

FINITE SAMPLE PROPERTIES OF ESTIMATORS FOR THE OPTIMAL PORTFOLIO WEIGHT

Harunori Mori*

This paper considers the problem of estimating the optimal portfolio weight to the mean-variance model in finance when parameters are unknown. For this purpose, we consider the following two classes of estimators. One is the class of proportional type estimators and the other is the class of Stein type estimators. First, we derive an unbiased estimator of the optimal portfolio weight, which belongs to the class of proportional type estimators. Second, we obtain dominance results within each class. From this, we showed that the unbiased proportional estimator and the maximum likelihood estimator are inadmissible.

Key words and phrases: Mean-variance model, simultaneous estimation, statistical decision theory, portfolio selection.

1. Introduction

The mean-variance model is a standard framework for defining optimal portfolio weight invested in risky assets. When we allow unrestricted short sales, the model is given by

$$\begin{aligned} & \text{Maximize} && \boldsymbol{\mu}'\boldsymbol{x} - a\boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x} \\ & \text{subject to} && \mathbf{1}'\boldsymbol{x} = 1, \end{aligned}$$

where \boldsymbol{x} is a portfolio weight, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are mean vector and positive definite covariance matrix of returns of the K risky assets, a is a risk aversion parameter satisfying $a > 0$ and $\mathbf{1}$ is a $K \times 1$ vector whose components are equal to 1. We denote the optimal portfolio weight by $\boldsymbol{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

In practice, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown and must be estimated from available historical information. We denote the return vector of K risky assets at period t by \boldsymbol{z}_t and assume that \boldsymbol{z}_t , $t = 1, \dots, n$ are independently distributed as a K -variate normal distribution $N_K(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given by

$$\boldsymbol{m} = \frac{1}{n} \sum_{t=1}^n \boldsymbol{z}_t \quad \text{and} \quad \boldsymbol{S} = \frac{1}{n} \sum_{t=1}^n (\boldsymbol{z}_t - \boldsymbol{m})(\boldsymbol{z}_t - \boldsymbol{m})',$$

respectively. The maximum likelihood estimator of the optimal portfolio weight is given by $\boldsymbol{x}^*(\boldsymbol{m}, \boldsymbol{S})$. For many analysts and investors, \boldsymbol{m} and \boldsymbol{S} seem obvious and reliable estimators (Michaud (1998)). However, many authors pointed out that estimation error can degrade the desirable properties of the optimal portfolio

when portfolio optimization is implemented using \mathbf{m} and \mathbf{S} . This problem is referred to as the problem of estimation risk in portfolio selection. Although many authors raised serious objections to the mean-variance model as a framework for defining optimal portfolio weight and proposed a number of alternatives, the mean-variance model serves as the standard optimization framework for modern asset management. Thus it is important to examine the problem statistically under the mean-variance model. However, previous attempts to consider the problem have been inconclusive.

From the point of view of statistics, the problem can be summarized in two issues. The first issue is an investigation of distributional properties of the maximum likelihood estimator $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$. Mori (2001) derived the asymptotic distribution of $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$. Jobson and Korkie (1980) also derived the asymptotic distribution of the maximum likelihood estimator of the optimal portfolio weight under the slightly different setup. However, little is known about the distributional properties of $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$ because it is a complicated function of \mathbf{m} and \mathbf{S} . Ledoit (1995) investigated, in the context of the mean-variance model, asymptotic properties of estimators of the covariance matrix Σ when the number of assets is large.

The second issue is an improvement on $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$. From a Monte Carlo simulation, Frost and Savarino (1986) and Jorion (1986) showed that the improvement can be obtained through the use of Stein type estimators of $\boldsymbol{\mu}$ and Σ instead of \mathbf{m} and \mathbf{S} , respectively. They defined the Stein type estimators as empirical Bayes estimators. Jobson *et al.* (1979) also showed similar results from a Monte Carlo simulation under the same setup as Jobson and Korkie (1980) although the use of Stein type estimator of Σ did not seem to be advantageous. Their results raise a statistically interesting question which is whether the Stein type estimator dominating $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$ exists. Hui *et al.* (1993) and Board and Sutcliffe (1994) supported their results by empirical studies.

This paper is concerned with the above two issues from the point of view of statistics. We consider the following quadratic programming problem with generalized linear restrictions rather than the mean-variance model:

$$(1.1) \quad \begin{aligned} & \text{Maximize} && \boldsymbol{\mu}'\mathbf{x} - a\mathbf{x}'\Sigma\mathbf{x} \\ & \text{subject to} && \mathbf{A}'\mathbf{x} = \mathbf{b}, \end{aligned}$$

where \mathbf{A} is a $K \times q$ matrix of rank q and \mathbf{b} is a $q \times 1$ vector. We denote the objective function of the problem (1.1) by $\pi(\mathbf{x})$, i.e.,

$$\pi(\mathbf{x}) \equiv \boldsymbol{\mu}'\mathbf{x} - a\mathbf{x}'\Sigma\mathbf{x}.$$

The optimal solution to the problem (1.1) is given by

$$\begin{aligned} \mathbf{x}^*(\boldsymbol{\mu}, \Sigma) = & \frac{1}{2a} \left(\Sigma^{-1} - \Sigma^{-1}\mathbf{A}(\mathbf{A}'\Sigma^{-1}\mathbf{A})^{-1}\mathbf{A}'\Sigma^{-1} \right) \boldsymbol{\mu} \\ & + \Sigma^{-1}\mathbf{A}(\mathbf{A}'\Sigma^{-1}\mathbf{A})^{-1}\mathbf{b}. \end{aligned}$$

In this paper, we focus on the problem of estimating $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. For this purpose, we consider the following two classes of estimators. One is the class of proportional type estimators. The maximum likelihood estimator $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$ belongs to this class. The other is the class of Stein type estimators. Jobson *et al.* (1979), Frost and Savarino (1986) and Jorion (1986) proposed an estimator $\mathbf{x}^*(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ where $\hat{\boldsymbol{\mu}}$ is a Stein type estimator of $\boldsymbol{\mu}$. Our definition of Stein type estimator is different from theirs and it is given in Section 2.1. We first derive an unbiased estimator of $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and its covariance matrix. Next we consider an improvement of $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$ under a certain loss function within each class. We show that there are estimators dominating $\mathbf{x}^*(\mathbf{m}, \mathbf{S})$ in each class.

In Section 2, we first formulate the problem of estimation risk from the point of view of statistics. The two classes of estimators and the loss function are defined in this section. The unbiased estimator and two dominance results are given in Theorem 2.1, 2.2 and 2.3, respectively. In Section 3, 4 and 5, we prove these theorems. Finally, brief conclusions are given in Section 6.

2. Estimation of the optimal solution

2.1. Formulation of the problem of estimation risk

Although previous studies dealing with the problem of estimation risk considered the problem of estimating $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, we need not estimate them separately. We focus on the problem of estimating $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The usual squared error loss function does not seem to be appropriate for this problem. If $\pi(\hat{\mathbf{x}})$ is much smaller than the maximum, then $\hat{\mathbf{x}}$ is not a good estimator even if $\hat{\mathbf{x}}$ is close to $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Because the objective of the problem (1.1) is to maximize $\pi(\mathbf{x})$, it is adequate to estimate $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ under the loss function

$$(2.1) \quad L(\hat{\mathbf{x}}) \equiv \pi(\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})) - \pi(\hat{\mathbf{x}}).$$

Frost and Savarino (1986) and Jorion (1986) used in their simulation studies the similar loss function to evaluate numerically their estimators of the optimal portfolio weight.

Next we define the following two functions for a convenience of notation.

DEFINITION 2.1. *Let \mathbf{W} be a $K \times K$ symmetric and nonsingular matrix and \mathbf{A} be a $K \times q$ matrix of rank q .*

$$(1) \quad \mathbf{F}_1(\mathbf{A}, \mathbf{W}) \equiv \mathbf{W}^{-1} - \mathbf{W}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{W}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{W}^{-1}$$

$$(2) \quad \mathbf{F}_2(\mathbf{A}, \mathbf{W}) \equiv \mathbf{W}^{-1}\mathbf{A}(\mathbf{A}'\mathbf{W}^{-1}\mathbf{A})^{-1}$$

By using these functions, we can represent the optimal solution as

$$\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2a}\mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma})\boldsymbol{\mu} + \mathbf{F}_2(\mathbf{A}, \boldsymbol{\Sigma})\mathbf{b}.$$

In this paper, we consider the following two types of estimators of $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ based on \mathbf{m} and \mathbf{S} .

DEFINITION 2.2. *Let c_P , c_S and d be positive constants.*

(1) *The proportional type estimator*

$$\hat{\boldsymbol{x}}_P(c_P) \equiv \frac{c_P}{2a} \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{S}) \boldsymbol{m} + \boldsymbol{F}_2(\boldsymbol{A}, \boldsymbol{S}) \boldsymbol{b}.$$

(2) *The Stein type estimator*

$$\hat{\boldsymbol{x}}_S(c_S, d) \equiv \frac{c_S}{2a} \left(1 - \frac{d}{\boldsymbol{m}' \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{S}) \boldsymbol{m}} \right) \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{S}) \boldsymbol{m} + \boldsymbol{F}_2(\boldsymbol{A}, \boldsymbol{S}) \boldsymbol{b}.$$

Jobson *et al.* (1979), Frost and Savarino (1986) and Jorion (1986) also proposed the use of the Stein type estimator of $\boldsymbol{\mu}$, although their estimators are different from $(1 - d/\boldsymbol{m}' \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{S}) \boldsymbol{m}) \boldsymbol{m}$. We derive $\hat{\boldsymbol{x}}_S(c_S, d)$ by taking account of the optimization problem (1.1). The derivation is explained in Section 5.

2.2. Main results

We first state some properties of the proportional type estimator $\hat{\boldsymbol{x}}_P(c_P)$. The following theorem gives the condition for $\hat{\boldsymbol{x}}_P(c_P)$ to be unbiased.

THEOREM 2.1. *Let $n > \max\{K + 1, K - q + 2\}$ and*

$$(2.2) \quad c_P^{UB} = \frac{n - K + q - 2}{n}.$$

Then $\hat{\boldsymbol{x}}_P(c_P^{UB})$ is unbiased. If $n > \max\{K + 1, K - q + 4\}$, then the covariance matrix of $\hat{\boldsymbol{x}}_P(c_P^{UB})$ exists and is given by

$$\begin{aligned} & V[\hat{\boldsymbol{x}}_P(c_P^{UB})] \\ &= \frac{n - K + q}{4a^2(n - K + q - 1)(n - K + q - 4)} \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu} \boldsymbol{\mu}' \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{\Sigma}) \\ &+ \left\{ \frac{n - K + q - 2}{4a^2(n - K + q - 1)(n - K + q - 4)} \left(\frac{n - 2}{n} + \boldsymbol{\mu}' \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu} \right) \right. \\ &\quad \left. + \frac{\boldsymbol{b}' (\boldsymbol{A}' \boldsymbol{\Sigma}^{-1} \boldsymbol{A})^{-1} \boldsymbol{b}}{n - K + q - 2} \right\} \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{\Sigma}). \end{aligned}$$

The proof is given in Section 3.

Note that the unbiased estimator $\hat{\boldsymbol{x}}_P(c_P^{UB})$ is the minimum variance unbiased estimator since \boldsymbol{m} and \boldsymbol{S} are complete sufficient statistics.

From Theorem 2.1, we can obtain the bias of $\hat{\boldsymbol{x}}_P(c_P)$. It is

$$\text{Bias}[\hat{\boldsymbol{x}}_P(c_P)] = \frac{nc_P - (n - K + q - 2)}{2a(n - K + q - 2)} \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu}.$$

We note that $\hat{\boldsymbol{x}}_P(1) = \boldsymbol{x}^*(\boldsymbol{m}, \boldsymbol{S})$, which is the maximum likelihood estimator of $\boldsymbol{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Hence the bias of $\hat{\boldsymbol{x}}_P(1)$ is given by

$$\text{Bias}[\hat{\boldsymbol{x}}_P(1)] = \frac{K - q + 2}{2a(n - K + q - 2)} \boldsymbol{F}_1(\boldsymbol{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu}.$$

Next we state a dominance result within the class of proportional type estimators.

THEOREM 2.2. *Let $n > \max\{K + 1, K - q + 4\}$ and*

$$(2.3) \quad 0 < c_P^* \leq \frac{(n - K + q - 1)(n - K + q - 4)}{n(n - 2)},$$

then $\hat{\mathbf{x}}_P(c_P^)$ dominates $\hat{\mathbf{x}}_P(c_P)$ relative to the loss (2.1) provided*

$$(2.4) \quad c_P > \frac{2(n - K + q - 1)(n - K + q - 4)}{n(n - 2)} - c_P^*.$$

If c_P^ and c_P do not satisfy the condition (2.4), $\hat{\mathbf{x}}_P(c_P^*)$ and $\hat{\mathbf{x}}_P(c_P)$ do not dominate each other.*

The proof is given in Section 4.

The constant c_P^{UB} satisfies the condition (2.4) if

$$c_P^* = \frac{(n - K + q - 1)(n - K + q - 4)}{n(n - 2)}.$$

Thus Theorem 2.2 shows that the unbiased estimator $\hat{\mathbf{x}}_P(c_P^{UB})$ is inadmissible. Similarly, we see that the maximum likelihood estimator $\hat{\mathbf{x}}_P(1)$ is also inadmissible.

We state the following theorem on the Stein type estimator $\hat{\mathbf{x}}_S(c, d)$.

THEOREM 2.3. *Let $n > K + 1$ and $K > q + 2$ and*

$$(2.5) \quad c_S = c_P = c \geq \frac{(n - K + q - 1)(n - K + q - 2)}{n(n - 2)},$$

then $\hat{\mathbf{x}}_S(c, d)$ dominates $\hat{\mathbf{x}}_P(c)$ provided

$$0 < d < \frac{2(K - q - 2)(n - K + q - 1)}{n(n - 2)c}.$$

The proof is given in Section 5.

Theorem 2.3 shows that there are estimators dominating the unbiased estimator $\hat{\mathbf{x}}_P(c_P^{UB})$ and the maximum likelihood estimator $\hat{\mathbf{x}}_P(1)$ in the class of Stein type estimators since the condition (2.5) holds when $c = c_P^{UB}$ and $c = 1$. It is also noted that any Stein type estimator and any proportional type estimator satisfying the condition (2.3) do not dominate each other since the right-hand side of (2.5) is greater than the right-hand side of (2.3). However, no comparison between $\hat{\mathbf{x}}_S(c_S, d)$ and $\hat{\mathbf{x}}_P(c_P)$ has been made yet when $c_S \neq c_P$.

3. Some results on F_1 and F_2 and proof of Theorem 2.1

3.1. Some properties of the matrix functions F_1 and F_2

Throughout this paper, we use the following notations. We let \mathbf{A} be a $K \times q$ matrix of rank q and \mathbf{b} be a $q \times 1$ vector. Let \mathbf{P} be an orthogonal matrix such that

$$(3.1) \quad \mathbf{P} \mathbf{A} \mathbf{A}' \mathbf{P}' = \mathbf{\Lambda},$$

where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of the matrix $\mathbf{A}\mathbf{A}'$. Let Λ and \mathbf{P} be partitioned as

$$\Lambda = \begin{pmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \end{pmatrix},$$

where Λ_1 is a $q \times q$ diagonal matrix with positive diagonal elements and \mathbf{P}_1 is a $q \times K$ matrix and \mathbf{P}_2 is a $(K - q) \times K$ matrix. We denote

$$(3.2) \quad \mathbf{L} = \mathbf{P}_1 \mathbf{A}.$$

We have the following lemma on the matrix functions \mathbf{F}_1 and \mathbf{F}_2 .

LEMMA 3.1. *Let us define $\mathbf{U} = \mathbf{P}\mathbf{W}\mathbf{P}'$ and $\mathbf{V} = \mathbf{U}^{-1}$. Let \mathbf{U} be partitioned into q and $K - q$ rows and columns*

$$(3.3) \quad \mathbf{U} = \begin{pmatrix} \mathbf{P}_1 \mathbf{W} \mathbf{P}'_1 & \mathbf{P}_1 \mathbf{W} \mathbf{P}'_2 \\ \mathbf{P}_2 \mathbf{W} \mathbf{P}'_1 & \mathbf{P}_2 \mathbf{W} \mathbf{P}'_2 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_{11} & \mathbf{U}'_{21} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{pmatrix}$$

and let \mathbf{V} be partitioned as \mathbf{U} has been partitioned, then

- (i) $\mathbf{F}_1(\mathbf{A}, \mathbf{W}) = \mathbf{P}'_2 \mathbf{U}_{22}^{-1} \mathbf{P}_2$,
- (ii) $\mathbf{F}_2(\mathbf{A}, \mathbf{W}) = \mathbf{P}'_1 (\mathbf{L}')^{-1} - \mathbf{P}'_2 \mathbf{U}_{22}^{-1} \mathbf{U}_{21} (\mathbf{L}')^{-1}$.

PROOF. From the definitions of \mathbf{U} and \mathbf{V} , we have

$$\mathbf{F}_1(\mathbf{A}, \mathbf{W}) = \mathbf{P}' (\mathbf{V} - \mathbf{V}\mathbf{P}\mathbf{A}(\mathbf{A}'\mathbf{P}'\mathbf{V}\mathbf{P}\mathbf{A})^{-1} \mathbf{A}'\mathbf{P}'\mathbf{V}) \mathbf{P}.$$

Since

$$\mathbf{A}'\mathbf{P}' = \begin{pmatrix} \mathbf{A}'\mathbf{P}'_1 & \mathbf{A}'\mathbf{P}'_2 \end{pmatrix} = \begin{pmatrix} \mathbf{L}' & \mathbf{0} \end{pmatrix},$$

we have

$$\begin{aligned} & \mathbf{V}\mathbf{P}\mathbf{A}(\mathbf{A}'\mathbf{P}'\mathbf{V}\mathbf{P}\mathbf{A})^{-1} \mathbf{A}'\mathbf{P}'\mathbf{V} \\ &= \mathbf{V} \begin{pmatrix} \mathbf{L}' \\ \mathbf{0} \end{pmatrix} \left(\begin{pmatrix} \mathbf{L}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}'_{21} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{L}' \\ \mathbf{0} \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{L}' & \mathbf{0} \end{pmatrix} \mathbf{V} \\ &= \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}'_{21} \\ \mathbf{V}_{21} & \mathbf{V}_{21} \mathbf{V}_{11}^{-1} \mathbf{V}'_{21} \end{pmatrix}. \end{aligned}$$

Here it is noted that

$$(3.4) \quad \mathbf{U} = \begin{pmatrix} \mathbf{V}_{11}^{-1} + \mathbf{V}_{11}^{-1} \mathbf{V}'_{21} \mathbf{V}_{22.1}^{-1} \mathbf{V}_{21} \mathbf{V}_{11}^{-1} & -\mathbf{V}_{11}^{-1} \mathbf{V}'_{21} \mathbf{V}_{22.1}^{-1} \\ -\mathbf{V}_{22.1}^{-1} \mathbf{V}_{21} \mathbf{V}_{11}^{-1} & \mathbf{V}_{22.1}^{-1} \end{pmatrix},$$

where $\mathbf{V}_{22.1} = \mathbf{V}_{22} - \mathbf{V}_{21} \mathbf{V}_{11}^{-1} \mathbf{V}'_{21}$. Since $\mathbf{U}_{22} = \mathbf{V}_{22.1}^{-1}$ from this, it follows that

$$\mathbf{V} - \mathbf{V}\mathbf{P}\mathbf{A}(\mathbf{A}'\mathbf{P}'\mathbf{V}\mathbf{P}\mathbf{A})^{-1} \mathbf{A}'\mathbf{P}'\mathbf{V} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{22}^{-1} \end{pmatrix}.$$

Hence $\mathbf{F}_1(\mathbf{A}, \mathbf{W})$ can be written as

$$\mathbf{F}_1(\mathbf{A}, \mathbf{W}) = \mathbf{P}'_2 \mathbf{U}_{22}^{-1} \mathbf{P}_2.$$

Thus (i) is proved.

Similarly to the proof of (i), we have

$$\begin{aligned} \mathbf{F}_2(\mathbf{A}, \mathbf{W}) &= \mathbf{P}' \mathbf{V} \mathbf{P} \mathbf{A} (\mathbf{A}' \mathbf{P}' \mathbf{V} \mathbf{P} \mathbf{A})^{-1} \\ &= \mathbf{P}'_1 (\mathbf{L}')^{-1} + \mathbf{P}'_2 \mathbf{V}_{21} \mathbf{V}_{11}^{-1} (\mathbf{L}')^{-1}. \end{aligned}$$

From (3.4), we have $-\mathbf{U}_{22} \mathbf{V}_{21} \mathbf{V}_{11}^{-1} = \mathbf{U}_{21}$ or $\mathbf{V}_{21} \mathbf{V}_{11}^{-1} = -\mathbf{U}_{22}^{-1} \mathbf{U}_{21}$. Therefore,

$$\mathbf{F}_2(\mathbf{A}, \mathbf{W}) = \mathbf{P}'_1 (\mathbf{L}')^{-1} - \mathbf{P}'_2 \mathbf{U}_{22}^{-1} \mathbf{U}_{21} (\mathbf{L}')^{-1}.$$

Hence this completes the lemma. \square

From the above lemma, we have

$$(3.5) \quad \mathbf{F}_1(\mathbf{A}, \mathbf{I}) = \mathbf{P}'_2 \mathbf{P}_2 \quad \text{and} \quad \mathbf{F}_2(\mathbf{A}, \mathbf{I}) = \mathbf{P}'_1 (\mathbf{L}')^{-1}$$

since \mathbf{P} is an orthogonal matrix.

The following lemma is easy to establish.

LEMMA 3.2. *Let \mathbf{T} be a matrix of same order as \mathbf{W} and $c \neq 0$, then*

- (i) $\mathbf{F}_1(\mathbf{A}, \mathbf{W}) = c \mathbf{F}_1(\mathbf{A}, c\mathbf{W})$,
- (ii) $\mathbf{F}_2(\mathbf{A}, \mathbf{W}) = \mathbf{F}_2(\mathbf{A}, c\mathbf{W})$,
- (iii) $\mathbf{F}_1(\mathbf{A}, \mathbf{W}) = (\mathbf{T}')^{-1} \mathbf{F}_1(\mathbf{T}^{-1} \mathbf{A}, \mathbf{T}^{-1} \mathbf{W} (\mathbf{T}')^{-1}) \mathbf{T}^{-1}$,
- (iv) $\mathbf{F}_2(\mathbf{A}, \mathbf{W}) = (\mathbf{T}')^{-1} \mathbf{F}_2(\mathbf{T}^{-1} \mathbf{A}, \mathbf{T}^{-1} \mathbf{W} (\mathbf{T}')^{-1})$.

3.2. Mean vectors and covariance matrices of $\mathbf{F}_1 \mathbf{X}$ and $\mathbf{F}_2 \mathbf{b}$

We need mean vectors and covariance matrices of $\mathbf{F}_1(\mathbf{A}, \mathbf{W}) \mathbf{X}$ and $\mathbf{F}_2(\mathbf{A}, \mathbf{W}) \mathbf{b}$ to prove Theorem 2.1. First we have the following lemma on the mean vectors.

LEMMA 3.3. *Let $\mathbf{X} \sim N_K(\boldsymbol{\mu}, \tau \mathbf{I})$ with $\tau > 0$ and \mathbf{W} be distributed as a Wishart distribution $W_K(m, \mathbf{I})$ with $m > K$, where \mathbf{X} and \mathbf{W} are independent, then*

- (i) $E[\mathbf{F}_1(\mathbf{A}, \mathbf{W}) \mathbf{X}] = \frac{1}{m - K + q - 1} \mathbf{F}_1(\mathbf{A}, \mathbf{I}) \boldsymbol{\mu}$,
- (ii) $E[\mathbf{F}_2(\mathbf{A}, \mathbf{W}) \mathbf{b}] = \mathbf{F}_2(\mathbf{A}, \mathbf{I}) \mathbf{b}$.

PROOF. From (3.3), we have $\mathbf{U}_{22} = \mathbf{P}_2 \mathbf{W} \mathbf{P}'_2 \sim W_{K-q}(m, \mathbf{I})$ since \mathbf{P} is an orthogonal matrix. Therefore

$$E[\mathbf{U}_{22}^{-1}] = \frac{1}{m - K + q - 1} \mathbf{I}.$$

Note that $\mathbf{F}_1(\mathbf{A}, \mathbf{W}) = \mathbf{P}'_2 \mathbf{U}_{22}^{-1} \mathbf{P}_2$ from Lemma 3.1. From this and (3.5), it follows that

$$E[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}] = \frac{1}{m - K + q - 1} \mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}.$$

Thus (i) is proved.

Let $\mathbf{y}_i \sim N_K(\mathbf{0}, \mathbf{I})$, $i = 1, \dots, m$ be independent and $\mathbf{Y}' = (\mathbf{y}_1 \cdots \mathbf{y}_m)$, then we can represent \mathbf{U} in the form $\mathbf{U} = \mathbf{Y}'\mathbf{Y}$ by definition of the Wishart distribution. Let \mathbf{Y} be partitioned as $\mathbf{Y} = (\mathbf{Y}_1 \mathbf{Y}_2)$, where \mathbf{Y}_1 is a $m \times q$ matrix and \mathbf{Y}_2 is a $m \times (K - q)$ matrix. Then \mathbf{U} is written as

$$(3.6) \quad \begin{pmatrix} \mathbf{U}_{11} & \mathbf{U}'_{21} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{Y}'_1 \mathbf{Y}_1 & \mathbf{Y}'_1 \mathbf{Y}_2 \\ \mathbf{Y}'_2 \mathbf{Y}_1 & \mathbf{Y}'_2 \mathbf{Y}_2 \end{pmatrix}.$$

Since \mathbf{Y}_1 and \mathbf{Y}_2 are independent and $E[\mathbf{Y}_1] = \mathbf{0}$, we have

$$(3.7) \quad E[\mathbf{U}_{22}^{-1} \mathbf{U}_{21}] = E_{\mathbf{Y}_2} [(\mathbf{Y}'_2 \mathbf{Y}_2)^{-1} \mathbf{Y}'_2 E_{\mathbf{Y}_1}[\mathbf{Y}_1]] = \mathbf{0}.$$

Hence it follows from Lemma 3.1 and (3.5) that

$$\begin{aligned} E[\mathbf{F}_2(\mathbf{A}, \mathbf{W})] &= \mathbf{P}'_1 (\mathbf{L}')^{-1} - \mathbf{P}'_2 E[\mathbf{U}_{22}^{-1} \mathbf{U}_{21}] (\mathbf{L}')^{-1} \\ &= \mathbf{F}_2(\mathbf{A}, \mathbf{I}). \end{aligned}$$

This proves (ii). \square

The following lemma gives the covariance matrices of $\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}$ and $\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}$.

LEMMA 3.4. *Let $\mathbf{X} \sim N_K(\boldsymbol{\mu}, \tau\mathbf{I})$ with $\tau > 0$ and $\mathbf{W} \sim W_K(m, \mathbf{I})$ with $m > \max\{K, K - q + 3\}$ are independent, then*

(i) $V[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}]$

$$\begin{aligned} &= \frac{m - K + q + 1}{(m - K + q)(m - K + q - 1)^2(m - K + q - 3)} \\ &\quad \times \mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I}) \\ &\quad + \frac{\tau(m - 1) + \boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}}{(m - K + q)(m - K + q - 1)(m - K + q - 3)} \mathbf{F}_1(\mathbf{A}, \mathbf{I}), \end{aligned}$$

(ii) $\text{Cov}[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}, \mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] = \mathbf{0}$,

(iii) $V[\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] = \frac{\mathbf{b}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{b}}{m - K + q - 1} \mathbf{F}_1(\mathbf{A}, \mathbf{I})$.

PROOF. First (i) is an immediate consequence of Theorem 3.2 (iii) of Haff (1979).

From Lemma 3.1, we have

$$\begin{aligned} & \mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{W}) \\ &= \mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{P}_2\mathbf{X}\mathbf{b}'\mathbf{L}^{-1}\mathbf{P}_1 - \mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{P}_2\mathbf{X}\mathbf{b}'\mathbf{L}^{-1}\mathbf{U}'_{21}\mathbf{U}_{22}^{-1}\mathbf{P}_2. \end{aligned}$$

From (3.6), we have

$$\begin{aligned} & E[\mathbf{U}_{22}^{-1}\mathbf{P}_2\mathbf{X}\mathbf{b}'\mathbf{L}^{-1}\mathbf{U}'_{21}\mathbf{U}_{22}^{-1}] \\ &= E_{\mathbf{Y}_2}[(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}\mathbf{P}_2\boldsymbol{\mu}\mathbf{b}'\mathbf{L}^{-1}E_{\mathbf{Y}_1}[\mathbf{Y}'_1]\mathbf{Y}_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}] = \mathbf{0}. \end{aligned}$$

Then the expectation of $\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{W})$ is written as

$$(3.8) \quad \begin{aligned} E[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{W})] &= \frac{\mathbf{P}'_2\mathbf{P}_2\boldsymbol{\mu}\mathbf{b}'\mathbf{L}^{-1}\mathbf{P}_1}{m - K + q - 1} \\ &= \frac{\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{I})}{m - K + q - 1}. \end{aligned}$$

Hence, from Lemma 3.3, we have

$$\begin{aligned} & Cov[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}, \mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] \\ &= E[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{W})] - \frac{\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{I})}{m - K + q - 1} \\ &= \mathbf{0}. \end{aligned}$$

Thus (ii) is proved.

Similarly, from (3.5), (3.6), (3.7) and Lemma 3.1, we can show that

$$\begin{aligned} & V[\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] \\ &= E[\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{W})] - \mathbf{F}_2(\mathbf{A}, \mathbf{I})\mathbf{b}\mathbf{b}'\mathbf{F}'_2(\mathbf{A}, \mathbf{I}) \\ &= \mathbf{P}'_2E[\mathbf{U}_{22}^{-1}\mathbf{U}_{21}(\mathbf{L}')^{-1}\mathbf{b}\mathbf{b}'\mathbf{L}^{-1}\mathbf{U}'_{21}\mathbf{U}_{22}^{-1}]\mathbf{P}_2 \\ &= \mathbf{P}'_2E_{\mathbf{Y}_2}[(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}\mathbf{Y}'_2E_{\mathbf{Y}_1}[\mathbf{Y}_1(\mathbf{L}')^{-1}\mathbf{b}\mathbf{b}'\mathbf{L}^{-1}\mathbf{Y}'_1]\mathbf{Y}_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}]\mathbf{P}_2. \end{aligned}$$

We denote $\mathbf{H} = E[\mathbf{Y}_1(\mathbf{L}')^{-1}\mathbf{b}\mathbf{b}'\mathbf{L}^{-1}\mathbf{Y}'_1]$ and the (i, j) -th element of \mathbf{H} by h_{ij} and the i -th row vector of \mathbf{Y}_1 by \mathbf{Y}_{1i} . Since \mathbf{Y}_{1i} is the subvector formed by the first q elements of $\mathbf{y}_i \sim N_K(\mathbf{0}, \mathbf{I})$,

$$E[\mathbf{Y}_{1i}] = \mathbf{0}, \quad V[\mathbf{Y}_{1i}] = \mathbf{I} \quad \text{and} \quad Cov[\mathbf{Y}_{1i}, \mathbf{Y}_{1j}] = \mathbf{0}$$

for $i \neq j$. Hence we have

$$h_{ii} = \text{tr}\left((\mathbf{L}')^{-1}\mathbf{b}\mathbf{b}'\mathbf{L}^{-1}\right) = \mathbf{b}'\mathbf{L}^{-1}(\mathbf{L}')^{-1}\mathbf{b} \quad \text{and} \quad h_{ij} = 0$$

for $i \neq j$. Note that $\mathbf{A}'\mathbf{A} = \mathbf{A}'\mathbf{P}'\mathbf{P}\mathbf{A} = \mathbf{L}'\mathbf{L}$. Therefore, we have

$$\begin{aligned} V[\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] &= (\mathbf{b}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{b})\mathbf{P}'_2E[(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}\mathbf{Y}'_2\mathbf{Y}_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}]\mathbf{P}_2 \\ &= (\mathbf{b}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{b})\mathbf{P}'_2E[\mathbf{U}_{22}^{-1}]\mathbf{P}_2 \\ &= \frac{\mathbf{b}'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{b}}{m - K + q - 1}\mathbf{F}_1(\mathbf{A}, \mathbf{I}). \end{aligned}$$

Thus (iii) is proved. \square

3.3. Proof of Theorem 2.1

Let \mathbf{T} be a nonsingular matrix such that $\boldsymbol{\Sigma} = \mathbf{T}\mathbf{T}'$. From Lemma 3.2, the proportional type estimator $\hat{\mathbf{x}}_P(c_P)$ is rewritten as

$$(3.9) \quad \hat{\mathbf{x}}_P(c_P) = \frac{nc_P}{2a} (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1}) \mathbf{T}^{-1} \mathbf{m} \\ + (\mathbf{T}')^{-1} \mathbf{F}_2 (\mathbf{T}^{-1} \mathbf{A}, n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1}) \mathbf{b}.$$

We note that

$$(3.10) \quad \mathbf{T}^{-1} \mathbf{m} \sim N_K \left(\mathbf{T}^{-1} \boldsymbol{\mu}, \frac{1}{n} \mathbf{I} \right) \quad \text{and} \quad n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1} \sim W_K(n-1, \mathbf{I}),$$

since $\mathbf{m} \sim N_K(\boldsymbol{\mu}, (1/n)\boldsymbol{\Sigma})$ and $n\mathbf{S} \sim W_K(n-1, \boldsymbol{\Sigma})$. From Lemma 3.3, we have

$$E[\hat{\mathbf{x}}_P(c_P)] = \frac{nc_P}{2a(n-K+q-2)} (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{T}^{-1} \boldsymbol{\mu} \\ + (\mathbf{T}')^{-1} \mathbf{F}_2 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{b} \\ = \frac{nc_P}{2a(n-K+q-2)} (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{T}^{-1} \boldsymbol{\Sigma} (\mathbf{T}')^{-1}) \mathbf{T}^{-1} \boldsymbol{\mu} \\ + (\mathbf{T}')^{-1} \mathbf{F}_2 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{T}^{-1} \boldsymbol{\Sigma} (\mathbf{T}')^{-1}) \mathbf{b} \\ = \frac{nc_P}{2a(n-K+q-2)} \mathbf{F}_1 (\mathbf{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu} + \mathbf{F}_2 (\mathbf{A}, \boldsymbol{\Sigma}) \mathbf{b}.$$

Hence, if $c_P = c_P^{UB}$, then $E[\hat{\mathbf{x}}_P(c_P)] = \mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. This implies that $\hat{\mathbf{x}}_P(c_P^{UB})$ is an unbiased estimator and completes the proof of (2.2).

Similarly, applying (3.9) and Lemma 3.4 to the covariance matrix of the proportional type estimator, we have

$$V[\hat{\mathbf{x}}_P(c_P)] \\ = \frac{n^2 c_P^2}{4a^2} (\mathbf{T}')^{-1} V [\mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1}) \mathbf{T}^{-1} \mathbf{m}] \mathbf{T}^{-1} \\ + \frac{nc_P}{a} (\mathbf{T}')^{-1} Cov [\mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1}) \mathbf{T}^{-1} \mathbf{m}, \\ \mathbf{F}_2 (\mathbf{T}^{-1} \mathbf{A}, n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1}) \mathbf{b}] \mathbf{T}^{-1} \\ + (\mathbf{T}')^{-1} V [\mathbf{F}_2 (\mathbf{T}^{-1} \mathbf{A}, n\mathbf{T}^{-1} \mathbf{S} (\mathbf{T}')^{-1}) \mathbf{b}] \mathbf{T}^{-1} \\ = \frac{n^2 (s+1) c_P^2}{4a^2 s (s-1)^2 (s-3)} (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{T}^{-1} \boldsymbol{\mu} \boldsymbol{\mu}' (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{T}^{-1} \\ + \frac{n^2 c_P^2}{4a^2 s (s-1) (s-3)} \left(\frac{n-2}{n} + \boldsymbol{\mu}' (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{T}^{-1} \boldsymbol{\mu} \right) \\ \times (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{T}^{-1} \\ + \frac{\mathbf{b}' (\mathbf{A}' (\mathbf{T}')^{-1} \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{b}}{s-1} (\mathbf{T}')^{-1} \mathbf{F}_1 (\mathbf{T}^{-1} \mathbf{A}, \mathbf{I}) \mathbf{T}^{-1},$$

where $s = n - K + q - 1$. Since Lemma 3.2, the covariance matrix is rewritten as

$$\begin{aligned} V[\hat{\mathbf{x}}_P(c_P)] &= \frac{n^2(s+1)c_P^2}{4a^2s(s-1)^2(s-3)} \mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu} \boldsymbol{\mu}' \mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma}) \\ &\quad + \left\{ \frac{n^2c_P^2}{4a^2s(s-1)(s-3)} \left(\frac{n-2}{n} + \boldsymbol{\mu}' \mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu} \right) \right. \\ &\quad \left. + \frac{\mathbf{b}'(\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{b}}{s-1} \right\} \mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma}). \end{aligned}$$

Hence, if $c_P = c_P^{UB}$, then we can obtain the covariance matrix of the unbiased estimator $V[\hat{\mathbf{x}}_P(c_P^{UB})]$. Thus the proof of Theorem 2.1 is complete.

4. Proof of Theorem 2.2

First we prepare expectations of some functions of $\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}$ and $\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}$ in the following lemma.

LEMMA 4.1. *Let $\mathbf{X} \sim N_K(\boldsymbol{\mu}, \tau\mathbf{I})$ with $\tau > 0$ and $\mathbf{W} \sim W_K(m, \mathbf{I})$ with $m > \max\{K, K - q + 3\}$ be independent, then*

(i) $E[\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}]$

$$= \frac{(m-1)(\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu} + \tau(K-q))}{(m-K+q)(m-K+q-1)(m-K+q-3)},$$

(ii) $E[\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] = 0.$

PROOF. Note that $\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X} = \text{tr} \mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})$. From Theorem 3.3 (i) and Theorem 3.4 (i), we have

$$\begin{aligned} &E[\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}] \\ &= \frac{\text{tr} \mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})}{(m-K+q)(m-K+q-3)} \\ &\quad + \frac{\tau(m-1) + \boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}}{(m-K+q)(m-K+q-1)(m-K+q-3)} \text{tr} \mathbf{F}_1(\mathbf{A}, \mathbf{I}). \end{aligned}$$

From (3.5), we observe that

$$\text{tr} \mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I}) = \boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})^2\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}$$

and

$$\text{tr} \mathbf{F}_1(\mathbf{A}, \mathbf{I}) = K - q.$$

Hence we can get (i).

Similarly, from (3.5) and (3.8), we have

$$\begin{aligned}
E[\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}] &= \text{tr } E[\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}\mathbf{b}'\mathbf{F}_2'(\mathbf{A}, \mathbf{W})] \\
&= \frac{1}{m-K+q-1} \text{tr } \mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}\mathbf{b}'\mathbf{F}_2'(\mathbf{A}, \mathbf{I}) \\
&= \frac{1}{m-K+q-1} \boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\mathbf{F}_2(\mathbf{A}, \mathbf{I})\mathbf{b} \\
&= 0.
\end{aligned}$$

This completes the proof of Lemma 4.1. \square

PROOF OF THEOREM 2.2. It follows from (3.9) that

$$\begin{aligned}
&\pi(\hat{\mathbf{x}}_P(c_P)) \\
&= -\frac{n^2 c_P^2}{4a} \mathbf{m}'(\mathbf{T}')^{-1} \mathbf{F}_1(\mathbf{T}^{-1}\mathbf{A}, n\mathbf{T}^{-1}\mathbf{S}(\mathbf{T}')^{-1})^2 \mathbf{T}^{-1}\mathbf{m} \\
&\quad + \frac{n c_P}{2a} \boldsymbol{\mu}'(\mathbf{T}')^{-1} \mathbf{F}_1(\mathbf{T}^{-1}\mathbf{A}, n\mathbf{T}^{-1}\mathbf{S}(\mathbf{T}')^{-1}) \mathbf{T}^{-1}\mathbf{m} \\
&\quad - n c_P \mathbf{m}'(\mathbf{T}')^{-1} \mathbf{F}_1(\mathbf{T}^{-1}\mathbf{A}, n\mathbf{T}^{-1}\mathbf{S}(\mathbf{T}')^{-1}) \mathbf{F}_2(\mathbf{T}^{-1}\mathbf{A}, n\mathbf{T}^{-1}\mathbf{S}(\mathbf{T}')^{-1}) \mathbf{b} \\
&\quad + \text{terms which do not contain } c_P,
\end{aligned}$$

where \mathbf{T} is a nonsingular matrix such that $\boldsymbol{\Sigma} = \mathbf{T}\mathbf{T}'$. From (3.10), Lemma 3.3 and Lemma 4.1, the expectation of $\pi(\hat{\mathbf{x}}_P(c_P))$ is written as

$$\begin{aligned}
(4.1) \quad &E[\pi(\hat{\mathbf{x}}_P(c_P))] \\
&= -\frac{n^2(n-2)}{4a(n-K+q-1)(n-K+q-2)(n-K+q-4)} \\
&\quad \times \left(\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma})\boldsymbol{\mu} + \frac{K-q}{n} \right) c_P^2 \\
&\quad + \frac{n}{2a(n-K+q-2)} \boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma})\boldsymbol{\mu} c_P \\
&\quad + \text{terms which do not contain } c_P.
\end{aligned}$$

The constant c_P maximizing $E[\pi(\hat{\mathbf{x}}_P(c_P))]$ is

$$(4.2) \quad c_P^{**} = \frac{(n-K+q-1)(n-K+q-4)}{n(n-2)} \theta,$$

where

$$\theta = \frac{\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma})\boldsymbol{\mu}}{\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma})\boldsymbol{\mu} + (K-q)/n}.$$

The maximizer (4.2) minimizes the risk of $\hat{\mathbf{x}}_P(c_P)$. The maximizer (4.2) is the function of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and thus $\hat{\mathbf{x}}_P(c_P^{**})$ is not a statistic. Instead of using c_P^{**} , we consider the constant c_P^* of the form

$$c_P^* = \frac{(n-K+q-1)(n-K+q-4)}{n(n-2)} \alpha,$$

where $0 < \alpha \leq 1$. We also consider the constant c_P of the form

$$c_p = \frac{(n - K + q - 1)(n - K + q - 4)}{n(n - 2)}\beta,$$

where $\beta > 0$. In what follows, we will derive the range of α and β for which $\hat{\mathbf{x}}_P(c_P^*)$ dominates $\hat{\mathbf{x}}_P(c_P)$. We denote the difference between the risks of $\hat{\mathbf{x}}_P(c_P^*)$ and $\hat{\mathbf{x}}_P(c_P)$ by

$$g(\beta) = E [L(\hat{\mathbf{x}}_P(c_P)) - L(\hat{\mathbf{x}}_P(c_P^*))]$$

From (4.1), the function $g(\beta)$ is written as

$$g(\beta) = \frac{(n - K + q - 1)(n - K + q - 4)}{4a(n - 2)(n - K + q - 2)} \\ \times \left(\boldsymbol{\mu}' \mathbf{F}_1(\mathbf{A}, \boldsymbol{\Sigma}) \boldsymbol{\mu} + \frac{K - q}{n} \right) (\beta - \alpha)(\beta + \alpha - 2\theta).$$

From this, if $\beta - \alpha > 0$ and $\beta + \alpha - 2\theta > 0$, then $g(\beta) > 0$. Since $0 \leq \theta < 1$, it holds that $\beta + \alpha - 2\theta > 0$ if $\beta + \alpha - 2 > 0$. Moreover, since $0 < \alpha \leq 1$, it holds that $\beta - \alpha > 0$ if $\beta + \alpha - 2 > 0$. Hence, if $\beta + \alpha - 2 > 0$, then $g(\beta) > 0$. This implies that $\hat{\mathbf{x}}_P(c_P^*)$ dominates any proportional type estimators satisfying the condition (2.4). On the other hand, it depends on θ whether the inequality $g(\beta) > 0$ holds for the case $\beta + \alpha - 2 \leq 0$. This completes the proof. \square

5. Derivation of the Stein type estimator and proof of Theorem 2.3

5.1. Derivation of the Stein type estimator

When $\boldsymbol{\Sigma}$ is unknown, the standard Stein type estimator of $\boldsymbol{\mu}$ is given by

$$\mathbf{m}_{ST} = \left(1 - \frac{\alpha}{(\mathbf{m} - \boldsymbol{\nu})' \mathbf{S}^{-1} (\mathbf{m} - \boldsymbol{\nu})} \right) (\mathbf{m} - \boldsymbol{\nu}) + \boldsymbol{\nu},$$

where $0 < \alpha < 2(K - 2)/(N - K + 3)$ and $\boldsymbol{\nu}$ is an arbitrary fixed vector (see Anderson (1984, Theorem 5.3.1)). The Stein type estimator of $\boldsymbol{\mu}$ proposed by Jobson *et al.* (1979) and Jorion (1986) are of the same form as \mathbf{m}_{ST} . By substituting \mathbf{m}_{ST} and \mathbf{S} into the optimal solution $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we can obtain an estimator of $\mathbf{x}^*(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, which is different from the Stein type estimator $\hat{\mathbf{x}}_S(c_S, d)$. Our derivation of $\hat{\mathbf{x}}_S(c_S, d)$ takes account of the optimization problem (1.1) and it is described below.

Let $\mathbf{P}' = \begin{pmatrix} \mathbf{P}'_1 & \mathbf{P}'_2 \end{pmatrix}$ be the orthogonal matrix defined in (3.1), then a general solution to $\mathbf{A}'\mathbf{x} = \mathbf{b}$ is written as

$$(5.1) \quad \mathbf{x} = \mathbf{P}'_2 \mathbf{y} + \mathbf{x}_0,$$

where \mathbf{y} is an arbitrary $(K - q) \times 1$ vector and \mathbf{x}_0 is a particular solution. Substituting the general solution into the quadratic programming problem (1.1), the problem reduces to the following problem with respect to \mathbf{y} :

$$\text{Maximize} \quad \left(\frac{1}{a} \boldsymbol{\mu} - 2\boldsymbol{\Sigma} \mathbf{x}_0 \right)' \mathbf{P}'_2 \mathbf{y} - \mathbf{y}' \mathbf{P}_2 \boldsymbol{\Sigma} \mathbf{P}'_2 \mathbf{y},$$

which is equivalent to the problem

$$\text{Minimize } (\mathbf{z} - \mathbf{P}_2\boldsymbol{\mu})'(\mathbf{P}_2\boldsymbol{\Sigma}\mathbf{P}_2')^{-1}(\mathbf{z} - \mathbf{P}_2\boldsymbol{\mu}),$$

where

$$(5.2) \quad \mathbf{z} = 2a(\mathbf{P}_2\boldsymbol{\Sigma}\mathbf{P}_2'\mathbf{y} + \mathbf{P}_2\boldsymbol{\Sigma}\mathbf{x}_0).$$

We denote the optimal solutions to these problems by \mathbf{y}^* and \mathbf{z}^* , respectively. Thus we see that the estimation problem for \mathbf{z}^* is equivalent to the estimation problem for mean vector of a $(K - q)$ -variate normal distribution $N_{K-q}(\mathbf{P}_2\boldsymbol{\mu}, \mathbf{P}_2\boldsymbol{\Sigma}\mathbf{P}_2')$ under the loss function

$$(\hat{\mathbf{z}} - \mathbf{P}_2\boldsymbol{\mu})'(\mathbf{P}_2\boldsymbol{\Sigma}\mathbf{P}_2')^{-1}(\hat{\mathbf{z}} - \mathbf{P}_2\boldsymbol{\mu}),$$

where $\hat{\mathbf{z}}$ is an estimator of \mathbf{z}^* . From Theorem 5.3.1 of Anderson (1984), the Stein type estimator of $\mathbf{P}_2\boldsymbol{\mu}$ is given by

$$\hat{\mathbf{z}}_S = \left(1 - \frac{d}{\mathbf{m}'\mathbf{P}_2'(\mathbf{P}_2\mathbf{S}\mathbf{P}_2')^{-1}\mathbf{P}_2\mathbf{m}}\right) \mathbf{P}_2\mathbf{m} = \left(1 - \frac{d}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right) \mathbf{P}_2\mathbf{m},$$

where d is a positive constant. Therefore we can obtain an estimator of \mathbf{y}^* from (5.2) by replacing $\boldsymbol{\Sigma}$ with \mathbf{S} . The estimator is

$$\hat{\mathbf{y}}_S = \frac{1}{2a} \left(1 - \frac{d}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right) (\mathbf{P}_2\mathbf{S}\mathbf{P}_2')^{-1}\mathbf{P}_2\mathbf{m} - (\mathbf{P}_2\mathbf{S}\mathbf{P}_2')^{-1}\mathbf{P}_2\mathbf{S}\mathbf{x}_0.$$

Substituting $\mathbf{y} = \hat{\mathbf{y}}_S$ into (5.1) gives an estimator of \mathbf{x}^* defined by

$$(5.3) \quad \hat{\mathbf{x}}_S = \frac{1}{2a} \left(1 - \frac{d}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right) \mathbf{P}_2'(\mathbf{P}_2\mathbf{S}\mathbf{P}_2')^{-1}\mathbf{P}_2\mathbf{m} \\ - \mathbf{P}_2'(\mathbf{P}_2\mathbf{S}\mathbf{P}_2')^{-1}\mathbf{P}_2\mathbf{S}\mathbf{x}_0 + \mathbf{x}_0.$$

Note that $\mathbf{x}_0 = \mathbf{P}_1'(\mathbf{L}')^{-1}\mathbf{b}$ is a particular solution to $\mathbf{A}'\mathbf{x} = \mathbf{b}$, where $\mathbf{L} = \mathbf{P}_1\mathbf{A}$. Substituting this \mathbf{x}_0 into (5.3) and using Lemma 3.1, we can obtain the Stein type estimator with $c_S = 1$ defined by Definition 2.2:

$$\hat{\mathbf{x}}_S(1, d) = \frac{1}{2a} \left(1 - \frac{d}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right) \mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m} + \mathbf{F}_2(\mathbf{A}, \mathbf{S})\mathbf{b}.$$

5.2. Expectations of certain functions of \mathbf{F}_1 and \mathbf{F}_2

To prove Theorem 2.3, we here evaluate expectations of the following functions

$$\frac{\boldsymbol{\mu}'\mathbf{F}_1\mathbf{X}}{\mathbf{X}'\mathbf{F}_1\mathbf{X}}, \quad \frac{\mathbf{X}'\mathbf{F}_1^2\mathbf{X}}{\mathbf{X}'\mathbf{F}_1\mathbf{X}}, \quad \frac{\mathbf{X}'\mathbf{F}_1^2\mathbf{X}}{(\mathbf{X}'\mathbf{F}_1\mathbf{X})^2} \quad \text{and} \quad \frac{\mathbf{X}'\mathbf{F}_1\mathbf{F}_2\mathbf{b}}{\mathbf{X}'\mathbf{F}_1\mathbf{X}},$$

where $\mathbf{F}_1 = \mathbf{F}_1(\mathbf{A}, \mathbf{W})$ and $\mathbf{F}_2 = \mathbf{F}_2(\mathbf{A}, \mathbf{W})$. First we state the following lemma on the fundamental result of the Wishart distribution.

LEMMA 5.1. *Let $\mathbf{W} \sim W_K(m, \mathbf{I})$ with $m > K$ and \mathbf{W} be partitioned into 1 and $K - 1$ rows and columns*

$$\mathbf{W} = \begin{pmatrix} w_{11} & \mathbf{W}'_{21} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix},$$

then

$$E [\mathbf{W}'_{21} \mathbf{W}_{22}^{-2} \mathbf{W}_{21}] = \frac{K - 1}{m - K}.$$

PROOF. In the proof of Lemma 3.3 (ii), let $q = 1$, then we have

$$\mathbf{W}'_{21} \mathbf{W}_{22}^{-2} \mathbf{W}_{21} = \mathbf{Y}'_1 \mathbf{Y}_2 (\mathbf{Y}'_2 \mathbf{Y}_2)^{-2} \mathbf{Y}'_2 \mathbf{Y}_1.$$

Hence, we have

$$\begin{aligned} E [\mathbf{W}'_{21} \mathbf{W}_{22}^{-2} \mathbf{W}_{21}] &= E_{\mathbf{Y}_2} [E_{\mathbf{Y}_1} [\mathbf{Y}'_1 \mathbf{Y}_2 (\mathbf{Y}'_2 \mathbf{Y}_2)^{-2} \mathbf{Y}'_2 \mathbf{Y}_1]] \\ &= E_{\mathbf{Y}_2} [\text{tr}(\mathbf{Y}'_2 \mathbf{Y}_2)^{-1}] \\ &= \frac{K - 1}{m - K}, \end{aligned}$$

which proves Lemma 5.1. \square

LEMMA 5.2. *Let Z be distributed as a Poisson distribution with parameter $1/(2\tau)\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}$ with $\tau > 0$. Let $\mathbf{X} \sim N_K(\boldsymbol{\mu}, \tau\mathbf{I})$ with $K > q + 2$ and $\mathbf{W} \sim W_K(m, \mathbf{I})$ with $m > K$ be independent, then*

$$\begin{aligned} \text{(i)} \quad E \left[\frac{\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} \right] &= 1 - \frac{K - q - 2}{2} E \left[\frac{1}{Z + (K - q - 2)/2} \right], \\ \text{(ii)} \quad E \left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} \right] &= \frac{m - 1}{(m - K + q)(m - K + q - 1)}, \\ \text{(iii)} \quad E \left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{(\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X})^2} \right] &= \frac{m - 1}{2\tau(m - K + q)} E \left[\frac{1}{Z + (K - q - 2)/2} \right], \\ \text{(iv)} \quad E \left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} \right] &= 0. \end{aligned}$$

PROOF. Let \mathbf{e}_1 be a $(K - q) \times 1$ vector whose first element is 1 and otherwise 0 and \mathbf{Q} be an orthogonal matrix satisfying

$$(5.4) \quad \mathbf{Q}\mathbf{P}_2\mathbf{X} = \sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}\mathbf{e}_1.$$

We define a $(K - q) \times (K - q)$ matrix $\tilde{\mathbf{U}}$ by

$$\tilde{\mathbf{U}} = \mathbf{Q}\mathbf{U}_{22}\mathbf{Q}',$$

where \mathbf{U}_{22} is defined in (3.3). We denote $\tilde{\mathbf{V}} = \tilde{\mathbf{U}}^{-1}$ and the (i, j) -th element of $\tilde{\mathbf{V}}$ by \tilde{v}_{ij} . From Lemma 3.1 and (5.4), we have

$$(5.5) \quad \begin{aligned} \mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X} &= \mathbf{X}'\mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{P}_2\mathbf{X} \\ &= \mathbf{X}'\mathbf{P}'_2\mathbf{Q}'\tilde{\mathbf{U}}^{-1}\mathbf{Q}\mathbf{P}_2\mathbf{X} \\ &= (\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X})\tilde{v}_{11}. \end{aligned}$$

Let denote the (i, j) -th element of $\tilde{\mathbf{U}}$ by \tilde{u}_{ij} and $\tilde{\mathbf{U}}$ be partitioned into 1 and $K - q - 1$ rows and columns

$$\tilde{\mathbf{U}} = \begin{pmatrix} \tilde{u}_{11} & \tilde{\mathbf{U}}'_{21} \\ \tilde{\mathbf{U}}_{21} & \tilde{\mathbf{U}}_{22} \end{pmatrix},$$

then $\tilde{\mathbf{V}}$ is written as

$$(5.6) \quad \tilde{\mathbf{V}} = \tilde{v}_{11} \begin{pmatrix} 1 & -\tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-1} \\ -\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21} & \tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21}\tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-1} + \tilde{v}_{11}^{-1}\tilde{\mathbf{U}}_{22}^{-1} \end{pmatrix},$$

where $\tilde{v}_{11} = (\tilde{u}_{11} - \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21})^{-1}$. Hence we have

$$\begin{aligned} \mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X} &= \sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}\mathbf{P}'_2\mathbf{Q}'\tilde{\mathbf{U}}^{-1}\mathbf{e}_1 \\ &= \tilde{v}_{11}\sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}\mathbf{P}'_2\mathbf{Q}' \begin{pmatrix} 1 \\ -\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21} \end{pmatrix}. \end{aligned}$$

Note that

$$(5.7) \quad \tilde{\mathbf{U}} \sim W_{K-q}(m, \mathbf{I})$$

for any choice of \mathbf{Q} since $\mathbf{U}_{22} \sim W_{K-q}(m, \mathbf{I})$ and \mathbf{Q} is an orthogonal matrix. From this and (3.7), we can observe that $E[\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21}] = \mathbf{0}$ for any choice of \mathbf{Q} . Hence we have

$$(5.8) \quad \begin{aligned} &E \left[\frac{\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} \mid \mathbf{X} \right] \\ &= E \left[\frac{\sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}\boldsymbol{\mu}'\mathbf{P}'_2\mathbf{Q}'}{(\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X})} \begin{pmatrix} 1 \\ -\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21} \end{pmatrix} \mid \mathbf{X} \right] \\ &= \frac{\boldsymbol{\mu}'\mathbf{P}'_2\mathbf{Q}'}{\sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}} \begin{pmatrix} 1 \\ -E[\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21}] \end{pmatrix} \\ &= \frac{\boldsymbol{\mu}'\mathbf{P}'_2\mathbf{Q}'\mathbf{e}_1}{\sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}}. \end{aligned}$$

From (5.4), we see that $\mathbf{Q}'\mathbf{e}_1 = \mathbf{P}_2\mathbf{X}/\sqrt{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}$. Hence we can represent the conditional expectation (5.8) as

$$E \left[\frac{\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} \mid \mathbf{X} \right] = \frac{\boldsymbol{\mu}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}.$$

Note that $\mathbf{P}_2\mathbf{X} \sim N_{K-q}(\mathbf{P}_2\boldsymbol{\mu}, \tau\mathbf{I})$. From Lemma 1 of Baranchik (1973), we have

$$E \left[\frac{\boldsymbol{\mu}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}} \right] = 1 - \frac{K-q-2}{2} E \left[\frac{1}{Z + (K-q-2)/2} \right],$$

where Z is distributed as a Poisson distribution with parameter $1/(2\tau)\boldsymbol{\mu}'\mathbf{P}'_2\mathbf{P}_2\boldsymbol{\mu} = 1/(2\tau)\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}$. This completes the proof of (i).

Similar to (5.5), we have

$$\begin{aligned} \mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X} &= \mathbf{X}'\mathbf{P}'_2\mathbf{U}_{22}^{-2}\mathbf{P}_2\mathbf{X} \\ &= (\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X})\mathbf{e}'_1\tilde{\mathbf{U}}^{-2}\mathbf{e}_1. \end{aligned}$$

Note that $\mathbf{e}'_1\tilde{\mathbf{U}}^{-2}\mathbf{e}_1 = \tilde{v}_{11}^2(1 + \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-2}\tilde{\mathbf{U}}_{21})$ from (5.6). Hence we have

$$(5.9) \quad \mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X} = \tilde{v}_{11}^2(\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}) \left(1 + \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-2}\tilde{\mathbf{U}}_{21} \right).$$

From (5.5) and (5.9), we have

$$\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} = \tilde{v}_{11} \left(1 + \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-2}\tilde{\mathbf{U}}_{21} \right).$$

From (5.7) and 8b.2 (iii) of Rao (1973), we see that \tilde{v}_{11}^{-1} is distributed as a central Chi-square distribution with $m - K + q + 1$ degrees of freedom, i.e., $\tilde{v}_{11}^{-1} \sim \chi^2(m - K + q + 1)$. Hence,

$$E[\tilde{v}_{11}] = \frac{1}{m - K + q - 1}.$$

From Theorem 7.3.6 of Anderson (1984), $\tilde{v}_{11}^{-1} = \tilde{u}_{11} - \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21}$ and $\tilde{\mathbf{U}}_{22}^{-1}\tilde{\mathbf{U}}_{21}$ are independent. Therefore, from Lemma 5.1, we have

$$\begin{aligned} E \left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} \right] &= E \left[\tilde{v}_{11} \left(1 + \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-2}\tilde{\mathbf{U}}_{21} \right) \right] \\ &= \frac{m-1}{(m-K+q)(m-K+q-1)}. \end{aligned}$$

Thus (ii) is proved.

From (5.5) and (5.9), we see that

$$\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{(\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X})^2} = \frac{1 + \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-2}\tilde{\mathbf{U}}_{21}}{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}.$$

Hence, from Lemma 5.1, we have

$$\begin{aligned} E \left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{(\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X})^2} \middle| \mathbf{X} \right] &= \frac{1}{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}} E \left[1 + \tilde{\mathbf{U}}'_{21}\tilde{\mathbf{U}}_{22}^{-2}\tilde{\mathbf{U}}_{21} \middle| \mathbf{X} \right] \\ &= \frac{m-1}{(m-K+q)} \frac{1}{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}. \end{aligned}$$

Since $\tau^{-1}\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}$ is distributed as a noncentral Chi-square distribution with $K - q$ degrees of freedom and noncentral parameter $\tau^{-1}\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}$, we have

$$\begin{aligned} E\left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})^2\mathbf{X}}{(\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X})^2}\right] &= \frac{m-1}{\tau(m-K+q)}E\left[\frac{\tau}{\mathbf{X}'\mathbf{P}'_2\mathbf{P}_2\mathbf{X}}\right] \\ &= \frac{m-1}{2\tau(m-K+q)}E\left[\frac{1}{Z+(K-q-2)/2}\right], \end{aligned}$$

where Z is distributed as a Poisson distribution with parameter $1/(2\tau)\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{I})\boldsymbol{\mu}$. Thus (iii) is proved.

From Lemma 3.1, we have

$$\begin{aligned} \frac{\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{F}_2(\mathbf{A}, \mathbf{W})}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} &= \frac{\mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{P}_2(\mathbf{P}'_1(\mathbf{L}')^{-1} - \mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{U}_{21}(\mathbf{L}')^{-1})}{\mathbf{X}'\mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{P}_2\mathbf{X}} \\ &= -\frac{\mathbf{P}'_2\mathbf{U}_{22}^{-2}\mathbf{U}_{21}(\mathbf{L}')^{-1}}{\mathbf{X}'\mathbf{P}'_2\mathbf{U}_{22}^{-1}\mathbf{P}_2\mathbf{X}}. \end{aligned}$$

From (3.6), this is rewritten as

$$\frac{\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{F}_2(\mathbf{A}, \mathbf{W})}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}} = -\frac{\mathbf{P}'_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-2}\mathbf{Y}'_2\mathbf{Y}_1(\mathbf{L}')^{-1}}{\mathbf{X}'\mathbf{P}'_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}\mathbf{P}_2\mathbf{X}}.$$

Since \mathbf{X} , \mathbf{Y}_1 , and \mathbf{Y}_2 are independent and $E[\mathbf{Y}_1] = \mathbf{0}$, we have

$$\begin{aligned} &E\left[\frac{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{F}_2(\mathbf{A}, \mathbf{W})\mathbf{b}}{\mathbf{X}'\mathbf{F}_1(\mathbf{A}, \mathbf{W})\mathbf{X}}\right] \\ &= -E\left[\frac{\mathbf{X}'\mathbf{P}'_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-2}\mathbf{Y}'_2}{\mathbf{X}'\mathbf{P}'_2(\mathbf{Y}'_2\mathbf{Y}_2)^{-1}\mathbf{P}_2\mathbf{X}}\right]E[\mathbf{Y}_1](\mathbf{L}')^{-1}\mathbf{b} \\ &= 0. \end{aligned}$$

This completes the proof of (iv). Thus Lemma 5.2 is proved. \square

5.3. Proof of Theorem 2.3

We denote the difference between the risks of $\hat{\mathbf{x}}_P(c)$ and $\hat{\mathbf{x}}_S(c, d)$ by

$$\begin{aligned} g(d) &= E[L(\hat{\mathbf{x}}_P(c)) - L(\hat{\mathbf{x}}_S(c, d))] \\ &= -\frac{c^2d^2}{4a}E\left[\frac{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\boldsymbol{\Sigma}\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}{(\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m})^2}\right] - \frac{cd}{2a}E\left[\frac{\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right] \\ &\quad + \frac{c^2d}{2a}E\left[\frac{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\boldsymbol{\Sigma}\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right] \\ &\quad + cdE\left[\frac{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\boldsymbol{\Sigma}\mathbf{F}_2(\mathbf{A}, \mathbf{S})\mathbf{b}}{\mathbf{m}'\mathbf{F}_1(\mathbf{A}, \mathbf{S})\mathbf{m}}\right]. \end{aligned}$$

From Lemma 3.2, we see that

$$\begin{aligned} \mathbf{F}_1(\mathbf{A}, \mathbf{S}) &= n(\mathbf{T}')^{-1}\mathbf{F}_1(\mathbf{T}^{-1}\mathbf{A}, n\mathbf{T}^{-1}\mathbf{S}(\mathbf{T}')^{-1})\mathbf{T}^{-1} \quad \text{and} \\ \mathbf{F}_2(\mathbf{A}, \mathbf{S}) &= (\mathbf{T}')^{-1}\mathbf{F}_2(\mathbf{T}^{-1}\mathbf{A}, n\mathbf{T}^{-1}\mathbf{S}(\mathbf{T}')^{-1}), \end{aligned}$$

where \mathbf{T} is a nonsingular matrix such that $\Sigma = \mathbf{T}\mathbf{T}'$. Here we substitute these into $g(d)$. From (3.10) and Lemma 5.2, we have

$$(5.10) \quad g(d) = -\frac{c^2 n(n-2)}{8a(n-K+q-1)} G d^2 - \frac{c}{2a} \left(1 - \frac{K-q-2}{2} G\right) d \\ + \frac{c^2 n(n-2)}{2a(n-K+q-1)(n-K+q-2)} d,$$

where G is the following expectation with respect to Z ,

$$G = E \left[\frac{1}{Z + (K-q-2)/2} \right],$$

where Z is distributed as a Poisson distribution with parameter $(n/2)\boldsymbol{\mu}'\mathbf{F}_1(\mathbf{A}, \Sigma)\boldsymbol{\mu}$. The roots of the quadratic equation $g(d) = 0$ are 0 and

$$(5.11) \quad d(c, G) = \frac{2(K-q-2)(n-K+q-1)}{n(n-2)c} \\ - \frac{4((n-K+q-1)(n-K+q-2) - n(n-2)c)}{n(n-2)(n-K+q-2)cG}.$$

If d belongs to the interval between these two roots, then $g(d) > 0$. We see that

$$\inf_{G>0} d(c, G) = \frac{2(K-q-2)(n-K+q-1)}{n(n-2)c} > 0$$

with $c \geq n^{-1}(n-2)^{-1}(n-K+q-1)(n-K+q-2)$. Hence, if $0 < d < \inf_{G>0} d(c, G)$, then $g(d) > 0$ for any G . This completes the proof of Theorem 2.3.

6. Conclusions

In this paper, we considered the problem of estimating the optimal solution to the problem (1.1). For the purpose, we examined two classes of estimators. One is the class of proportional type estimators and the other is the class of Stein type estimators. We showed that there are estimators dominating the maximum likelihood estimator in each class. However, there are no simple dominance results between two classes.

There are two problems in our results from the point of view of practical applications. First, the problem (1.1) does not include inequality constraints, which are important in practical applications. Second, we assumed that the return of a risky asset is distributed as a normal distribution. However, many empirical studies in finance have concluded that the normality assumption is suspect. The extension of this study, which takes account of nonnegativity constraints and nonnormality assumptions, is important in both finance and statistics. Thus this extension awaits future studies.

Acknowledgements

I am grateful to Professor Jun-ichiro Fukuchi for helpful discussions. I am also grateful to anonymous referees for helpful comments that improved this paper considerably.

REFERENCES

- Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed., John Wiley.
- Baranchik, A. J. (1973). Inadmissibility of maximum likelihood estimator in some multiple regression problems with three or more independent variables, *Annals of Statistics*, **1**, 312–321.
- Board, J. L. G. and Sutcliffe, C. M. S. (1994). Estimation methods in portfolio selection and the effectiveness of short sales restrictions: UK evidence, *Management Science*, **40**, 516–534.
- Frost, P. A. and Savarino, J. E. (1986). An empirical Bayes approach to efficient portfolio selection, *Journal of Financial and Quantitative Analysis*, **21**, 293–305.
- Haff, L. R. (1979). An identity for the Wishart distribution with applications, *Journal of Multivariate Analysis*, **9**, 531–544.
- Hui, T., Kwan, E. K. and Lee, C. (1993). Optimal portfolio diversification: Empirical Bayes versus classical approach, *Journal of Operational Research*, **44**, 1155–1159.
- Jobson, J. D. and Korkie, B. (1980). Estimation for Markowitz efficient portfolios, *Journal of the American Statistical Association*, **75**, 544–554.
- Jobson, J. D., Korkie, B. and Ratti, V. (1979). Improved estimation for Markowitz portfolios using James-Stein type estimators, *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, **71**, 279–284.
- Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis*, **21**, 279–292.
- Ledoit, O. (1995). Essays on risk and return in the stock market, Doctoral theses, Massachusetts Institute of Technology.
- Michaud, R. O. (1998). *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*, Oxford University Press.
- Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., John Wiley.
- Mori, H. (2001). Asymptotic inadmissibility of maximum likelihood estimator in a quadratic programming problem, *Gakushuin Economic Paper*, **38**, 35–49 (in Japanese).