{L})^{-1}\boldsymbol{\delta}^{*\prime}\boldsymbol{\Xi}_1\boldsymbol{\delta}^{*})))]\,\mathrm{trace}(\boldsymbol{\delta}^{*\prime}\mathbf{W}\boldsymbol{\delta}^{*}).
\end{aligned}$$

Proof. We have

$$\mathrm{ADR}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}; \mathbf{W}) = \mathrm{E}\big\{\mathrm{trace}\big[\big(\boldsymbol{\eta} + h(\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2})\boldsymbol{\xi}\big)'\mathbf{W}\big(\boldsymbol{\eta} + h(\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2})\boldsymbol{\xi}\big)\big]\big\}$$

and since h is a real-valued function, we get

$$\begin{aligned}
\mathrm{ADR}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}; \mathbf{W}) = \mathrm{E}\{\mathrm{trace}[\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}]\} &+ 2\mathrm{E}\{h(\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2})\,\mathrm{trace}[\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\eta}]\} \\
&+ \mathrm{E}\{h^2(\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2})\,\mathrm{trace}[\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\xi}]\}.
\end{aligned}$$

Then, combining Theorem 2.1, Theorem 2.2 and Proposition 2.1, we get

$$\begin{aligned}
\mathrm{E}\{\mathrm{trace}[\boldsymbol{\eta}'\mathbf{W}\boldsymbol{\eta}]\} &= \mathrm{trace}\big(\mathbf{W}(\boldsymbol{\Lambda} - \Sigma)\big)\,\mathrm{trace}(\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}) + \mathrm{trace}(\boldsymbol{\delta}^{*\prime}\mathbf{W}\boldsymbol{\delta}^{*}) \\
\mathrm{E}\{h(\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2})\,\mathrm{trace}[\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\eta}]\} &= -\mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\delta}^{*\prime}\boldsymbol{\Xi}_1\boldsymbol{\delta}^{*})))] \\
&\qquad \times \mathrm{trace}(\boldsymbol{\delta}^{*\prime}\mathbf{W}\boldsymbol{\delta}^{*})
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}\{h^2(\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Xi}_1,\boldsymbol{\Xi}_2})\,\mathrm{trace}[\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\xi}]\} \\
&\quad = \mathrm{E}[h^2(\chi^2_{pq+4}(\mathrm{trace}((\mathbf{L}'\boldsymbol{\Omega}\mathbf{L})^{-1}\boldsymbol{\delta}^{*\prime}\boldsymbol{\Xi}_1\boldsymbol{\delta}^{*})))]\,\mathrm{trace}(\boldsymbol{\delta}^{*\prime}\mathbf{W}\boldsymbol{\delta}^{*}) \\
&\qquad + \mathrm{E}[h^2(\chi^2_{pq+2}(\mathrm{trace}((\mathbf{L}'\boldsymbol{\Omega}\mathbf{L})^{-1}\boldsymbol{\delta}^{*\prime}\boldsymbol{\Xi}_1\boldsymbol{\delta}^{*})))]\,\mathrm{trace}(\mathbf{W}\boldsymbol{\Sigma})\,\mathrm{trace}(\mathbf{L}'\boldsymbol{\Omega}\mathbf{L}),
\end{aligned}$$

that completes the proof. □

4 Conclusion

In this article, we derive a formula for the asymptotic distributional risk function of a class of estimator concerning the common mean parameter matrix of a random sample of n matrices variables. For the sake of simplicity, we assume that the random matrices are independent and identically distributed. However, the result holds for a general ergodic and strictly process which satisfies the conditions for the central limit theorem for stationary processes. Also, we present two results which extend Theorem 1 and Theorem 2 in Bock and Judge (1978). Finally, the established result is useful in evaluating the performance of a class of Shrinkage-pretest types estimators.

Appendix

In this section, we present some technical details that are used in establishing the main result. The following lemma is useful in deriving the asymptotic joint distribution of the estimators $\overline{\mathbf{X}}$ and $\widetilde{\boldsymbol{\theta}}$ given in (1.1).

Lemma A.1. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be* $m \times p$ *and* $n \times q$ *matrices respectively. Further, let* $\mathbf{X} \sim \mathcal{N}_{p\times n}(\boldsymbol{\mu}, \Omega \otimes \Phi)$. *Then,* $\mathbf{AXB} \sim \mathcal{N}_{m\times q}(\mathbf{A}\boldsymbol{\mu}\mathbf{B}, (\mathbf{A}\Omega\mathbf{A}') \otimes (\mathbf{B}'\Phi\mathbf{B}))$.

The proof of Lemma A.1 follows directly from the following well-known algebraic property on matrices vectorization (see, e.g., Izenman, 2008 among others).

Lemma A.2. *Let* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *be matrices such that* $\mathbf{ABC}$ *is well defined. Also, let* $\mathbf{I}$ *denote an identity matrix. Then,* $\mathrm{Vec}(\mathbf{ABC}) = (\mathbf{AB} \otimes \mathbf{I})\,\mathrm{Vec}(\mathbf{C}) = (\mathbf{A} \otimes \mathbf{C}')\,\mathrm{Vec}(\mathbf{B})$.

Proof of Proposition 2.1. To simplify the proof, let us replace $\widetilde{\boldsymbol{\theta}}$ by $\widetilde{\boldsymbol{\theta}}_0 = \overline{\mathbf{X}} - \boldsymbol{\Lambda}\mathbf{F}'(\mathbf{F}\boldsymbol{\Lambda}\mathbf{F}')^{-1}(\mathbf{F}\overline{\mathbf{X}}\mathbf{L} - \mathbf{d})(\mathbf{L}'\boldsymbol{\Omega}\mathbf{L})^{-1}\mathbf{L}'\boldsymbol{\Omega}$. Indeed, since $\widehat{\boldsymbol{\Lambda}}$ and $\widehat{\boldsymbol{\Omega}}$ are strongly consistent, the estimators $\widetilde{\boldsymbol{\theta}}_0$ and $\widetilde{\boldsymbol{\theta}}$ are asymptotically equivalent, as n tends to infinity. Also, let $\boldsymbol{\varrho}_n = \sqrt{n}(\overline{\mathbf{X}} - \boldsymbol{\theta})\mathbf{L}$, let $\boldsymbol{\xi}_{0,n} = \sqrt{n}(\overline{\mathbf{X}} - \widetilde{\boldsymbol{\theta}}_0)\mathbf{L}$ and let $\boldsymbol{\zeta}_{0,n} = \sqrt{n}(\widetilde{\boldsymbol{\theta}}_0 - \boldsymbol{\theta})\mathbf{L}$. We have

$$(\boldsymbol{\varrho}_n', \boldsymbol{\zeta}_{0,n}')' = (\mathbf{I}_k, \mathbf{I}_k - \mathbf{F}'\mathbf{J}')'\boldsymbol{\varrho}_n - \mathbf{J}\boldsymbol{\delta}.$$

Then, using the central limit theorem along with Slutsky theorem, we get $(\boldsymbol{\varrho}_n', \boldsymbol{\zeta}_{0,n}')' \xrightarrow{\mathcal{L}}_{n\to\infty} (\mathbf{I}_k, \mathbf{I}_k - \mathbf{F}'\mathbf{J}')'\boldsymbol{\varrho} - \mathbf{J}\boldsymbol{\delta}$ where $\boldsymbol{\varrho} \sim \mathcal{N}_{k\times m}(\mathbf{0}, \boldsymbol{\Lambda} \otimes \mathbf{L}'\boldsymbol{\Omega}\mathbf{L})$. Hence,

$$(\boldsymbol{\varrho}_n', \boldsymbol{\zeta}_{0,n}')' \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}_{2k\times 2m}\big(-\mathbf{J}\boldsymbol{\delta}, (\mathbf{I}_k, \mathbf{I}_k - \mathbf{F}'\mathbf{J}')'\boldsymbol{\Lambda}(\mathbf{I}_k, \mathbf{I}_k - \mathbf{F}\mathbf{J}) \otimes \mathbf{L}'\boldsymbol{\Omega}\mathbf{L}\big). \quad \text{(A.1)}$$

Further,

$$\begin{pmatrix} \boldsymbol{\varrho}_n \\ \boldsymbol{\xi}_{0,n} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{I}_k & -\mathbf{I}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\varrho}_n \\ \boldsymbol{\zeta}_{0,n} \end{pmatrix}. \quad \text{(A.2)}$$

Therefore, by combining (A.1) and (A.2), we get the first statement of the proposition. The second statement of the proposition is proved by noting that

$$\begin{pmatrix} \boldsymbol{\zeta}_{0,n} \\ \boldsymbol{\xi}_{0,n} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_k & -\mathbf{I}_k \\ \mathbf{0} & \mathbf{I}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\varrho}_n \\ \boldsymbol{\xi}_{0,n} \end{pmatrix}, \quad \text{(A.3)}$$

and then, the proof follows directly from the first statement of the proposition and Slutsky theorem. □

Theorem A.1. *Let* $\mathbf{X} \sim \mathcal{N}_{q\times k}(\mathbf{M}, \boldsymbol{\Upsilon} \otimes \boldsymbol{\Lambda})$ *where* $\boldsymbol{\Lambda}$ *is a positive definite matrix, and* $\boldsymbol{\Upsilon}$ *is a nonnegative definite matrix with rank* $p \leq k$. *Also, let* $\boldsymbol{\Xi}$ *be a symmetric and positive definite matrix which satisfies the two following conditions*:

(i) $\boldsymbol{\Upsilon\Xi}$ *is idempotent matrix*;
(ii) $\boldsymbol{\Xi\Upsilon\Xi M} = \boldsymbol{\Xi M}$.

Then, for any h Borel measurable and real-valued integrable function, we have

$$\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi\Upsilon\Xi}\mathbf{X}))\mathbf{X}] = \mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{M}'\boldsymbol{\Xi\Upsilon\Xi}\mathbf{M})))]\mathbf{M}.$$

Proof. Since $\boldsymbol{\Upsilon\Xi}$ is idempotent matrix, $\boldsymbol{\Xi}^{1/2}\boldsymbol{\Upsilon}\boldsymbol{\Xi}^{1/2}$ is a symmetric and idempotent matrix and then, there exists an orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}\boldsymbol{\Xi}^{1/2}\boldsymbol{\Upsilon}\boldsymbol{\Xi}^{1/2}\mathbf{Q}' = \begin{pmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{A.4}$$

Moreover, let $\mathbf{V} = \mathbf{Q}\boldsymbol{\Xi}^{1/2}\mathbf{X}\boldsymbol{\Lambda}^{-1/2}$. Then, $\mathrm{Vec}(\mathbf{V}) = (\mathbf{Q}\boldsymbol{\Xi}^{1/2} \otimes \boldsymbol{\Lambda}^{-1/2})\,\mathrm{Vec}(\mathbf{X})$ and hence,

$$\mathrm{Vec}(\mathbf{V}) = \begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{pmatrix} \sim \mathcal{N}_{q\times k}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \otimes \mathbf{I}_q\right), \tag{A.5}$$

with

$$\boldsymbol{\mu}_1 = [\mathbf{I}_{pq}, \mathbf{0}]\mathrm{E}(\mathrm{Vec}(\mathbf{V})) = ([\mathbf{I}_p, \mathbf{0}] \otimes \mathbf{I}_q)(\mathbf{Q}\boldsymbol{\Xi}^{1/2} \otimes \boldsymbol{\Lambda}^{-1/2})\,\mathrm{Vec}(\mathbf{M}). \tag{A.6}$$

Therefore,

$$\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi\Upsilon\Xi}\mathbf{X}) = \mathrm{trace}(\mathbf{V}'\mathbf{Q}\boldsymbol{\Xi}^{1/2}\boldsymbol{\Upsilon}\boldsymbol{\Xi}^{1/2}\mathbf{Q}'\mathbf{V}) = \mathbf{V}_1'\mathbf{V}_1.$$

Hence,

$$\mathrm{Vec}(\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi\Upsilon\Xi}\mathbf{X}))\mathbf{X}]) = \begin{pmatrix} \mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\mathbf{V}_1] \\ \mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\mathbf{V}_2] \end{pmatrix}. \tag{A.7}$$

Also, from (A.5), $\mathbf{V}_2$ is zero with probability one and then,

$$\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\mathbf{V}_2] = \mathbf{0}, \tag{A.8}$$

and using Theorem 1 in Judge and Bock (1978), we have

$$\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\mathbf{V}_1] = \boldsymbol{\mu}_1\mathrm{E}(h(\chi^2_{pq+2}(\boldsymbol{\mu}_1'\boldsymbol{\mu}_1))), \tag{A.9}$$

where $\boldsymbol{\mu}_1$ is given by (A.6). Further, from (A.6),

$$\boldsymbol{\mu}_1'\boldsymbol{\mu}_1 = \mathrm{Vec}(\mathbf{M})'(\boldsymbol{\Xi}^{1/2}\mathbf{Q}' \otimes \boldsymbol{\Lambda}^{-1/2})\left(\begin{pmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \otimes \mathbf{I}_q\right)(\mathbf{Q}\boldsymbol{\Xi}^{1/2} \otimes \boldsymbol{\Lambda}^{-1/2})\,\mathrm{Vec}(\mathbf{M}),$$

and using (A.4), we get $\boldsymbol{\mu}_1'\boldsymbol{\mu}_1 = \mathrm{Vec}(\mathbf{M})'(\boldsymbol{\Xi\Upsilon\Xi} \otimes \boldsymbol{\Lambda}^{-1})\,\mathrm{Vec}(\mathbf{M})$, and then

$$\boldsymbol{\mu}_1'\boldsymbol{\mu}_1 = \mathrm{trace}(\mathbf{M}'\boldsymbol{\Xi\Upsilon\Xi}\mathbf{M}\boldsymbol{\Lambda}^{-1}). \tag{A.10}$$

Further, combining (A.7), (A.8) and (A.9), we have

$$\begin{aligned}&\mathrm{Vec}(\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))\mathbf{X}])\\ &\quad=\mathrm{E}(h(\chi^2_{pq+2}(\boldsymbol{\mu}_1'\boldsymbol{\mu}_1)))(\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\otimes\boldsymbol{\Lambda}^{1/2})\left(\begin{pmatrix}\mathbf{I}_p & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\otimes\mathbf{I}_q\right)\\ &\qquad\times(\mathbf{Q}\boldsymbol{\Xi}^{1/2}\otimes\boldsymbol{\Lambda}^{-1/2})\,\mathrm{Vec}(\mathbf{M}),\end{aligned}$$

that gives

$$\mathrm{Vec}(\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))\mathbf{X}])=\mathrm{E}(h(\chi^2_{pq+2}(\boldsymbol{\mu}_1'\boldsymbol{\mu}_1)))\,\mathrm{Vec}(\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{M}).$$

Then, since $\boldsymbol{\Xi}$ is nonsingular matrix, we have $\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{M}=\mathbf{M}$. Therefore,

$$\mathrm{Vec}(\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))\mathbf{X}])=\mathrm{Vec}(\mathrm{E}(h(\chi^2_{pq+2}(\boldsymbol{\mu}_1'\boldsymbol{\mu}_1)))\mathbf{M}),$$

that completes the proof. □

Proof of Theorem 2.2. From the above computations, we have

$$h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))=h(\mathbf{V}_1'\mathbf{V}_1).$$

Further, as in Theorem 1, we have $\mathbf{V}=\mathbf{Q}\boldsymbol{\Xi}^{1/2}\mathbf{X}\boldsymbol{\Lambda}^{-1/2}$ where $\mathbf{Q}$ is the same as in (A.4). We have $\mathbf{X}=\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\mathbf{V}\boldsymbol{\Lambda}^{1/2}$ and hence,

$$\mathbf{X}'\mathbf{A}\mathbf{X}=\boldsymbol{\Lambda}^{1/2}\mathbf{V}'\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\mathbf{V}\boldsymbol{\Lambda}^{1/2}. \tag{A.11}$$

Therefore,

$$\begin{aligned}&\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))\,\mathrm{trace}(\mathbf{X}'\mathbf{A}\mathbf{X})]\\ &\quad=\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\,\mathrm{trace}(\boldsymbol{\Lambda}^{1/2}\mathbf{V}'\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\mathbf{V}\boldsymbol{\Lambda}^{1/2})],\end{aligned}$$

then,

$$\begin{aligned}&\mathrm{E}[h(\mathrm{trace}(\boldsymbol{\Lambda}^{-1}\mathbf{X}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\boldsymbol{\Xi}\mathbf{X}))\,\mathrm{trace}(\mathbf{X}'\mathbf{A}\mathbf{X})]\\ &\quad=\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\,\mathrm{trace}(\boldsymbol{\Lambda}^{1/2}\mathbf{V}'\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\mathbf{V}\boldsymbol{\Lambda}^{1/2})]\\ &\quad=\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\,\mathrm{trace}(\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}')].\end{aligned}$$

Also, we have

$$\mathrm{trace}(\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}')=(\mathrm{Vec}(\mathbf{V}))'(\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}'\otimes\boldsymbol{\Lambda})\,\mathrm{Vec}(\mathbf{V}).$$

Further, let

$$(\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}\otimes\boldsymbol{\Lambda})=\mathbf{G}=\begin{pmatrix}\mathbf{G}_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{pmatrix}. \tag{A.12}$$

Therefore, since $\mathrm{Vec}(\mathbf{V})=(\mathbf{V}_1',\mathbf{0})'$, we have

$$(\mathrm{Vec}(\mathbf{V}))'(\boldsymbol{\Lambda}\otimes\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q})\,\mathrm{Vec}(\mathbf{V})=\mathbf{V}_1'\mathbf{G}_{11}\mathbf{V}_1,$$

and then,

$$\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)(\mathrm{Vec}(\mathbf{V}))'(\boldsymbol{\Lambda}\otimes\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q})\,\mathrm{Vec}(\mathbf{V})] \\ =\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\mathbf{V}_1'\mathbf{G}_{11}\mathbf{V}_1],$$

and using Theorem 2 in Judge and Bock (1978), we have

$$\mathrm{E}[h(\mathbf{V}_1'\mathbf{V}_1)\mathbf{V}_1'\mathbf{G}_{11}\mathbf{V}_1]=\mathrm{E}[h(\chi^2_{pq+2}(\mathrm{trace}(\boldsymbol{\mu}_1'\boldsymbol{\mu}_1)))]\,\mathrm{trace}(\mathbf{G}_{11}) \\ +\mathrm{E}[h(\chi^2_{pq+4}(\mathrm{trace}(\boldsymbol{\mu}_1'\boldsymbol{\mu}_1)))](\boldsymbol{\mu}_1'\mathbf{G}_{11}\boldsymbol{\mu}_1), \tag{A.13}$$

where $\boldsymbol{\mu}_1$ is given by (A.6). Further, note that

$$\mathbf{G}_{11}=([\mathbf{I}_p,\mathbf{0}]\otimes\mathbf{I}_q)\mathbf{G}\left[\begin{pmatrix}\mathbf{I}_p\\ \mathbf{0}\end{pmatrix}\otimes\mathbf{I}_q\right] \tag{A.14}$$

and then, combining this relation with (A.6), we get

$$\boldsymbol{\mu}_1'\mathbf{G}_{11}\boldsymbol{\mu}_1=\mathrm{Vec}(\mathbf{M})'(\boldsymbol{\Xi}^{1/2}\mathbf{Q}'\otimes\boldsymbol{\Lambda}^{-1/2})\left(\begin{pmatrix}\mathbf{I}_p & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\otimes\mathbf{I}_q\right)\mathbf{G}\left(\begin{pmatrix}\mathbf{I}_p & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\otimes\mathbf{I}_q\right) \\ \times(\mathbf{Q}\boldsymbol{\Xi}^{1/2}\otimes\boldsymbol{\Lambda}^{-1/2})\,\mathrm{Vec}(\mathbf{M}),$$

that gives

$$\boldsymbol{\mu}_1'\mathbf{G}_{11}\boldsymbol{\mu}_1=\mathrm{trace}(\mathbf{M}\mathbf{M}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}\mathbf{A}).$$

Also, under Assumption (ii), $\mathbf{M}'\boldsymbol{\Xi}\boldsymbol{\Upsilon}=\mathbf{M}'$, and then

$$\boldsymbol{\mu}_1'\mathbf{G}_{11}\boldsymbol{\mu}_1=\mathrm{trace}(\mathbf{M}\mathbf{M}'\mathbf{A})=\mathrm{trace}(\mathbf{M}'\mathbf{A}\mathbf{M}). \tag{A.15}$$

Further, using (A.4), (A.12) and (A.14), we have

$$\mathrm{trace}(\mathbf{G}_{11})=\mathrm{trace}[(\mathbf{Q}\boldsymbol{\Xi}^{-1/2}\mathbf{A}\boldsymbol{\Xi}^{-1/2}\mathbf{Q}\otimes\boldsymbol{\Lambda})(\mathbf{Q}\boldsymbol{\Xi}^{1/2}\boldsymbol{\Upsilon}\boldsymbol{\Xi}^{1/2}\mathbf{Q}'\otimes\mathbf{I}_q)] \\ =\mathrm{trace}(\mathbf{A}\boldsymbol{\Upsilon})\,\mathrm{trace}(\boldsymbol{\Lambda}). \tag{A.16}$$

Finally, the proof is completed by combining (A.10), (A.13), (A.15) and (A.16). □

Acknowledgments

The author is thankful to the Editor-in-Chief and anonymous referees for their valuable time spent on this paper. Also, the author acknowledges the financial support received from Natural Sciences and Engineering Research Council of Canada.

References

Ahmed, S. E. and Saleh, A. K. M. E. (1999). Improved nonparametric estimation of location vectors in a multivariate regression models. *Nonparametric Statistics* **11**, 51–78. MR1719441

Ahmed, S. E. (2001). Shrinkage estimation of regression coefficients from censored data with multiple observations. In *Empirical Bayes and and Likelihood Inference* (S. E. Ahmed and N. Reid, eds.). *Lecture Notes in Statist.* 103–120. NewYork: Springer. MR1855562

Eagleson, G. K. (1975). On Gordin's central limit theorem for stationary processes. *Journal of Applied Probability* **12**, 176–179. MR0383501

Izenman, A. J. (2008). *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*. *Springer Texts in Statistics.* New York: Springer. MR2445017

Judge, G. G. and Bock, M. E. (1978). *The statistical Implication of Pre-test and Stein-Rule Estimators in Econometrics*. Amsterdam: North Holland. MR0483199

Department of Mathematics and Statistics
University of Windsor
401 Sunset Avenue, Windsor, Ontario
Canada N9B 3P4
E-mail: severien@uwindsor.ca