The Korean Communications
in Statistics Vol. 10 No. 2, 2003
pp. 603-606

Note on Estimating the Eigen System of $\Sigma_1^{-1}\Sigma_2$

Myung Geun Kim[1)]

Abstract

The maximum likelihood estimators of the eigenvalues and eigenvectors of $\Sigma_1^{-1}\Sigma_2$ are shown to be the eigenvalues and eigenvectors of $S_1^{-1}S_2$ under multivariate normality and are explicitly derived. The nature of the eigenvalues and eigenvectors of $\Sigma_1^{-1}\Sigma_2$ or their estimators will be uncovered.

Keywords: Eigenvalue, eigenvector, maximum likelihood estimator.

1. Introduction

Let X_i be a p-variate random vector having a normal distribution $N(\mu_i, \Sigma_i)$ with mean vector μ_i and positive definite covariance matrix Σ_i for $i=1,2$. The maximum likelihood estimator of Σ_i based on a sample of size n_i drawn from $N(\mu_i, \Sigma_i)$ is denoted by S_i. The sample covariance matrix S_i is positive definite with probability one if and only if $n_i > p$ (Dykstra, 1970), which will be assumed throughout. In any multivariate textbooks, no explicit derivation of the maximum likelihood estimators of the eigenvalues and eigenvectors of $\Sigma_1^{-1}\Sigma_2$ is made. The eigenvalues and eigenvectors of $S_1^{-1}S_2$ are implicitly used as the maximum likelihood estimators of the eigenvalues and eigenvectors of $\Sigma_1^{-1}\Sigma_2$. It may be due to the invariance property of the maximum likelihood estimator. However, a routine use of the eigenvalues and eigenvectors of $S_1^{-1}S_2$ does not give any idea about the nature of the eigenvalues and eigenvectors of $\Sigma_1^{-1}\Sigma_2$. The eigenvalues of $S_1^{-1}S_2$ are important, for example, in testing the hypothesis $\Sigma_1 = \Sigma_2$ (Muirhead, 1982, Section 8.2). In light of decision theory, Muirhead and Verathaworn (1985) considered an estimation of the eigenvalues of

1) Professor, Department of Applied Statistics, Seowon University, Cheongju, 361-742.
E-mail: mgkim@seowon.ac.kr

$\Sigma_1^{-1}\Sigma_2$.

The purpose of this note is to show that the eigenvalues and eigenvectors of $S_1^{-1}S_2$ are the normal theory maximum likelihood estimators of those of $\Sigma_1^{-1}\Sigma_2$ and is to find explicit forms of the maximum likelihood estimators. Further, we will uncover the nature of the eigenvalues and eigenvectors of $\Sigma_1^{-1}\Sigma_2$ or their estimators.

2. Eigen system of $\Sigma_1^{-1}\Sigma_2$

2.1 Usual derivation

For easy reference and the clarity of our argument, we consider the usual method for finding the eigenvalues and eigenvectors of $S_1^{-1}S_2$. Let b be the eigenvector of $S_1^{-1}S_2$ associated with the eigenvalue λ. Then $S_2b=\lambda S_1b$. We can write $S_2=MM^T$ where M is a nonsingular matrix that can be chosen by using the spectral decomposition theorem or the Cholesky decomposition theorem. Define a vector a by $b=S_1^{-1}Ma$. This gives $(M^TS_1^{-1}M)a=\lambda a$. Then λ and b can be found. Usually the eigenvector b is normalized such that $b^TS_1b=1$. To this end, a should satisfy $a^Ta=1/\lambda$. This method is based on pure matrix algebra and does not give any idea about the nature of λ and b.

2.2 Maximum likelihood estiamtors

First, we state a well-known fact about a diagonalization of two positive definite covariance matrices (Muirhead, 1982, p.592) before considering the maximum likelihood estimation.

Lemma 1. *Two positive definite symmetric matrices Σ_1 and Σ_2 are always diagonalized by a nonsingular matrix B as follows*

$$B^T\Sigma_1B=I_p \quad \text{and} \quad B^T\Sigma_2B=\Lambda \tag{1}$$

where Λ is a positive definite diagonal matrix of dimension p and I_p is the identity matrix of dimension p.

In Lemma 1 B is uniquely determined up to sign changes if the diagonal elements of Λ are distinct. Lemma 1 implies that each column vector of B normalized with respect to Σ_1 is the eigenvector of $\Sigma_1^{-1}\Sigma_2$ and its associated eigenvalue is the corresponding diagonal element of Λ.

If we define A by $A=(B^{-1})^T$, then Lemma 1 gives $\Sigma_1=AA^T$ and $\Sigma_2=A\Lambda A^T$. The i-th column of B and the i-th diagonal element of Λ are denoted by b_i and λ_i, respectively and the hat notation indicates the corresponding maximum likelihood estimator. Let $r_i=n_i/n_+$ and $n_+=n_1+n_2$. Then the likelihood equations under the reparametrization (1) are easily computed as

$$I_p=diag(\hat{B}^TS_1\hat{B}) \quad \text{and} \quad \hat{\Lambda}=diag(\hat{B}^TS_2\hat{B}) \tag{2}$$

$$\hat{B}^T(r_1S_1)\hat{B}+\hat{B}^T(r_2S_2)\hat{B}\hat{\Lambda}^{-1}=I_p \tag{3}$$

Next we will show that $\hat{B}^TS_1\hat{B}$ and $\hat{B}^TS_2\hat{B}$ become diagonal matrices. To this end suppose that $\hat{b}_i^T(r_1S_1)\hat{b}_j=x$ for a fixed pair $1 \le i \neq j \le p$, where x is assumed to be nonzero. Then the symmetry of $\hat{B}^T(r_1S_1)\hat{B}$ gives $\hat{b}_j^T(r_1S_1)\hat{b}_i=x$. If we compare the (i,j)th and (j,i)th elements of both sides of equation (3) and use the symmetry of $\hat{B}^T(r_2S_2)\hat{B}$, then we have

$$\hat{b}_i^T(r_2S_2)\hat{b}_j=-\hat{\lambda}_ix=-\hat{\lambda}_jx.$$

Hence we have $\hat{\lambda}_i=\hat{\lambda}_j$. Since the probability that $\hat{\lambda}_i=\hat{\lambda}_j$ is zero (Okamoto, 1973), we have $x=0$. Thus the likelihood equations (2) and (3) reduce to

$$\hat{B}^TS_1\hat{B}=I_p \quad \text{and} \quad \hat{B}^TS_2\hat{B}=\hat{\Lambda} \tag{4}$$

which is just Lemma 1 with unknown parameters replaced by their respective maximum likelihood estimators. Thus we see that the eigenvalues and eigenvectors of $S_1^{-1}S_2$ are the maximum likelihood estimators of those of $\Sigma_1^{-1}\Sigma_2$. Note that the maximum likelihood estimators $\hat{\Lambda}$ and $\hat{B}$ are equivariant estimators under the group of affine transformations. The procedure above does not depend on group labelling.

2.3 Explicit solutions

We will find explicit forms of the maximum likelihood estimators $\hat{\Lambda}$ and $\hat{B}$. The spectral decomposition theorem gives an expression $S_+=S_1+S_2=ULU^T$, where U is an orthogonal matrix of the eigenvectors of S_+ and L is a positive definite diagonal matrix of the eigenvalues. Let $\Psi = L^{-1/2}U^TS_2UL^{-1/2}$. By the spectral decomposition theorem, we have $\Psi = VGV^T$, where V and G should be interpreted as usual. Since Ψ is positive definite, the diagonal elements of G are all positive. Note that $|\Psi - \delta I_p| = 0$ if and only if $|S_+^{-1}S_2 - \delta I_p| = 0$. Hence Ψ and $S_+^{-1}S_2$ have the same eigenvalues δ so that $S_1^{-1}S_2$ has $\lambda = \delta/(1-\delta)$ as its eigenvalues. Since S_1 and S_2 are positive definite, we have $tr(S_1^{-1}S_2) > 0$. Hence the largest eigenvalue λ_1 of $S_1^{-1}S_2$ is positive. Since $\delta = \lambda/(1+\lambda)$ is a strictly increasing function of λ, the largest eigenvalue δ_1 of $S_+^{-1}S_2$ satisfies $0 < \delta_1 < 1$. Since $|S_1 + S_2| = |L| > |S_1| = |L||G|$, it is clear that $0 < |G| < 1$. Thus the diagonal elements of G are positive and less than 1, which implies that G and $I_p - G$ are positive definite diagonal matrices. Therefore $\hat{B} = UL^{-1/2}V(I_p - G)^{-1/2}$ is a nonsingular matrix and $\hat{\Lambda} = G(I_p - G)^{-1}$ is a positive definite diagonal matrix, and they solve the likelihood equations (4).

References

[1] Dykstra, R.L. (1970) Establishing the positive definiteness of the sample covariance matrix, *Annals of Mathematical Statistics*, **41**, 2153-2154.

[2] Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc.

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

[4] Okamoto, M. (1973) Distinctness of the eigenvalues of a quadratic form in a multivariate sample, *Annals of Statistics*, **1**, 763-765.

[Received May 2003, Accepted August 2003]