Ann. Inst. Statist. Math.
Vol. 43, No. 1, 157–165 (1991)

A NOTE ON ESTIMATING EIGENVALUES OF SCALE MATRIX OF THE MULTIVARIATE F-DISTRIBUTION

YOSHIHIKO KONNO

Faculty of Business Administration, Ishinomaki Senshu University, Ishinomaki, Miyagi 986, Japan

(Received January 9, 1989; revised July 23, 1990)

Abstract. Let $F_{p\times p}$ have the multivariate F-distribution with a scale matrix Δ and degrees of freedom n_1 and n_2. In this paper the problem of estimating eigenvalues of Δ is considered. By constructing the improved orthogonally invariant estimators $\hat{\Delta}(F)$ of Δ, which are analogous to Haff-type estimators of a normal covariance matrix, new estimators of eigenvalues of Δ are given. This is because the eigenvalues of $\hat{\Delta}(F)$ are taken as estimates of the eigenvalues of Δ.

Key words and phrases: Estimation of eigenvalues, multivariate F-distribution, covariance matrix, orthogonally invariant estimators.

1. Introduction

Suppose that a $p \times p$ positive definite random matrix F has the probability density function

$$\text{(1.1)} \qquad \frac{\Gamma_p\left(\frac{1}{2}(n_1+n_2)\right)}{\Gamma_p\left(\frac{1}{2}n_1\right)\Gamma_p\left(\frac{1}{2}n_2\right)}(\det\Delta)^{-n_1/2}\frac{(\det F)^{(n_1-p-1)/2}}{\det(I_p+\Delta^{-1}F)^{n/2}}$$

where $n_1 > p+1$, $n_2 > p+1$, $n = n_1+n_2$, Δ is a $p \times p$ positive definite matrix and $\Gamma_p(a) = \pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma(a-(i-1)/2)$. The matrix F having the distribution given above does not arise in a natural way. However, the eigenvalues of F are important as they have the same distribution as the eigenvalues of $S_1S_2^{-1}$ where S_1 and S_2 are independent Wishart matrices with $S_i \sim W_p(n_i, \Sigma_i)(i = 1, 2)$ and this particular distribution depends only on the eigenvalues of Δ, say $\delta_1, \ldots, \delta_p(\delta_1 \geq \cdots \geq \delta_p > 0)$, which may be interpreted as the eigenvalues of $\Sigma_1\Sigma_2^{-1}$. In this paper our concern is the problem of estimating $\delta_1, \ldots, \delta_p$ in a decision theoretic way. These eigenvalues are important in the problem of testing H: $\Sigma_1 = \Sigma_2$ against K: $\Sigma_1 \neq \Sigma_2$, as the power function of any invariant test statistic depends only on $\delta_1, \ldots, \delta_p$. For details, see Muirhead (1982) and Muirhead and Verathaworn (1985).

To give an alternative estimator of $\delta_1, \ldots, \delta_p$, we concentrate on estimating a matrix Δ based on the matrix F. As long as we restrict our attention to the orthogonally invariant estimators $\hat{\Delta}(F)$ of Δ, the eigenvalues of $\hat{\Delta}(F)$ may be taken as estimates of $\delta_1, \ldots, \delta_p$ and we may expect that the eigenvalues of good estimators of Δ perform well as estimates of $\delta_1, \ldots, \delta_p$. This approach was first considered by Muirhead and Verathaworn (1985). Subsequently, Dey (1988) considered the problem of estimating $\delta_1, \ldots, \delta_p$ directly under the squared error loss. In a different setup, Loh (1988) dealt with the problem of estimating the matrix $\Sigma_1\Sigma_2^{-1}$.

Here we employ two loss functions defined by $L_1(\hat{\Delta}, \Delta) = \mathrm{tr}(\hat{\Delta}\Delta^{-1}) - \log\det(\hat{\Delta}\Delta^{-1}) - p$ and $L_2(\hat{\Delta}, \Delta) = \mathrm{tr}(\hat{\Delta}\Delta^{-1} - I_p)^2$. The corresponding risk functions are given by $R_i(\hat{\Delta}, \Delta) = E_F[L_i(\hat{\Delta}, \Delta)]$, $i = 1, 2$. Following an approach similar to that by Haff (1982) in the problem of estimating a normal covariance matrix, Muirhead and Verathaworn (1985) developed an approximation to Bayes rule under the loss function L_1. Later, using an approximation to the risk R_1 (due to Muirhead and Verathaworn (1985)), Gupta and Krishnamoorthy (1987) and Dey (1989) found new estimators under the loss function L_1. Konno (1991) showed that, for the case $p = 2$, the orthogonally invariant estimator given in Gupta and Krishnamoorthy (1987) is minimax under the loss L_1. Leung and Muirhead (1988) obtained the best scalar multiple of F under the loss function L_2.

Our estimators have the form

$$\hat{\Delta}(F) = a[F + ut(u)I_p] \tag{1.2}$$

where a is a constant and $t(u)$ is an absolutely continuous, nonincreasing and nonnegative function of $u = 1/\mathrm{tr}(F^{-1})$, which is analogous to the empirical Bayes estimator of the normal covariance matrix by Haff (1980). Note that our estimators are orthogonally invariant estimators of Δ, so that their eigenvalues may be considered as estimates of $\delta_1, \ldots, \delta_p$. We show that the proposed estimators dominate the best scalar multiple of F under L_1 and L_2 losses.

But the modification of estimators as done in this paper is not the best one since the proposed estimators move all eigenvalues of the best scalar multiple in the same direction. To remedy this shortcoming, one could apply the method of the Wishart matrix (Perron (1989)), which makes use of a doubly stochastic matrix to estimation of the F matrix.

In Section 2 of this paper, we state the integration by parts formula, called an F identity, to help in our risk computations. In Sections 3 and 4, sufficient conditions are given under which our estimators dominate the best scalar multiple of F under the loss functions L_1 and L_2.

Finally, in Section 5, numerical studies are carried out to indicate percentage improvements in average loss for our proposed estimators compared with the unbiased estimator, which is the best scalar multiple of F under the L_1 loss.

2. Preliminaries: the F identity

For $Q = (q_{ij})$ and a constant r, put $Q_{(r)} = rQ + (1-r)\mathrm{diag}(Q)$ where $\mathrm{diag}(Q)$ is a diagonal matrix with diagonal elements equal to those of Q. Let $D_{(p\times p)} = (d_{ij})$ be a matrix of differential operators given by $((1/2)(1+\delta_{ij})\partial/\partial f_{ij})$ for the Kronecker delta δ_{ij} and $F = (f_{ij})$. We define $DQ = \left(\sum_{k=1}^{p} d_{ik}q_{kj}\right)$ as a formal product followed by differentiation at the component level and $(\partial/\partial F)h(F) = (\partial h(F)/\partial f_{ij})$ for a real-valued function $h(F)$. Now we state the F identity which follows directly from (5) in Muirhead and Verathaworn (1985). This is useful to evaluate the risk difference between the usual estimators and our proposed estimators.

LEMMA 2.1. *Let F follow an $F_p(n_1, n_2;\ \Delta)$ distribution defined by* (1.1). *Under certain regularity conditions, we have*

$$\text{(2.1)}\quad E[h(F)\mathrm{tr}(\Delta+F)^{-1}V] = E\left[\frac{2}{n}h(F)\mathrm{tr}(DV) + \frac{2}{n}\mathrm{tr}\left(\frac{\partial h(F)}{\partial F}V_{(1/2)}\right) + \frac{n_1-p-1}{n}h(F)\mathrm{tr}(F^{-1}V)\right]$$

for a matrix $V = (v_{ij}(F,\ \Delta))$, a scalar $h(F)$ and $n = n_1 + n_2$.

3. Improved estimators under L_1 loss

Muirhead and Verathaworn (1985) showed that, for the L_1 loss, the unbiased estimate $\hat{\Delta}_U(F) = a_1F$ (being $a_1 = (n_2-p-1)/n_1$) is the best among the estimators aF where a is a constant.

Our main results in this section concern the estimators $\hat{\Delta}_1 = a_1(F + ut(u)I_p)$. For the purpose of proving that $\hat{\Delta}_1$ dominates $\hat{\Delta}_U$ under certain conditions on t, we need the following lemma.

LEMMA 3.1. *Let F have an $F_p(n_1, n_2;\ \Delta)$ distribution. Then we have the inequality*

$$E\left[\frac{t(u)\mathrm{tr}\Delta^{-1}}{\mathrm{tr}F^{-1}}\right] \le E\left[\frac{t(u)(n_1-p+1)}{n_2-2}\right],$$

which holds iff $p = 1$ and $t(u)$ is a constant.

PROOF. Put $h(F) = t(u)/\mathrm{tr}(F^{-1})$ and $V(\Delta,\ F) = (\Delta+F)\Delta^{-1}$ in (2.1). Then, noting that $\mathrm{tr}(DF\Delta^{-1}) = ((p+1)/2)\mathrm{tr}\Delta^{-1}$ (see Haff (1979)), the first term on the r.h.s. of (2.1) is equal to $[(p+1)/n]E[t(u)\mathrm{tr}\Delta^{-1}/\mathrm{tr}(F^{-1})]$. Using $(\partial/\partial F)\mathrm{tr}F^{-1} = -F_{(2)}^{-2}$ (see Haff (1980)), we get

$$\frac{\partial}{\partial F}h(F) = \frac{t(u)F_{(2)}^{-2}}{(\mathrm{tr}F^{-1})^2} + \frac{t'(u)F_{(2)}^{-2}}{(\mathrm{tr}F^{-1})^3}.$$

From these and the equation $\mathrm{tr}A_{(r)}B_{(1/r)} = \mathrm{tr}AB$ for any $p \times p$ matrices A and B, direct calculations show that (2.1) provides

$$(3.1)\quad n_2 E\left[\frac{t(u)\mathrm{tr}\Delta^{-1}}{\mathrm{tr}F^{-1}}\right] = E\left[\frac{2t(u)\mathrm{tr}(F^{-1}\Delta^{-1} + F^{-2})}{(\mathrm{tr}F^{-1})^2} + \frac{2t'(u)\mathrm{tr}(F^{-2} + F^{-1}\Delta^{-1})}{(\mathrm{tr}F^{-1})^3} + t(u)(n_1 - p - 1)\right].$$

Note that $t(u) \geq 0$. Applying $\mathrm{tr}(F^{-1}\Delta^{-1}) \leq (\mathrm{tr}F^{-1})(\mathrm{tr}\Delta^{-1})$ and $\mathrm{tr}F^{-2}/(\mathrm{tr}F^{-1})^2 \leq 1$ to the first term on the r.h.s. of (3.1) and noting that the second term on the r.h.s. of (3.1) is less than zero because $t'(u) \leq 0$, we get the desired result.

THEOREM 3.1. *For $p \geq 2$, the estimators of the form (1.2) given by $a = (n_2 - p - 1)/n_1$ and $t(u)$, an absolutely continuous and nonincreasing function bounded by*

$$(3.2)\qquad 0 \leq t(u) \leq \frac{2(p-1)(n_1 + n_2 - p - 1)}{n_1(n_2 - 2)},$$

beat the unbiased estimator $\hat{\Delta}_U$ under the L_1 loss.

PROOF. Put $\alpha_1(\Delta) = R_1(\hat{\Delta}_1, \Delta) - R_1(\hat{\Delta}_U, \Delta)$. Noting that $\log|I + A| \geq \mathrm{tr}A - (1/2)\mathrm{tr}A^2$ for any positive definite matrix A, the condition for $\alpha_1(\Delta) \leq 0$ may be written as $E[(1/2)t^2(u) - t(u) + (a_1 t(u)\mathrm{tr}\Delta^{-1}/\mathrm{tr}F^{-1})] \leq 0$. Using Lemma 3.1 it is seen that the condition (3.2) is sufficient for $\alpha_1(\Delta) \leq 0$.

Remark 3.1. Since $S = n_2 F$ converges weakly to the Wishart distribution $W_p(n_1, \Delta)$ as n_2 tends to infinity, the estimators in Theorem 3.1 with $t^*(u) = t(n_2 u)$ become reduced to the estimators of the covariance matrix Δ given by $\hat{\Delta}_1 = (1/n_1)(S + ut^*(u)I_p)$ where $t^*(u)$ is an absolutely continuous and nonincreasing function of $u = 1/\mathrm{tr}S^{-1}$ bounded by $0 \leq t^*(u) \leq 2(p-1)/n_1$. Theorem 3.1 implies that $\hat{\Delta}_1$ dominates $\hat{\Delta}_U = S/n_1$, which was the result obtained by Haff (1980).

4. Improved estimators under L_2 loss

It is shown in Leung and Muirhead (1988) that, for the L_2 loss and $n_2 > p+3$, the best estimator of the form aF is given by $\hat{\Delta}_B = a_2 F$ where

$$(4.1)\qquad a_2 = \frac{(n_2 - p)(n_2 - p - 3)}{(n_1 + p + 1)(n_2 - p - 1) + pn_1 + 2}.$$

Assume that $t(u)$ in (1.2) is a constant so that we lack the generality of t under the L_2 loss. Now our goal is to find a sufficient condition under which the estimators $\hat{\Delta}_2 = a_2(F + utI_p)$, where $u = 1/\mathrm{tr}F^{-1}$, dominate $\hat{\Delta}_B$ under the L_2 loss. Put

$$(4.2)\quad \alpha_2(\Delta) = R_2(\hat{\Delta}_2, \Delta) - R_2(\hat{\Delta}_B, \Delta) = 2a_2^2 tE\left[\frac{\mathrm{tr}(F\Delta^{-2})}{\mathrm{tr}F^{-1}}\right] - 2a_2 tE\left[\frac{\mathrm{tr}\Delta^{-1}}{\mathrm{tr}F^{-1}}\right] + a_2^2 t^2 E\left[\frac{\mathrm{tr}\Delta^{-2}}{(\mathrm{tr}F^{-1})^2}\right].$$

To evaluate (4.2), we need the following lemma given by the application of the identity (2.1). Their proofs add little insight into the problem. Therefore, the proofs have been put into the Appendix.

LEMMA 4.1. *Let F have an $F_p(n_1, n_2; \Delta)$ distribution with $n_2 > p+3$. Then the following inequalities hold:*

$$\text{(i)} \qquad E\left[\frac{\mathrm{tr}\Delta^{-2}}{(\mathrm{tr}F^{-1})^2}\right] \leq \frac{(n_1 - p_1 + 1)(n_1 - p + 3)}{(n_2 - 2)(n_2 - 4)} E\left[\frac{\mathrm{tr}F^{-2}}{(\mathrm{tr}F^{-1})^2}\right],$$

$$\text{(ii)} \qquad E\left[\frac{\mathrm{tr}\Delta^{-1}}{\mathrm{tr}F^{-1}}\right] \geq \frac{2}{n_2} E\left[\frac{\mathrm{tr}F^{-2}}{(\mathrm{tr}F^{-1})^2}\right] + \frac{n_1 - p - 1 + 2\epsilon}{n_2}$$

where

$$\epsilon = \frac{p(n_1 - p - 1) + 2}{p^2(pn_2 - 2)}, \tag{4.3}$$

$$\begin{aligned}\text{(iii)} \quad E\left[\frac{\mathrm{tr}(F\Delta^{-2})}{\mathrm{tr}F^{-1}}\right] \leq{} & \frac{2}{(n_2 - p - 1)(n_2 - 2)} \\ & \cdot\left[\frac{(n_1 + n_2)(n_1 - p + 1) + n_1(n_2 - 2)}{n_2} \right.\\ & \left. + \frac{(n_1 - p + 1)(n_1 - p + 3)}{n_2 - 4}\right] \\ & \cdot E\left[\frac{\mathrm{tr}F^{-2}}{(\mathrm{tr}F^{-1})^2}\right] + \frac{n_1(n_1 - p - 1)}{n_2(n_2 - p - 1)}.\end{aligned}$$

THEOREM 4.1. *Assume that $n_2 > p + 3$ and $p \geq 2$. Let*

$$\begin{aligned}\beta = {} & \frac{2(n_2 - 2)(n_2 - 4)}{(n_1 - p + 1)(n_1 - p + 3)}\left[\frac{n_1 - p + 1 + 2\epsilon}{a_2 n_2} - \frac{n_1(n_1 - p - 1)}{n_2(n_2 - p - 1)}\right. \\ & - \frac{2}{(n_2 - p - 1)(n_2 - 2)}\left\{\frac{(n_1 + n_2)(n_1 - p + 1) + n_1(n_2 - 2)}{n_2}\right. \\ & \left.\left. + \frac{(n_1 - p + 1)(n_1 - p + 3)}{n_2 - 4}\right\}\right]\end{aligned}$$

where a_2 and ϵ are defined by (4.1) and (4.3), respectively. If $\beta > 0$, then the estimators of the form (1.2) given by $a = a_2$ and $0 \leq t \leq \beta$ beat $\hat{\Delta}_B$ under the L_2 loss.

PROOF. From (ii) and (iii) of Lemma 4.1, the coefficient of t in (4.2) is bounded above by

$$\begin{aligned}& \left[\frac{4a_2^2}{(n_2 - p - 1)(n_2 - 2)}\left\{\frac{(n_1 + n_2)(n_1 - p + 1) + n_1(n_2 - 2)}{n_2}\right.\right. \\ & \left.\left. + \frac{(n_1 - p + 1)(n_1 - p + 3)}{n_2 - 4}\right\} - \frac{4a_2}{n_2}\right] E\left[\frac{\mathrm{tr}F^{-2}}{(\mathrm{tr}F^{-1})^2}\right] \\ & + \frac{2a_2}{n_2}\left\{\frac{a_2 n_1(n_1 - p - 1)}{n_2 - p - 1} - (n_1 - p - 1 + 2\epsilon)\right\}.\end{aligned} \tag{4.4}$$

Noting that $\epsilon \geq 0$ and that the term inside the second curly bracket of (4.4) is bounded by

$$(n_1 - p - 1)\left\{\frac{(n_2 - p - 1)^2 - (n_2 - p + 1)}{(n_2 - p - 1)^2} - 1\right\} < 0,$$

it is seen that (4.4) can be bounded above by

$$- \frac{a_2^2\beta(n_1 - p + 1)(n_1 - p + 3)}{(n_2 - 2)(n_2 - 4)} E\left[\frac{\mathrm{tr}F^{-2}}{(\mathrm{tr}F^{-1})^2}\right]. \tag{4.5}$$

Using (i) of Lemma 4.1 and (4.5), a straightforward calculation shows that the sufficient condition for $\alpha_2(\Delta) \leq 0$ becomes $t^2 - \beta t \leq 0$, which completes the proof.

Remark 4.1. Similar to Remark 3.1, we can see that Theorem 4.1 implies Theorem 4.6 in Haff (1980).

Unfortunately β is not always positive when $p = 2$. We carry out numerical calculations to see whether n_1 and n_2 satisfy $\beta > 0$. It indicates that β is monotonically decreasing in n_1 for each fixed n_2, which follows that β has just one sign change. Table 1 shows that, for example, β is not positive for $n_1 \geq 43$ when $n_2 = 15$. It also shows that the minimum of n_1 such that β does not take a positive value for each n_2 first goes down and then goes up as n_2 increases. When $p \geq 3$, our numerical calculation shows that β is always positive.

Table 1. For fixed n_2, minimum of n_1 under which β is not positive.

n_2	10	11	12	13	14	15	16	17	18	19	20	30	40	50
n_1	298	74	53	46	44	43	43	43	44	46	47	64	83	102

5. Numerical studies

In this section, we use the Monte Carlo simulation method to compute average loss and percentage reductions in average loss for $\hat{\Delta}_1$ (where $t(u) = (p-1)(n_1 + n_2 - p - 1)/\{n_1(n_2 - 2)\}$) compared with $\hat{\Delta}_U$ under the loss function

$$L(\Delta_e, \hat{\Delta}_e) = \mathrm{tr}(\hat{\Delta}_e\Delta_e^{-1}) - \log\det(\hat{\Delta}_e\Delta_e^{-1}) - p, \tag{5.1}$$

where for a matrix A, $A_e = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, λ_i's are the eigenvalues of A.

For $p = 4$ and $n_1, n_2 = 10, 25, 50$, a sample of 100 S_1's and 1000 S_2's is generated, where $S_i \sim W_4(n_i, I_4)$, $i = 1, 2$. For each (n_1, n_2), the 100 pairs (S_1, S_2) are transformed into $F = \Delta^{1/2}S_1^{1/2}S_2^{-1}S_1^{1/2}\Delta^{1/2}$, for each 3 choices of Δ.

The eigenvalues of F's are then obtained to form the estimates of the eigenvalues of Δ.

Table 2 shows average loss and percentage reductions in average loss for $\hat{\Delta}_1$ compared to $\hat{\Delta}_U$ under the loss (5.1) when $p = 4$. It indicates that for all choices of Δ, percentage reductions range between 5% and 20%.

Table 2. Average loss and improvements for $\hat{\Delta}_1$ over $\hat{\Delta}_U$ when $p = 4$.

	$n_1 = 10$ $n_2 = 10$	$n_1 = 25$ $n_2 = 25$	$n_1 = 50$ $n_2 = 50$
	$\Delta = \mathrm{diag}(1, 1, 1, 1)$		
$\hat{\Delta}_U$	2.78	.870	.403
$\hat{\Delta}_1$	2.46	.805	.383
	11.7%	7.36%	5.1%
	$\Delta = \mathrm{diag}(8, 4, 2, 1)$		
$\hat{\Delta}_U$	1.57	.350	.165
$\hat{\Delta}_1$	1.26	.292	.147
	19.6%	16.8%	11.3%
	$\Delta = \mathrm{diag}(10, 10, 1, 1)$		
$\hat{\Delta}_U$	2.78	.539	.268
$\hat{\Delta}_1$	2.46	.481	.250
	11.7%	10.7%	6.8%

Acknowledgements

The author wishes to thank Professor N. Sugiura for his valuable comments and suggestions. He is also indebted to the referees for their helpful comments and suggestions.

Appendix

PROOF OF LEMMA 4.1. (i) Take $h(F) = (\mathrm{tr}F^{-1})^{-2}$ and $V = (\Delta + F)\Delta^{-2}$ in the F identity (2.1). From similar calculations to those for the proof of Lemma 3.1, we may see that (2.1) provides

$$E\left[\frac{\mathrm{tr}\Delta^{-2}}{(\mathrm{tr}F^{-1})^2}\right] \leq \frac{n_1 - p + 3}{n_2 - 4} E\left[\frac{\mathrm{tr}(F^{-1}\Delta^{-1})}{(\mathrm{tr}F^{-1})^2}\right]. \tag{A.1}$$

Hence it suffices to show that

$$E\left[\frac{\mathrm{tr}(F^{-1}\Delta^{-1})}{(\mathrm{tr}F^{-1})^2}\right] \leq \frac{n_1 - p + 1}{n_2 - 2} E\left[\frac{\mathrm{tr}F^{-2}}{(\mathrm{tr}F^{-1})^2}\right]. \tag{A.2}$$

Set $G = F^{-1}$. Then it is seen that G follows an $F_p(n_2, n_1; \Delta^{-1})$ distribution. Applying Lemma 2.1 (being $h(G) = (\mathrm{tr}G)^{-2}$ and $V(G, \Delta) = (\Delta^{-1} + G)G^2$ in (2.1)) with a distribution of G instead of F and noting that $\mathrm{tr}D\Delta^{-1}G^2 = ((p+2)/2)\mathrm{tr}(\Delta^{-1}G) + (1/2)(\mathrm{tr}\Delta^{-1})(\mathrm{tr}G)$ and $\mathrm{tr}G^3 = ((2p+3)/2)\mathrm{tr}G^2 + (1/2)(\mathrm{tr}G)^2$, a similar argument leads to (A.2), which completes the proof of (i).

(ii) Let $t(u)$ be a constant in (3.1). Then the remainder of the proof is to evaluate $E[\mathrm{tr}(F^{-1}\Delta^{-1})/(\mathrm{tr}F^{-1})^2]$. Using the fact that $p\mathrm{tr}F^{-2} \geq (\mathrm{tr}F^{-1})^2$ and making the transformation $T = \Delta^{-1/2}F\Delta^{-1/2}$, we can see that the term is bounded below by

$$\text{(A.3)} \qquad \frac{1}{p}E\left[\frac{\mathrm{tr}(T^{-1}\Delta^{-2})}{\mathrm{tr}(T^{-1}\Delta^{-1})^2}\right] \geq \frac{1}{p}E\left[\frac{1}{\mathrm{tr}T^{-1}}\right].$$

Noting that T^{-1} has an $F_p(n_2, n_1; I_p)$ distribution and using Lemma 3.4 in Leung and Muirhead (1987), the r.h.s. of (A.3) is bounded below by $\{p(n_1 - p - 1) + 2\}/\{p^2(pn_2 - 2)\}$, which completes the proof of (ii).

(iii) Set $h(F) = (\mathrm{tr}F^{-1})^{-1}$ and $V(F, \Delta) = (\Delta + F)\Delta^{-2}F$ in (2.1). From the equality $\mathrm{tr}(DF\Delta^{-2}F) = (p+1)\mathrm{tr}(\Delta^{-2}F)$ (see Konno (1988)) and similar argument in the proof of Lemma 3.1, we may see that (2.1) gives

$$\text{(A.4)} \quad (n_2 - p - 1)E\left[\frac{\mathrm{tr}(F\Delta^{-2})}{\mathrm{tr}F^{-1}}\right] = n_1 E\left[\frac{\mathrm{tr}\Delta^{-1}}{\mathrm{tr}F^{-1}}\right] + 2E\left[\frac{\mathrm{tr}(\Delta^{-1}F^{-1} + \Delta^{-2})}{(\mathrm{tr}F^{-1})^2}\right].$$

First put (3.1) ($t(u)$) being a constant) into the first term on the r.h.s. of (A.3) and use (A.1) and (A.2). Then we may get the desired result.

REFERENCES

Dey, D. K. (1988). Simultaneous estimation of eigenvalues, *Ann. Inst. Statist. Math.*, **40**, 137–147.

Dey, D. K. (1989). On estimation of the scale matrix of the multivariate F distribution, *Comm. Statist. A—Theory Methods*, **18**, 1373–1383.

Gupta, A. K. and Krishnamoorthy, K. (1987). Improved estimators of eigenvalues of $\Sigma_1\Sigma_2^{-1}$, Tech. Report No. 87-11, Bowling Green State University, Ohio.

Haff, L. R. (1979). An identity for the Wishart distribution with applications, *J. Multivariate Anal.*, **9**, 531–544.

Haff, L. R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix, *Ann. Statist.*, **8**, 586–597.

Haff, L. R. (1982). Solutions of the Euler-Lagrange equations for certain multivariate normal estimation problem (unpublished manuscript).

Konno, Y. (1988). Exact moments of the multivariate F and beta distributions, *J. Japan Statist. Soc.*, **18**, 123–130.

Konno, Y. (1991). On estimating eigenvalues of the scale matrix of the multivariate F distribution, *Sankhyā Ser. A*, **53** (to appear).

Leung, P. L. and Muirhead, R. J. (1987). Estimation of parameter matrices and eigenvalues in MANOVA and canonical correlation analysis, *Ann. Statist.*, **15**, 1651–1666.

Leung, P. L. and Muirhead, R. J. (1988). On estimating characteristic roots in a two-sample problem, *Statistical Decision Theory and Related Topics IV* (eds. S. S. Gupta and J. O. Berger), 397–402, Springer, New York.

Loh, W.-L. (1988). Estimating covariance matrices, Ph. D. Thesis, Stanford University.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.

Muirhead, R. J. and Verathaworn, T. (1985). On estimating the latent roots of $\Sigma_1\Sigma_2^{-1}$, *Multivariate Analysis—VI* (ed. P. R. Krishnaiah), 431–447, North-Holland, Amsterdam.

Perron, F. (1989). Minimax estimators of a covariance matrix, Rapports de recherche du départment de mathématiques et de statistiques, D.M.S. No. 89-23, Université de Montréal.