DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 7 | Issue 1 Article 16

5-1-2008

Estimation of Covariance Matrix in Signal Processing When the Noise Covariance Matrix is Arbitrary

Madhusudan Bhandary
Columbus State University, bhandary_madhusudan@columbusstate.edu

Journal of Modern Applied Statistical Methods
May, 2008, Vol.7, No. 1, 198-204

1538 – 9472/08/$95.00

Estimation of Covariance Matrix in Signal Processing When the Noise Covariance Matrix is Arbitrary

Madhusudan Bhandary
Columbus State University

An estimator of the covariance matrix in signal processing is derived when the noise covariance matrix is arbitrary based on the method of maximum likelihood estimation. The estimator is a continuous function of the eigenvalues and eigenvectors of the matrix $\hat{\Sigma}_1^{-\frac{1}{2}} S^* \hat{\Sigma}_1^{-\frac{1}{2}}$, where S^* is the sample covariance matrix of observations consisting of both noise and signals and $\hat{\Sigma}_1$ is the estimator of covariance matrix based on observations consisting of noise only. Strong consistency and asymptotic normality of the estimator are briefly discussed.

Key words: Maximum likelihood estimator, signal processing, white noise, colored noise.

Introduction

The covariance and correlation matrices are used for a variety of purposes. They give a simple description of the overall shape of a point-cloud in p-space. They are used in principal component analysis, factor analysis, discriminant analysis, canonical correlation analysis, tests of independence etc. In signal processing, estimation of covariance matrix is important because it helps to discriminate between signals and noise (filtering).

The problem of estimation of the dispersion matrix of the form $\Gamma + \sigma^2\Sigma_1$ is considered, where the unknown matrix Γ is n.n.d. of rank q(< p), σ^2 (> 0) is unknown and Σ_1 is some arbitrary positive matrix. In general, the model is signal processing is

$$\mathbf{X}(t) = AS(t) + \mathbf{n}(t) \qquad (1.1)$$

where, $\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_p(t))'$ is the px1 observation vector at time t, $\mathbf{S}(t) = (S_1(t), S_2(t), \ldots, S_q(t))'$ is the qx1 vector of unknown random signals at time t, $\mathbf{n}(t) = (n_1(t), n_2(t), \ldots, n_p(t))'$ is the px1 random noise vector at time t, and $A = (A(\Phi_1), A(\Phi_2), \ldots, A(\Phi_q))$ is the pxq matrix of unknown coefficients, $A(\Phi_r)$ is the px1 vector of functions of the elements of unknown vector Φ_r associated with the r^{th} signal and q < p.

In model (1.1), $\mathbf{X}$(t) is assumed to be distributed as p-variate normal distribution with mean vector zero and dispersion matrix $A\Psi A' + \sigma^2\Sigma_1 = \Gamma + \sigma^2\Sigma_1$, where $\Gamma = A\Psi A'$ is unknown n.n.d. matrix of rank q(<p) and Ψ = covariance matrix of $\mathbf{S}$(t), σ^2 (>0) is unknown, $\sigma^2\Sigma_1$ is the covariance matrix of the noise vector $\mathbf{n}$(t) and Σ_1 is some arbitrary positive definite matrix. In the above situation, when the covariance matrix of the noise vector $\mathbf{n}$(t) is $\sigma^2 I_p$, where I_p denotes identity matrix of order pxp, the model is called white noise model. If the covariance matrix of $\mathbf{n}$(t) is $\sigma^2\Sigma_1$, where Σ_1 is some arbitrary positive definite matrix, the model is colored noise model.

One of the important problems that arise in the area of signal processing is to estimate q, the number of signals transmitted. The problem

Madhusudan Bhandary is on the faculty of the Department of Mathematics. Mailing address: Columbus State University, 4225 University Avenue, Columbus, GA 31907. E-mail: bhandary_madhusudan@colstate.edu.

is equivalent to estimate the multiplicity of the smallest eigen value of the covariance matrix of the observation vector. Anderson (1963), Krishnaiah (1976), Rao (1983), Wax and Kailath (1984), Zhao et.al (1986a,b) considered the above problem. Chen (2001), Chen (2002) and Kundu (2000) developed procedures for estimating the number of signals.

Another important problem in this area is to have some idea about covariance and correlation matrix. The estimation of the dispersion matrix of the form $\Gamma + \sigma^2\Sigma_1$ is of interest, and then, the derivation of the estimator is discussed. Strong consistency and asymptotic normality of the estimator are then discussed.

Derivation of the Estimator

Let the observations **x**(t_1), **x**(t_2), …, **x**(t_n) be n observed p-component signals at n different time points which are independently and identically distributed as p-variate normal distribution with mean vector zero and dispersion matrix $\Gamma + \sigma^2\Sigma_1$, where $\Gamma = A\Psi A'$ and is n.n.d. of rank q(<p) and Σ_1 is some arbitrary positive definite matrix.

Because Γ is n.n.d. of rank q(<p), it can be assumed that $\Gamma = BB'$, where B is a pxq matrix of rank q and

$$B'B = Diag.(\theta_1, \theta_2, ..., \theta_q), \qquad (2.1)$$

where $\theta_1 \geq \theta_2 \geq ... \geq \theta_q$ are the non-zero eigen values of Γ.

The log-likelihood of the observations based on $\mathbf{x_i}$'s, apart from a constant term, can be written as follows :

$$\log L = -\frac{n}{2}\log\left|BB' + \sigma^2\Sigma_1\right|$$

$$-\frac{1}{2}tr.(BB' + \sigma^2\Sigma_1)^{-1}S \qquad (2.2)$$

where, S= $\sum_{i=1}^{n} x_i x_i', x_i = x(t_i), i = 1,2,...,n$

Following Lawley and Maxel (1963, Chapter 2):

$$\frac{\partial \log L}{\partial B} =$$

$$\left[-\frac{n}{2}(BB' + \sigma^2\Sigma_1)^{-1} + \frac{1}{2}(BB' + \sigma^2\Sigma_1)^{-1}S(BB' + \sigma^2\Sigma_1)^{-1}\right]$$

$$2B = 0$$

$$\text{i.e. } \Sigma_2^{-1}(\Sigma_2 - S^*)\Sigma_2^{-1}B = 0 \qquad (2.3)$$

where, $\Sigma_2 = BB' + \sigma^2\Sigma_1$ and $S^* = \frac{S}{n}$.

Using Rao(1983, p.33)

$$\Sigma_2^{-1} = (BB' + \sigma^2\Sigma_1)^{-1} =$$

$$\left(\frac{\Sigma_1^{-1}}{\sigma^2} - \frac{\Sigma_1^{-1}}{\sigma^2}B\left(\frac{B'\Sigma_1^{-1}B}{\sigma^2} + I_q\right)^{-1}\frac{B'\Sigma_1^{-1}}{\sigma^2}\right) =$$

$$\frac{1}{\sigma^2}\left(\Sigma_1^{-1} - \Sigma_1^{-1}B(I_q + D)^{-1}\frac{B'\Sigma_1^{-1}}{\sigma^2}\right) \qquad (2.4)$$

where, $D = \frac{B'\Sigma_1^{-1}B}{\sigma^2}$ and I_p denotes identity matrix of order pxp. Using (2.4) in (2.3),

$$(\Sigma_2 - S^*)\frac{1}{\sigma^2}\left[\Sigma_1^{-1} - \Sigma_1^{-1}B(I_q + D)^{-1}\frac{B'\Sigma_1^{-1}}{\sigma^2}\right]B = 0$$

$$\text{i.e. } (\Sigma_2 - S^*)\frac{\Sigma_1^{-1}B}{\sigma^2}\left[I_q - (I_q + D)^{-1}D\right] = 0$$

$$\text{i.e. } (\Sigma_2 - S^*)\frac{\Sigma_1^{-1}B}{\sigma^2}(I_q + D)^{-1} = 0$$

$$\text{i.e. } (\Sigma_2 - S^*)\Sigma_1^{-1}B = 0 \qquad (2.5)$$

which after substitution of Σ_2 from (2.3) and rearrangement of terms gives

$$S^*\Sigma_1^{-1}B = B(\sigma^2 I_q + B'\Sigma_1^{-1}B)$$

$$\text{i.e. } (\Sigma_1^{-\frac{1}{2}}S^*\Sigma_1^{-\frac{1}{2}})(\Sigma_1^{-\frac{1}{2}}B) =$$

$$(\Sigma_1^{-\frac{1}{2}}B)(\sigma^2 I_q + B'\Sigma_1^{-1}B) \qquad (2.6)$$

It can be seen that the right hand side of (2.2) remains the same, if the matrix B is replaced by BP where P is an orthogonal matrix and hence $B'\Sigma_1^{-1}B$ can be reduced to $P'B'\Sigma_1^{-1}BP$ which can be reduced to a diagonal form because $B'\Sigma_1^{-1}B$ is a real symmetric matrix (See Bellman (1960) p.54).

From (2.6) it is trivial that columns of $\Sigma_1^{-\frac{1}{2}}B$ are eigenvectors of the matrix $\Sigma_1^{-\frac{1}{2}}S^*\Sigma_1^{-\frac{1}{2}}$ and the diagonal elements of $\sigma^2 I_q + B'\Sigma_1^{-1}B$ are the corresponding eigenvalues (2.7).

Let $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_p$ be the ordered eigen values of $\Sigma_1^{-\frac{1}{2}}S^*\Sigma_1^{-\frac{1}{2}}$ and let $\Theta = Diag.(\alpha_1, \alpha_2, ..., \alpha_q)$. Since the diagonal elements of $B'\Sigma_1^{-1}B$ are the column sum of squares of $\Sigma_1^{-\frac{1}{2}}B$, each eigenvector should be normalized so that the sum of squares equal the corresponding eigenvalue minus σ^2. Let $\widetilde{B}$ be a pxq matrix whose columns are $w_1, w_2, ..., w_q$, where $w_1, w_2, ..., w_q$ are a set of unit-length eigen vectors corresponding to the q largest eigen values of $\Sigma_1^{-\frac{1}{2}}S^*\Sigma_1^{-\frac{1}{2}}$. Then,

$$\widetilde{B}'\widetilde{B} = I_q$$

and

$$\Sigma_1^{-\frac{1}{2}}\hat{B} = \widetilde{B}(\Theta - \sigma^2 I_q)^{\frac{1}{2}} \qquad (2.8)$$

Another likelihood equation can be written as follows:

$$\frac{\partial \log L}{\partial \sigma^2} =$$

$$tr.(\Sigma_2^{-1}(\Sigma_2 - S^*)\Sigma_2^{-1}\Sigma_1) = 0 \qquad (2.9)$$

From (2.4) and (2.9),

$$tr.\left[(I_p - \Sigma_2^{-1}S^*)(\frac{1}{\sigma^2}(I_p - \Sigma_1^{-1}B(I_q + D)^{-1}\frac{B'}{\sigma^2}))\right] = 0,$$

$$tr.\left[\begin{array}{l}\frac{1}{\sigma^2}(I_p - \Sigma_1^{-1}B(I_q + D)^{-1}\frac{B'}{\sigma^2}) - \\ \frac{\Sigma_2^{-1}S^*}{\sigma^2} + \frac{1}{\sigma^2}\Sigma_2^{-1}S^*\Sigma_1^{-1}B(I_q + D)^{-1}\frac{B'}{\sigma^2}\end{array}\right] = 0$$

$$tr.\left[\begin{array}{l}\frac{I_p}{\sigma^2} - \frac{\Sigma_1^{-1}B}{\sigma^2}(I_q + D)^{-1}\frac{B'}{\sigma^2} - \\ \frac{\Sigma_2^{-1}S^*}{\sigma^2} + \frac{\Sigma_1^{-1}B}{\sigma^2}(I_q + D)^{-1}\frac{B'}{\sigma^2}\end{array}\right] = 0 \text{ (using (2.5))}$$

$$tr.\left[\frac{I_p}{\sigma^2} - \frac{\Sigma_2^{-1}S^*}{\sigma^2}\right] = 0$$

$$tr.\left[\frac{I_p}{\sigma^2} - \frac{1}{\sigma^2}\left\{\Sigma_1^{-1} - \Sigma_1^{-1}B(I_q + D)^{-1}\frac{B'\Sigma_1^{-1}}{\sigma^2}\right\}\frac{S^*}{\sigma^2}\right] = 0 \text{ (using (2.4))}$$

i.e.,

$$tr.\left[\frac{I_p}{\sigma^2} - \frac{\Sigma_1^{-1}S^*}{\sigma^4} + \frac{\Sigma_1^{-1}B(I_q + D)^{-1}B'\Sigma_1^{-1}S^*}{\sigma^6}\right] = 0$$

$$\frac{p}{\sigma^2} - \frac{tr.(\Sigma_1^{-\frac{1}{2}}S^*\Sigma_1^{-\frac{1}{2}})}{\sigma^4} + \frac{tr.(B'\Sigma_1^{-1}B)}{\sigma^4} = 0 \qquad (2.10)$$

(2.10) is obtained due to the fact that

$$\frac{\Sigma_1^{-1}B(I_q + D)^{-1}B'\Sigma_1^{-1}S^*}{\sigma^6} =$$

$$\frac{\Sigma_1^{-1}B(I_q + D)^{-1}B'\Sigma_1^{-1}\Sigma_2}{\sigma^6} \quad \text{(using (2.5))}$$

$$= \frac{\Sigma_1^{-1}B(I_q + D)^{-1}(I_q + D)\sigma^2 B'}{\sigma^6}$$

$$= \frac{\Sigma_1^{-1}BB'}{\sigma^4}$$

(because $B'\Sigma_1^{-1}\Sigma_2 = B'\Sigma_1^{-1}(BB' + \sigma^2\Sigma_1)$) $= B'\Sigma_1^{-1}BB' + \sigma^2 B' = (D + I_q)\sigma^2 B'$)

From (2.10),

$$\frac{p}{\sigma^2} - \frac{\sum_{i=1}^{p}\alpha_i}{\sigma^4} + \frac{tr.(\Theta - \sigma^2 I_q)}{\sigma^4}$$

= 0 (using (2.8))

$$\text{i.e. } \frac{p}{\sigma^2} - \frac{\sum_{i=1}^{p}\alpha_i}{\sigma^4} + \frac{\sum_{i=1}^{q}(\alpha_i - \sigma^2)}{\sigma^4} = 0$$

$$\text{i.e. } \hat{\sigma}^2 = \frac{\sum_{i=q+1}^{p}\alpha_i}{p-q} \qquad (2.11)$$

It remains to estimate the matrix Σ_1.An independent set of observations on noise is necessary to be found only to estimate Σ_1. Let $y(t_1), y(t_2),\ldots, y(t_m)$ be i.i.d. ~ $N_p(0, \sigma^2\Sigma_1)$. Let $y(t_i) = y_i = (y_{i1}, y_{i2},\ldots, y_{ip})'$ for convenience. Then the trivial estimator of the covariance matrix

$$\Sigma_1 \text{ is } \hat{\Sigma}_1 = \frac{1}{m}\sum_{i=1}^{m} \underset{\sim}{y_i}\, \underset{\sim}{y_i'} \qquad (2.12)$$

Hence, final estimator of the covariance matrix can be written as follows:

$$\text{Estimator of } (\Gamma + \sigma^2\Sigma_1) = \hat{B}\hat{B}' + \hat{\sigma}^2\hat{\Sigma}_1$$

$$= \hat{\Sigma}_1^{\frac{1}{2}}\widetilde{B}(\Theta - \hat{\sigma}^2 I_q)\widetilde{B}'\hat{\Sigma}_1^{\frac{1}{2}} + \hat{\sigma}^2\hat{\Sigma}_1 \qquad (2.13)$$

where

$\widetilde{B} = (w_1: w_2: \ldots: w_q)$

$\Theta = Diag.(\alpha_1, \alpha_2, \ldots, \alpha_q)$

α_r = r^{th} ordered eigen value of $\hat{\Sigma}_1^{-\frac{1}{2}} S^* \hat{\Sigma}_1^{-\frac{1}{2}}$

$\mathbf{w}_r$ = r^{th} orthonormal eigenvector of $\hat{\Sigma}_1^{-\frac{1}{2}} S^* \hat{\Sigma}_1^{-\frac{1}{2}}$ corresponding to α_r

$\hat{\sigma}^2$ is given by (2.11)

and $\hat{\Sigma}_1$ can be obtained from (2.12).

Strong Consistency of the Estimator

Lemma 3.1.

Let the observations $y_1, y_2,\ldots, y_m$ be i.i.d. ~ $N_p(0, \sigma^2\Sigma_1)$, where Σ_1 is some arbitrary positive definite matrix. Let $\hat{\Sigma}_1$ be the estimator of Σ_1 given by (2.12). Then $\hat{\Sigma}_1$ is a strongly consistent estimator of Σ_1.

Proof.

The proof of Lemma 3.1 is trivial from Strong Law of Large Number Theory.

Lemma 3.2

Suppose A, A_n, n = 1, 2, …, are all pxp symmetric matrices such that A_n-A = O(α_n) and $\alpha_n \to 0$ as n $\to \infty$. Denote by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ and $\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \ldots \geq \lambda_p^{(n)}$ the eigenvalues of A and A_n, respectively. Then,

$$\lambda_i^{(n)} - \lambda_i = O(\alpha_n) \text{ as n } \to \infty,\ i = 1,\ldots, p.$$

Proof.

The proof of Lemma 3.2 is given in Zhao, Krishnaiah and Bai (1986a).

Lemma 3.3

Suppose A, A_n, n = 1, 2, …, are all pxp symmetric matrices such that A_n-A = O(β_n) and $\beta_n \to 0$ as n $\to \infty$. Denote $f_1, f_2,\ldots, f_p$ and $f_1^{(n)}$, $f_2^{(n)}$, …, $f_p^{(n)}$ the eigenvectors of A and A_n respectively, corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_p$ and $\lambda_1^{(n)}, \lambda_2^{(n)}, \ldots, \lambda_p^{(n)}$ respectively.

Then, $\left\| f_i^{(n)} - f_i \right\| = O(\beta_n)$ as n $\to \infty, i = 1,\ldots, p.$

Note: Lemma 3.3 may not be true, if the symmetric matrix A has same eigenvalues. But it is true for those eigenvectors corresponding to distinct eigenvalues of A.

Proof.

The proof of Lemma 3.3 can be done similar way as in Zhao, Krishnaiah and Bai (1986a).

Theorem 3.1

Let $\Gamma + \hat{\sigma^2 \Sigma_1}$ be an estimator of $\Gamma + \sigma^2 \Sigma_1$ obtained from (2.13). Then $\Gamma + \hat{\sigma^2 \Sigma_1} \xrightarrow{a.s.} \Gamma + \sigma^2 \Sigma_1$ as $n \to \infty$ and $m \to \infty$.

Proof.
Using Lemma 3.1,

$$\hat{\Sigma}_1 \xrightarrow{a.s.} \Sigma_1 \text{ as } m \to \infty \qquad (3.1)$$

From Strong Law of Large Number Theory,

$$S^* = \frac{1}{n}\sum_{i=1}^{n} \underset{\sim}{x_i} \underset{\sim}{x_i'} \xrightarrow{a.s.} E(\underset{\sim}{x_1} \underset{\sim}{x_1'}) \text{ as } n \to \infty$$
$$= V(\underset{\sim}{x_1}) + \underset{\sim}{0}\underset{\sim}{0}'$$
$$= \Gamma + \sigma^2 \Sigma_1$$

Hence,

$$\hat{\Sigma}_1^{-\frac{1}{2}} S^* \hat{\Sigma}_1^{-\frac{1}{2}} \xrightarrow{a.s.} \Sigma_1^{-\frac{1}{2}} (\Gamma + \sigma^2 \Sigma_1) \Sigma_1^{-\frac{1}{2}} \text{ as } n \to \infty \text{ and } m \to \infty$$
$$= \Sigma_1^{-\frac{1}{2}} \Gamma \Sigma_1^{-\frac{1}{2}} + \sigma^2 I_p \qquad (3.2)$$

Let $l_1 > l_2 > \ldots > l_q > \sigma^2$ be the ordered eigenvalues of $\Sigma_1^{-\frac{1}{2}} \Gamma \Sigma_1^{-\frac{1}{2}} + \sigma^2 I_p$ and $\underset{\sim}{d_1}, \underset{\sim}{d_2}, \ldots, \underset{\sim}{d_p}$ be the corresponding orthonormal eigenvectors of $\Sigma_1^{-\frac{1}{2}} \Gamma \Sigma_1^{-\frac{1}{2}} + \sigma^2 I_p$. Then, using (3.2) and Lemma 3.2,

$$\alpha_i \xrightarrow{a.s.} l_i \ ; \ i = 1,2,\ldots,q$$

and

$$\alpha_i \xrightarrow{a.s.} \sigma^2 \text{ for } i = q+1, \ldots, p \text{ as } n \to \infty \qquad (3.3)$$

Because the eigenvalues $l_1, l_2, \ldots, l_q$ of $\Sigma_1^{-\frac{1}{2}} \Gamma \Sigma_1^{-\frac{1}{2}} + \sigma^2 I_p$ are not the same, using (3.2) and Lemma 3.3,

$$\underset{\sim}{w_i} \xrightarrow{a.s.} \underset{\sim}{d_i} \ ; \ i = 1,2,\ldots,q \text{ as } n \to \infty \qquad (3.4)$$

where α_i's and $\underset{\sim}{w_i}$'s are explained in (2.13).

Now, $\hat{\sigma}^2 = \dfrac{\sum_{i=q+1}^{p} \alpha_i}{p-q} \xrightarrow{a.s.} \sigma^2$ as $n \to \infty$ (using (3.3)) (3.5)

and

$$\Gamma + \hat{\sigma^2 \Sigma_1} = \hat{\Sigma}_1^{\frac{1}{2}} \widetilde{B} (\Theta - \hat{\sigma}^2 I_q) \widetilde{B}' \hat{\Sigma}_1^{\frac{1}{2}} + \hat{\sigma}^2 \hat{\Sigma}_1$$
$$= \hat{\Sigma}_1^{\frac{1}{2}} \left(\sum_{i=1}^{q} (\alpha_i - \hat{\sigma}^2) \underset{\sim}{w_i} \underset{\sim}{w_i'}\right) \hat{\Sigma}_1^{\frac{1}{2}} + \hat{\sigma}^2 \hat{\Sigma}_1$$
$$\xrightarrow{a.s.} \Sigma_1^{\frac{1}{2}} \left(\sum_{i=1}^{q} (l_i - \sigma^2) \underset{\sim}{d_i} \underset{\sim}{d_i'}\right) \Sigma_1^{\frac{1}{2}} + \sigma^2 \Sigma_1 \qquad (3.6)$$

Because $\underset{\sim}{d_1}, \underset{\sim}{d_2}, \ldots, \underset{\sim}{d_p}$ are orthonormal eigenvectors,

$$DD' = I_p \text{ where } \underset{pxp}{D} = (\underset{\sim}{d_1} : \underset{\sim}{d_2} : \ldots : \underset{\sim}{d_p})$$

Hence,

$$\sigma^2 I_p = \sigma^2 \sum_{i=1}^{q} \underset{\sim}{d_i} \underset{\sim}{d_i'} + \sigma^2 \sum_{i=q+1}^{p} \underset{\sim}{d_i} \underset{\sim}{d_i'} \qquad (3.7)$$

Again, from Spectral Decomposition,

$$\Sigma_1^{-\frac{1}{2}} \Gamma \Sigma_1^{-\frac{1}{2}} + \sigma^2 I_p = \sum_{i=1}^{q} l_i \underset{\sim}{d_i} \underset{\sim}{d_i'} + \sigma^2 \sum_{i=q+1}^{p} \underset{\sim}{d_i} \underset{\sim}{d_i'} \qquad (3.8)$$

Therefore,

$$\sum_{i=1}^{q}(l_i-\sigma^2)\,\underset{\sim}{d_i}\,\underset{\sim}{d_i'}$$

$$=\sum_{i=1}^{q} l_i\,\underset{\sim}{d_i}\,\underset{\sim}{d_i'} - \sigma^2\sum_{i=1}^{q}\underset{\sim}{d_i}\,\underset{\sim}{d_i'}$$

$$=(\Sigma_1^{-\frac{1}{2}}\Gamma\Sigma_1^{-\frac{1}{2}}+\sigma^2 I_p - \sigma^2\sum_{i=q+1}^{p}\underset{\sim}{d_i}\,\underset{\sim}{d_i'}) -$$

$$(\sigma^2 I_p - \sigma^2\sum_{i=q+1}^{p}\underset{\sim}{d_i}\,\underset{\sim}{d_i'})$$

(using (3.7) and (3.8))

$$=\Sigma_1^{-\frac{1}{2}}\Gamma\Sigma_1^{-\frac{1}{2}} \qquad (3.9)$$

Using (3.9) in (3.6), we get Theorem 3.1.

Asymptotic Normality of the Estimator

Theorem 4.1

Let $\Gamma+\hat{\sigma^2\Sigma_1}$ be an estimator of $\Gamma+\sigma^2\Sigma_1$ obtained from (2.13).

Then the limiting distribution of $\sqrt{n}$ ($\Gamma+\hat{\sigma^2\Sigma_1}$ - $\Gamma+\sigma^2\Sigma_1$) is normal with mean 0 and variance B where B is given by (4.5) later.

Proof.

From (3.1) $\hat{\Sigma}_1 \xrightarrow{a.s.} \Sigma_1$ as $m\to\infty$. Because

$$S^*=\frac{1}{n}\sum_{i=1}^{n}\underset{\sim}{x_i}\,\underset{\sim}{x_i'},$$

where

$$\underset{\sim}{x_i}\sim N_p(\underset{\sim}{0},\Gamma+\sigma^2\Sigma_1);\ i=1,2,...,n,$$

using Theorem 3.4.4 of Anderson (1984), p.81, the limiting distribution of

$$C(n)=\sqrt{n}(\hat{\Sigma}_1^{-\frac{1}{2}}S^*\hat{\Sigma}_1^{-\frac{1}{2}}-\Sigma_1^{-\frac{1}{2}}\Gamma\Sigma_1^{-\frac{1}{2}}+\sigma^2 I_p)$$

is normal with mean 0 and covariance

$$E(C_{ij}(n)C_{kl}(n))=\sigma_{ik}\sigma_{jl}+\sigma_{il}\sigma_{jk} \qquad (4.1)$$

where $\sigma_{ij}=(i,j)^{th}$ element of

$$\Sigma_1^{-\frac{1}{2}}\Gamma\Sigma_1^{-\frac{1}{2}}+\sigma^2 I_p.$$

(4.1) is obtained due to the fact that

$$S^{**}=\hat{\Sigma}_1^{-\frac{1}{2}}S^*\hat{\Sigma}_1^{-\frac{1}{2}}=\frac{1}{n}\sum_{i=1}^{n}\underset{\sim}{u_i^*}\,\underset{\sim}{u_i^{*\prime}}$$

asymptotically (using 3.1) and

$$\underset{\sim}{u_i^*}=\Sigma_1^{-\frac{1}{2}}\underset{\sim}{x_i}\sim N_p(\underset{\sim}{0},\Sigma_1^{-\frac{1}{2}}\Gamma\Sigma_1^{-\frac{1}{2}}+\sigma^2 I_p)$$

From (2.13), estimator of $\Gamma+\sigma^2\Sigma_1$ is

$$\Gamma+\hat{\sigma^2\Sigma_1} = \hat{\Sigma}_1^{\frac{1}{2}}\widetilde{B}(\Theta-\hat{\sigma}^2 I_q)\widetilde{B}'\hat{\Sigma}_1^{\frac{1}{2}}+\hat{\sigma}^2\hat{\Sigma}_1$$

$$=\hat{\Sigma}_1^{\frac{1}{2}}(\sum_{i=1}^{q}(\alpha_i-\hat{\sigma}^2)\,\underset{\sim}{w_i}\,\underset{\sim}{w_i'})\hat{\Sigma}_1^{\frac{1}{2}}+\hat{\sigma}^2\hat{\Sigma}_1$$

where α_i's, $\underset{\sim}{w_i}$'s and $\hat{\sigma}^2$ are explained in (2.13).

Because

$$\hat{\Sigma}_1 \xrightarrow{a.s.} \Sigma_1 \text{ as } m\to\infty$$

(using (3.1))

$$\underset{\sim}{w_i}\xrightarrow{a.s.}\underset{\sim}{d_i}\ ;\ i=1,2,...,q \text{ as } n\to\infty$$

(using (3.4))

and

$$\hat{\sigma}^2\xrightarrow{a.s.}\sigma^2 \text{ as } n\to\infty$$

(using (3.5)),

the limiting distribution of $\Gamma+\hat{\sigma^2\Sigma_1}$ is same as that of

$$\Sigma_1^{\frac{1}{2}}(\sum_{i=1}^{q}\alpha_i-\sigma^2)\,\underset{\sim}{d_i}\,\underset{\sim}{d_i'})\Sigma_1^{\frac{1}{2}}+\sigma^2\Sigma_1$$

(see Rao, 1983, p.122, (x)(b)) (4.2)

Using the result of Anderson (1984) p.468,

$$E(\alpha_i)=l_i\ ;\ i=1,2,...,q$$

asymptotically. Hence, from (4.2),

$$E(\Sigma_1^{\frac{1}{2}}(\sum_{i=1}^{q}\alpha_i - \sigma^2)\underset{\sim}{d_i}\,\underset{\sim}{d_i'})\Sigma_1^{\frac{1}{2}} + \sigma^2\Sigma_1)$$
$$= \Sigma_1^{\frac{1}{2}}(\sum_{i=1}^{q}(l_i - \sigma^2)\underset{\sim}{d_i}\,\underset{\sim}{d_i'})\Sigma_1^{\frac{1}{2}} + \sigma^2\Sigma_1$$
$$= \Gamma + \sigma^2\Sigma_1 \text{ (see 3.6 and 3.9).}$$

From (4.2), the asymptotic variance of the estimator is same as that of

$$\sum_{i=1}^{q}\alpha_i \underset{\sim}{f_i}\,\underset{\sim}{f_i'} \text{, where } \underset{\sim}{f_i} = \Sigma_1^{\frac{1}{2}}\underset{\sim}{d_i} \qquad (4.3)$$

From the result of Anderson (1984) p.468, $\sqrt{n}(\alpha_i - l_i)$; $i = 1,2,...,q$ are independently distributed and

$$\sqrt{n}(\alpha_i - l_i) \sim N(0,2l_i^2) \;;\; i = 1,2,...,q \quad (4.4)$$

Hence, asymptotic variance of $\sum_{i=1}^{q}\alpha_i \underset{\sim}{f_i}\,\underset{\sim}{f_i'}$ can be obtained using (4.4). Call the asymptotic variance as

$$\text{V}(\sum_{i=1}^{q}\alpha_i \underset{\sim}{f_i}\,\underset{\sim}{f_i'}) = B. \qquad (4.5)$$

References

Anderson, T. W. (1963). Asymptotic theory for principal component analysis, *Annals of Mathematical Statistics, 34*, 122-138.

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd Ed.). *NY: Wiley*.

Bellman, R. (1960). *Introduction to matrix analysis. New York: McGraw-Hill.*

Chen, P. (2002). A selection procedure for estimating the number of signal components. *J*ournal of Statistical Planning and Inference, 105, 299-301.

Chen, P., Wicks, M.C., & Adve, R. S. (2001). Development of a statistical procedure for detecting the number of signals in a radar measurement. *IEEE Proceedings of Radar, Sonar and Navigations*, *148*(4), 219-226.

Krishnaiah, P.R. (1976). Some recent developments on complex multivariate distributions, *Journal of Multivariate Analysis*, *6*, 1-30.

Kundu, D. (2000). Estimating the number of signals in the presence of white noise. *Journal of Statistical Planning and Inference*, 90, 57-68.

Lawley, D. N., & Maxwell, A.E. (1963). Factor Analysis as a Statistical Method, *Butterworths, London*.

Rao, C.R. (1983). Likelihood ratio tests for relationships between two covariance matrices. *In: T. Amemiya, S. Karlin and L. Goodman, eds. Studies in Econometrics*, *Time Series and Multivariate Statistics*. NY: Academic Press.

Rao, C.R. (1983). Linear statistical inference and its applications. NY: *Wiley Eastern Limited.*

Wax, M., T. Kailath (1985). Determination of the number of signals by information theoretic criteria, *IEEE Trans. Acoustics Speech Signal Processing ASSP*-33, 387-392.

Wax, M., Shan, T. J., & Kailath, T. (1984). Spatio temporal spectral analysis by eigen structure methods, *IEEE Trans. Acoustics Speech Signal Processing ASSP*-32, 817-827.

Zhao, L. C., Krishnaiah, P. R., & Bai, Z. D. (1986a). On detection of number of signals in presence of white noise, *Journal of Multivariate Analysis*, 20, 1-25.

Zhao, L. C., Krishnaiah, P. R., & Bai, Z. D. (1986b). On detection of the number of signals when the noise covariance matrix is arbitrary, *Journal of Multivariate Analysis*, *20*, 26-49.