Ann. Inst. Statist. Math.
38 (1986), Part A, 57-68

CONSISTENCY CONDITIONS ON THE LEAST SQUARES ESTIMATOR IN SINGLE COMMON FACTOR ANALYSIS MODEL

YUTAKA KANO

(Received Sept. 7, 1983; revised Apr. 16, 1985)

Summary

This paper is concerned with the consistency of estimators in a single common factor analysis model when the dimension of the observed vector is not fixed. In the model several conditions on the sample size n and the dimension p are established for the least squares estimator (L.S.E.) to be consistent. Under some assumptions, $p/n \to 0$ is a necessary and sufficient condition that the L.S.E. converges in probability to the true value. A sufficient condition for almost sure convergence is also given.

1. Introduction

A common factor analysis model belongs to the family of covariance structure models and asymptotic properties of their estimators have been studied by many authors. Anderson and Rubin [1] first formulated factor analysis as a mathematical problem and proved the consistency and the asymptotic normality of the maximum likelihood estimator (M.L.E.) for the structural parameter. Asymptotic properties of the generalized least squares estimator (G.L.S.E.) were investigated by Browne [2].

On the other hand, Tumura and Fukutomi [6] and Fukutomi [3] reported, by using numerical examples, that the M.L.E. can be a discontinuous function of the sample variance matrix. Kano [4], however, showed that each of the M.L.E. and G.L.S.E., as a function of the sample variance matrix, is continuous at least at the true variance matrix, and hence is consistent.

When the distance between the sample and true variance matrices is not small enough or the sample size is not sufficiently large, these estimators are quite likely to get far from the true value because they

Key words: Consistency, dimension of observed vector, least squares estimator, sample size, single common factor analysis model.

are continuous but complicated functions of the sample. The sample size which is required to yield the good estimate depends on the true value and the structure of the model—especially the dimension, i.e. the number of items.

The previous authors studied the properties of the consistency under the assumption that the dimension p is fixed but did not consider the influence of p. Assuming a single common factor analysis model, this paper investigates how the sample size n should be increased so that the least squares estimator (L.S.E.) is to be near the true value as the dimension p varies. Particularly, in Section 2 some conditions on n and p are given under which the L.S.E. converges to the true value in probability or almost surely. These results are extensions of the results in Kano [4] although we restrict the number of common factors to one and the estimation method to the least squares method in this paper (see Lemmas 1 and 2 in Section 3).

2. Main results

Consider a common factor analysis model given by Williams [7] which is known to be free from factor indeterminacy under some conditions. An infinite dimensional random vector $x=(X_1, X_2, \cdots)'$ is said to conform to a single common factor analysis model if there exist scalars $u_i \geqq 0$, a_i $(i=1, 2, \cdots)$ and uncorrelated random variables Y and E_i $(i=1, 2, \cdots)$ with zero mean and unit variance such that

$$X_i = a_i Y + u_i E_i \qquad (i=1, 2, \cdots),$$

where $(a_1, a_2, \cdots, u_1, u_2, \cdots)'$ is a structural parameter (vector) and Y is a common factor of x. Since the model is scale invariant, we can assume without loss of generality that

$$\text{(A.1)} \qquad a_i^2 + u_i^2 = 1 \qquad (i=1, 2, \cdots).$$

For any positive integer p, let $x_p=(X_1, \cdots, X_p)'$, $a_p=(a_1, \cdots, a_p)'$ and $U_p=\text{diag}(u_1, \cdots, u_p)$, which stands for the diagonal matrix with diagonal elements $u_1, \cdots, u_p$. Then we have

$$\text{(2.1)} \qquad \text{Var}\{x_p\} = a_p a_p' + U_p^2 \quad (=\Sigma_p, \text{ say}),$$

which means that x_p conforms to an ordinary single common factor analysis model with dimension p.

Let $x_{pk}=(X_{1k}, \cdots, X_{pk})'$ for $k=1, \cdots, n$ be a (projective) random sample taken from the single common factor analysis model. The purpose of this paper is to find conditions that the estimator $(\hat{a}_{np}, \hat{U}_{np})$ for the (projective) parameter (a_p, U_p), which is constructed from the above sample, converges to the true value in probability (or almost surely), i.e.,

$$\|\hat{a}_{np}-a_p\| \xrightarrow{\text{P (a.s.)}} 0 \quad \text{and} \quad \|\hat{U}^2_{np}-U^2_p\| \xrightarrow{\text{P (a.s.)}} 0 ,$$

where the symbol $\|\cdot\|$ denotes the usual Euclidean norm of matrices and vectors. Since the above quantities are double sequences with indices n and p, conditions on the pair (n, p) will be considered in this paper.

Let $S_{np}=n^{-1}\sum_{k=1}^{p} x_{pk}x'_{pk}$. For simplicity we omit all subscripts of Σ_p, S_{np}, a_p, $\hat{a}_{np}$, U_p and $\hat{U}_{np}$. We assume that $p\geqq 3$ throughout this paper because of nonidentifiability for $p<3$ (see Anderson and Rubin [1], Theorem 5.5). Then the least squares estimator (L.S.E.) for the structural parameter (a, U) is determined by the following relation:

$$\|\hat{\Sigma}-S\|=\min_{\bar{a},\bar{U}} \|\bar{\Sigma}-S\| , \tag{2.2}$$

where $\hat{\Sigma}=\hat{a}\hat{a}'+\hat{U}^2$ and $\bar{\Sigma}=\bar{a}\bar{a}'+\bar{U}^2$. The L.S.E. exists with probability one because the parameter space can be compactified. By the definition of the estimator we have easily

$$\|\hat{\Sigma}-\Sigma\|\leqq\|\hat{\Sigma}-S\|+\|S-\Sigma\|\leqq 2\|S-\Sigma\| , \tag{2.3}$$

where Σ is the true variance matrix.

Write $S=(s_{ij})$, $\hat{a}=(\hat{a}_1,\cdots,\hat{a}_p)'$ and $\hat{U}=\text{diag}\,(\hat{u}_1,\cdots,\hat{u}_p)$. In order to remove the indeterminacy due to the sign of the loading vector a and its estimator $\hat{a}$, we pose the following restrictions:

$$a_1>0 \quad \text{and} \quad \hat{a}_1\geqq 0 . \tag{2.4}$$

We impose the following assumptions about the underlying probability space and the true value of the parameter:

(A.2) there exists a scalar $M_0>0$ such that $E(X_i^4)<M_0$ for all $i \in N$,

(A.3) there exists a scalar $a_0>0$ such that $|a_i|>a_0$ for all $i \in N$,

where N denotes the set of all positive integers. Under the above assumptions we have the following

THEOREM 1. *The L.S.E.* $(\hat{a}, \hat{U})$ *is determined by* (2.2). *Then under* (A.1)–(A.3), *if* $n^{-1}p\to 0$, *then* $\|\hat{a}-a\|\xrightarrow{\text{P}}0$ *and* $\|\hat{U}^2-U^2\|\xrightarrow{\text{P}}0$, *i.e., for any* $\varepsilon_1, \varepsilon_2>0$ *there exists a scalar* $\delta>0$ *such that if* $n^{-1}p<\delta$, *then* $\Pr\{\|\hat{a}-a\|<\varepsilon_1$ *and* $\|\hat{U}^2-U^2\|<\varepsilon_1\}>1-\varepsilon_2$.

In order to establish the converse of Theorem 1 we further assume the following

(A.4) there exists a scalar $u_0>0$ such that $u_i>u_0$ for all $i \in N$,

(A.5) there exists a scalar $m_0>0$ such that $\mathrm{Cov}(X_i^2, X_j^2)>m_0$ for all $i, j \in N$,

(A.6) there exists a scalar $\gamma>0$ such that

$$\overline{\lim_{p\to\infty}} E\left\{\left|p^{-1}\sum_{i=1}^{p}(X_i^2-1)\right|^{2+\gamma}\right\}<\infty .$$

Remark. The assumption (A.5) is satisfied if for any i and j $(i\neq j)$, (X_i, X_j) has the bivariate normal distribution with $E(X_i)=0$ and $\inf_{i,j\in N} |E(X_iX_j)|>0$. The assumption (A.6) guarantees the Liapunov condition and it can be replaced by the following assumption:

(A.6)′ there exists a random variable X with $E(X^4)<\infty$ such that $|X_i| <X$ (w.p.l.) for all $i \in N$,

which is sufficient for the Lindeberg-Feller condition.

THEOREM 2. *The L.S.E.* $(\hat{a}, \hat{U})$ *is determined by* (2.2). *Under* (A.1) *and* (A.4)-(A.6), *if* $\|\hat{a}-a\|\xrightarrow{P}0$ *and* $\|\hat{U}^2-U^2\|\xrightarrow{P}0$, *then* $n^{-1}p\to 0$, *i.e., for any* $\varepsilon>0$ *there exist scalars* $\delta_1, \delta_2>0$ *such that if* $\Pr\{\|\hat{a}-a\|\leqq\delta_1$ *and* $\|\hat{U}^2-U^2\|\leqq\delta_1\}>1-\delta_2$, *then* $n^{-1}p<\varepsilon$.

Theorems 1 and 2 show that under (A.1)-(A.6), $n^{-1}p\to 0$ is a necessary and sufficient condition for $\|\hat{a}-a\|$ and $\|\hat{U}^2-U^2\|$ to converge to zero in probability.

Finally we give a condition under which the L.S.E. converges to the true value almost surely.

THEOREM 3. *The L.S.E. is determined by* (2.2). *Under* (A.1)-(A.3) *for every* $\alpha>0$, *if* $n^{-1}p^{2+\alpha}\to 0$ *then* $\|\hat{a}-a\|\xrightarrow{\text{a.s.}}0$ *and* $\|\hat{U}^2-U^2\|\xrightarrow{\text{a.s.}}0$, *i.e.*,

$$\Pr\{\lim_{f(n,p)\to 0}\|\hat{a}-a\|=0 \text{ and } \lim_{f(n,p)\to 0}\|\hat{U}^2-U^2\|=0\}=1 ,$$

where $f(n, p)=n^{-1}p^{2+\alpha}$.

3. Proofs of the theorems

The proofs of Theorems 1–3 depend on the following lemmas, which will be proved in Appendix.

LEMMA 1. *Under* (A.1) *and* (A.3), *there exist scalars* δ *and* $M_1>0$ *such that if* $p^{-1}\sum_{i\neq j}^{p}(\hat{a}_i\hat{a}_j-a_ia_j)^2<\delta^2$, *then*

$$\sum_{i=1}^{p} (\hat{a}_i - a_i)^2 \leqq p^{-1} M_1 \sum_{i \neq j}^{p} (\hat{a}_i \hat{a}_j - a_i a_j)^2 .$$

LEMMA 2. (i) *Under* (A.1),

$$\sum_{i=1}^{p} (\hat{u}_i^2 - u_i^2)^2 \leqq 2 \sum_{i=1}^{p} \{(s_{ii} - 1)^2 + (\hat{a}_i^2 - a_i^2)^2\} .$$

(ii) *Under* (A.1) *and* (A.4) *there exists a scalar* $M_2 > 0$ *such that*

$$\sum_{i=1}^{p} (s_{ii} - 1)^2 \leqq M_2 \sum_{i=1}^{p} \{(\hat{a}_i^2 - a_i^2)^2 + (\hat{u}_i^2 - u_i^2)^2\} .$$

We are now in a position to prove theorems.

PROOF OF THEOREM 1. Assume that $4p^{-1}\|S - \Sigma\|^2 < \delta^2$. Then from (2.3) and Lemma 1 we have

$$\|\hat{a} - a\|^2 \leqq 4p^{-1} M_1 \|S - \Sigma\|^2 < M_1 \delta^2 . \tag{3.1}$$

When p is fixed, Lemmas 1, 2 and the relation (3.1) imply that the L.S.E. is a continuous function of S, as was shown by Kano [4].

From the assumption (A.2) we have

$$E\{\|S - \Sigma\|^2\} < n^{-1} p^2 M_0 , \tag{3.2}$$

which implies from Markov's inequality that if $4n^{-1}pM_0/\varepsilon < \delta^2$, then

$$\Pr\{4p^{-1}\|S - \Sigma\|^2 < \delta^2\} > 1 - \varepsilon .$$

This implies from (3.1) that

$$\Pr\{\|\hat{a} - a\|^2 < M_1 \delta^2\} > 1 - \varepsilon ,$$

which proves that

$$\|\hat{a} - a\| \xrightarrow{\text{P}} 0 \qquad \text{as } n^{-1}p \to 0 . \tag{3.3}$$

Next we shall prove that $\|\hat{U}^2 - U^2\| \xrightarrow{\text{P}} 0$. Since $E\left\{\sum_{i=1}^{p} (s_{ii} - 1)^2\right\} < n^{-1}pM_0$, we have similarly

$$\sum_{i=1}^{p} (s_{ii} - 1)^2 \xrightarrow{\text{P}} 0 \qquad \text{as } n^{-1}p \to 0 . \tag{3.4}$$

By (A.1) and Schwarz' inequality,

$$\begin{aligned} \sum_{i=1}^{p} (\hat{a}_i^2 - a_i^2)^2 &\leqq 2 \sum_{i=1}^{p} (\hat{a}_i - a_i)^4 + 8 \sum_{i=1}^{p} (\hat{a}_i - a_i)^2 \\ &\leqq 2\|\hat{a} - a\|^4 + 8\|\hat{a} - a\|^2 , \end{aligned} \tag{3.5}$$

which implies that

$$\sum_{i=1}^{p}(\hat{a}_i^2-a_i^2)^2 \xrightarrow{P} 0 \quad \text{as } n^{-1}p \to 0 . \tag{3.6}$$

It follows from (3.4), (3.6) and Lemma 2 (i) that

$$\sum_{i=1}^{p}(\hat{u}_i^2-u_i^2)^2 \xrightarrow{P} 0 \quad \text{as } n^{-1}p \to 0 .$$ Q.E.D.

PROOF OF THEOREM 2. In this part we write $S=(s_{nij})$. From (3.5) and Lemma 2 (ii) we have

$$\sum_{i=1}^{p}(s_{nii}-1)^2 \leqq M_2\{2\|\hat{a}-a\|^4+8\|\hat{a}-a\|^2+\|\hat{U}^2-U^2\|^2\}$$

and hence the assumptions of Theorem 2 imply that

$$\sum_{i=1}^{p}(s_{nii}-1)^2 \xrightarrow{P} 0 .$$

It is sufficient to prove that for any $\varepsilon>0$ there exist scalars $\delta_1, \delta_2>0$ such that if $\Pr\left\{\sum_{i=1}^{p}(s_{nii}-1)^2<\delta_1\right\}>1-\delta_2$, then $n^{-1}p<\varepsilon$.

We shall show a contradiction by assuming that there exist a scalar $\varepsilon_0>0$, sequences $\{n_l\}$, $\{p_l\}$, $\{\delta_1^l\}$ and $\{\delta_2^l\}$ such that
(i) $\delta_1^l, \delta_2^l>0$ and $\delta_1^l, \delta_2^l \to 0$ as $l\to\infty$,
(ii) $\Pr\left\{\sum_{i=1}^{p_l}(s_{n_l ii}-1)^2<\delta_1^l\right\}>1-\delta_2^l$ for all $l\in N$ and that
(iii) $n_l^{-1}p_l \geqq \varepsilon_0$ for all $l\in N$.
From (ii) we have

$$\Pr\{(s_{n_l 11}-1)^2<\delta_1^l\}>1-\delta_2^l \quad \text{for all } l\in N. \tag{3.7}$$

It follows from (i), (3.7) and (A.5) that

$$n_l \to \infty \quad \text{as } l\to\infty . \tag{3.8}$$

By Schwarz' inequality,

$$\sum_{i=1}^{p}(s_{nii}-1)^2=\sum_{i=1}^{p}\left\{n^{-1}\sum_{k=1}^{n}(X_{ik}^2-1)\right\}^2 \geqq p^{-1}\left\{\sum_{i=1}^{p}n^{-1}\sum_{k=1}^{n}(X_{ik}^2-1)\right\}^2 ,$$

which implies that

$$\Pr\left[\left\{p_l^{-1/2}n_l^{-1}\sum_{k=1}^{n_l}\sum_{i=1}^{p_l}(X_{ik}^2-1)\right\}^2<\delta_1^l\right]>1-\delta_2^l \quad \text{for all } l\in N. \tag{3.9}$$

Define

$$Y_{lk}=\sum_{i=1}^{p_l}(X_{ik}^2-1), \quad T_l=\sum_{k=1}^{n_l}Y_{lk} \quad \text{and} \quad s_l^2=\sum_{k=1}^{n_l}\mathrm{Var}(Y_{lk}) .$$

Then we can rewrite (3.9) as

(3.10) $$\Pr[\{p_l^{-1/2}n_l^{-1}s_l(s_l^{-1}T_l)\}^2<\delta_1^l]>1-\delta_2^l \qquad \text{for all } l\in N\,.$$

First we shall show that $s_l^{-1}T_l \xrightarrow{L} N(0,1)$ as $l\to\infty$. The variable T_l is the sum of independent and identically distributed variables and hence under (A.5) we have

(3.11) $$\mathrm{Var}\,(Y_{l1})=\sum_{i=1}^{p_l}\sum_{j=1}^{p_l}\mathrm{Cov}\,(X_{i1}^2, X_{j1}^2)>p_l^2m_0 \qquad \text{and}$$

(3.12) $$s_l^2=n_l\,\mathrm{Var}\,(Y_{l1})>n_lp_l^2m_0\,.$$

From (3.8), (3.12) and (A.6) we have

$$\begin{aligned} s_l^{-(2+r)}\sum_{k=1}^{n_l}E\{|Y_{lk}|^{2+r}\} &= n_ls_l^{-(2+r)}E\{|Y_{l1}|^{2+r}\} \\ &< m_0^{-(2+r)/2}n_l^{-r/2}p_l^{-(2+r)}E\{|Y_{l1}|^{2+r}\}\to 0 \\ &\qquad\qquad \text{as } l\to\infty\,, \end{aligned}$$

which implies that

(3.13) $$s_l^{-1}T_l \xrightarrow{L} N(0,1)$$

by using the Liapunov Theorem. Under the assumption (A.6)′ we can use the Lindeberg-Feller Theorem. Since

$$Y_{l1}^2=\left\{\sum_{i=1}^{p_l}(X_{i1}^2-1)\right\}^2\leqq p_l\sum_{i=1}^{p_l}(X_{il}-1)^2<p_l\sum_{i=1}^{p_l}(X^2+1)^2<p_l^2(X^2+1)^2\,,$$

we have from (3.8), (3.11) and (3.12)

$$\begin{aligned} L(\varepsilon) &\equiv s_l^{-2}\sum_{k=1}^{n_l}\int_{|Y_{lk}|\geq\varepsilon s_l}Y_{lk}^2dP \\ &= \{\mathrm{Var}\,(Y_{l1})\}^{-1}\times\int_{|Y_{l1}|\geq\varepsilon s_l}Y_{l1}^2dP \\ &< (p_l^2m_0)^{-1}\times\int_{p_l(X^2+1)\geq\varepsilon(n_lp_l^2m_0)^{1/2}}p_l^2(X^2+1)^2dP \\ &= m_0^{-1}\times\int_{X^2+1>\varepsilon(n_lm_0)^{1/2}}(X^2+1)^2dP\to 0 \qquad \text{as } l\to\infty\,, \end{aligned}$$

where P is a probability measure on the underlying probability space. It follows from (i), (3.10) and (3.13) that $p_l^{-1/2}n_l^{-1}s_l\to 0$ as $l\to\infty$ so that by using (3.12) $n_l^{-1}p_l\to 0$ as $l\to\infty$. This contradicts (iii). Q.E.D.

PROOF OF THEOREM 3. If p is bounded, we have $p^{-1}\|S-\Sigma\|^2\xrightarrow{\text{a.s.}} 0$ because the assumptions of Theorem 3 imply that $n\to\infty$ and S is the mean of independent and identically distributed variables $x_{pk}x_{pk}'$ $(k=1,\cdots,n)$ with $E(x_{pk}x_{pk}')=\Sigma$.

We assume therefore that p increases unboundedly. The assumptions of Theorem 3 imply that $n^{-1}p^{2+\alpha}<c$ for some $c>0$. For any $p\in$

N, let n be any element of the set $\{n \in N \mid n^{-1}p^{2+\alpha} < c\}$. Since

$$\sum_{p=1}^{\infty} p^{-1}E\{\|S-\Sigma\|^2\} < M_0 \sum_{p=1}^{\infty} n^{-1}p < M_0 c \sum_{p=1}^{\infty} p^{-(1+\alpha)} < \infty \ ,$$

we have $p^{-1}\|S-\Sigma\|^2 \xrightarrow{\text{a.s.}} 0$ by using Theorem (iii) on p. 111 of Rao [5]. Therefore it follows from (2.3) and Lemma 1 that

$$\|\hat{a}-a\| \xrightarrow{\text{a.s.}} 0 \ , \tag{3.14}$$

and hence we have from (3.5)

$$\sum_{i=1}^{p} (\hat{a}_i^2 - a_i^2)^2 \xrightarrow{\text{a.s.}} 0 \ . \tag{3.15}$$

Similarly we have

$$\sum_{p=1}^{\infty} \left[E\left\{ \sum_{i=1}^{p} (s_{ii}-1)^2 \right\} \right] < M_0 \sum_{p=1}^{\infty} n^{-1}p < \infty \ ,$$

which implies

$$\sum_{i=1}^{p} (s_{ii}-1)^2 \xrightarrow{\text{a.s.}} 0 \ . \tag{3.16}$$

It follows from (3.15), (3.16) and Lemma 2 (i) that

$$\|\hat{U}^2 - U^2\| \xrightarrow{\text{a.s.}} 0 \ . \qquad \text{Q.E.D.}$$

Appendix

In Appendix we give proofs of Lemmas 1 and 2. For simplicity we write b_i and v_i instead of $\hat{a}_i$ and $\hat{u}_i$ $(i=1, \cdots, p)$, respectively.

PROOF OF LEMMA 1. First we shall show some preliminary results (R.1) through (R.4), which are related to the quantities

$$c_{ij} = b_i b_j - a_i a_j \qquad (i, j = 1, \cdots, p) \ . \tag{1}$$

(R.1) If $|b_1^2 - a_1^2| < \eta_1$, then $|b_1 - a_1| < a_0^{-1}\eta_1$.

PROOF. From (2.4) we have

$$|b_1 - a_1| = (b_1 + a_1)^{-1} |b_1^2 - a_1^2| < a_1^{-1}\eta_1 < a_0^{-1}\eta_1 \ . \qquad \text{Q.E.D.}$$

(R.2) If $|c_{ij}| < \eta_2$ and $|b_j - a_j| < L\eta_2$ for some $L>0$, then $|b_i - a_i| < a_0^{-1}(L|b_i| + 1)\eta_2$.

PROOF. $|b_i - a_i| \leqq |a_j|^{-1}\{|a_j b_i - b_j b_i| + |b_j b_i - a_j a_i|\} < a_0^{-1}(L|b_i|+1)\eta_2$. Q.E.D.

(R.3) If $|c_{ij}|, |c_{jk}|, |c_{ki}| < 2^{-1}a_0^2$ (<1) for different i, j and k, then $(b_i^2 - a_i^2)^2$

$\leqq 48a_0^{-4}(c_{ij}^2+c_{jk}^2+c_{ki}^2)$.

PROOF. From (1) we have

(2) $$(b_ib_jb_k)^2=(a_ia_j+c_{ij})(a_ja_k+c_{jk})(a_ka_i+c_{ki}) \qquad \text{and}$$

(3) $$(b_jb_k)^2=(a_ja_k+c_{jk})^2 .$$

Since

(4) $$|a_ja_k+c_{jk}|\geqq|a_ja_k|-|c_{jk}|>2^{-1}a_0^2 ,$$

dividing (2) by (3) we have

(5) $$b_i^2=(a_ja_k+c_{jk})^{-1}(a_ia_j+c_{ij})(a_ka_i+c_{ki}) .$$

From (4) and (5) we have (noting that $|c_{ij}|<1$)

$$|b_i^2-a_i^2|\leqq 2a_0^{-2}(|c_{ij}|+|c_{jk}|+|c_{ki}|+|c_{ij}||c_{ki}|)\leqq 4a_0^{-2}(|c_{ij}|+|c_{jk}|+|c_{ki}|) ,$$

which implies that

$$(b_i^2-a_i^2)^2\leqq 48a_0^{-4}(c_{ij}^2+c_{jk}^2+c_{ki}^2) . \qquad \text{Q.E.D.}$$

(R.4) There exist scalars η_3, $L'>0$ which depend only on a_0 such that for any $p\geqq 3$ if $|c_{ij}|<\eta_3$ for all $1\leqq i,\ j\leqq p$ and $i\neq j$, then

$$\sum_{i=1}^{p}(b_i-a_i)^2\leqq p^{-1}L'\sum_{i\neq j}^{p}c_{ij}^2 .$$

PROOF. Taking $\eta_3<2^{-1}a_0^2$, we have from (R.3)

(6) $$(b_i^2-a_i^2)^2\leqq 48a_0^{-4}(c_{ij}^2+c_{jk}^2+c_{ki}^2)$$

for all different $1\leqq i,\ j,\ k\leqq p$, which implies that

$$\sum(b_i^2-a_i^2)^2\leqq 48a_0^{-4}\sum(c_{ij}^2+c_{jk}^2+c_{ki}^2) ,$$

where the summation is extended over all different subscripts $1\leqq i,\ j,\ k\leqq p$. Then we have

$$(p-1)(p-2)\sum_{i=1}^{p}(b_i^2-a_i^2)^2\leqq 3\times 48a_0^{-4}(p-2)\sum_{i\neq j}^{p}c_{ij}^2 ,$$

which implies that

(7) $$\sum_{i=1}^{p}(b_i^2-a_i^2)^2\leqq p^{-1}L_1\sum_{i\neq j}^{p}c_{ij}^2$$

for some $L_1>0$.

From (6) and the assumptions of (R.4) we have

$$(b_i^2-a_i^2)^2\leqq 48\times 3a_0^{-4}\eta_3^2$$

which implies that $|b_i|$'s are bounded and hence we have from (R.1) and (R.2) $|b_i - a_i| < a_0$, provided η_3 is very small. Therefore we have

$$|b_i + a_i| > a_0 ,$$

which implies that

$$(8) \qquad (b_i - a_i)^2 = (b_i + a_i)^{-2}(b_i^2 - a_i^2)^2 \leqq a_0^{-2}(b_i^2 - a_i^2)^2 \qquad \text{for } 1 \leqq i \leqq p .$$

Therefore (R.4) follows from (7) and (8). Q.E.D.

Let us prove Lemma 1. Since the assumption of this lemma is written as

$$p^{-1} \sum_{i=1}^{p} \Bigl(\sum_{\substack{j=1 \\ j \neq i}}^{p} c_{ij}^2 \Bigr) < \delta^2 ,$$

there exist at least two different integers k and l ($1 \leqq k, l \leqq p$) such that

$$\sum_{\substack{j=1 \\ j \neq k}}^{p} c_{kj}^2 < 2\delta^2 \qquad \text{and} \qquad \sum_{\substack{j=1 \\ j \neq l}}^{p} c_{lj}^2 < 2\delta^2 ,$$

which imply that

$$(9) \qquad \begin{aligned} &|c_{kj}| < 2\delta \qquad \text{for every } j \text{ with } 1 \leqq j \leqq p \text{ and } j\ (\neq k), \quad \text{and} \\ &|c_{lj}| < 2\delta \qquad \text{for every } j \text{ with } 1 \leqq j \leqq p \text{ and } j\ (\neq l) . \end{aligned}$$

Take $2\delta < 2^{-1}a_0^2$. Then using (R.3), we have from (9)

$$(b_j^2 - a_j^2)^2 \leqq 48 \times 3a_0^{-4} \times (2\delta)^2 \qquad \text{for } 1 \leqq j \leqq p ,$$

which implies that there exists $L_2 > 0$ such that

$$(10) \qquad |b_i| < L_2 \qquad \text{for } 1 \leqq i \leqq p$$

and that

$$(11) \qquad |b_1 - a_1| < 24a_0^{-3}\delta$$

in view of (R.1). We can assume that $k \neq 1$ without loss of generality. Then using (R.2), we have from (9)–(11)

$$(12) \qquad |b_k - a_k| < a_0^{-1}(12a_0^{-3}L_2 + 1) \times 2\delta \quad (= L_3 \times 2\delta, \text{ say}) .$$

Using (R.2) again, we have from (9), (10) and (12)

$$(13) \qquad |b_j - a_j| < a_0^{-1}(L_2L_1 + 1) \times 2\delta \qquad \text{for } 1 \leqq j \leqq p \text{ and } j\ (\neq k) .$$

From (10), (12), (13) and (A.1) there exists a scalar $L_4 > 0$ such that

$$(14) \qquad |c_{ij}| = |b_ib_j - a_ia_j| < L_4\delta \qquad \text{for } 1 \leqq i,\ j \leqq p \text{ and } i \neq j .$$

Therefore Lemma 1 follows from (14) and (R.4). Q.E.D.

PROOF OF LEMMA 2. Since the objective function (2.2) is written as

$$\|\Sigma-S\|^2=\sum_{i\neq j}^{p}(a_ia_j-s_{ij})^2+\sum_{i=1}^{p}(a_i^2+u_i^2-s_{ii})^2\,,$$

we have

$$v_i^2=\begin{cases} s_{ii}-b_i^2 & \text{if } b_i^2<s_{ii}\,, \\ 0 & \text{if } b_i^2\geqq s_{ii}\,. \end{cases} \tag{15}$$

When $b_i^2<s_{ii}$, it follows from (15) and (A.1) that

$$v_i^2-u_i^2=(s_{ii}-1)-(b_i^2-a_i^2)\,. \tag{16}$$

When $b_i^2\geqq s_{ii}$, it follows from (15) and (A.1) that

$$0\leqq u_i^2-v_i^2\leqq 1-b_i^2+|b_i^2-a_i^2|\leqq|1-s_{ii}|+|b_i^2-a_i^2|\,. \tag{17}$$

From (16) and (17) we have

$$(v_i^2-u_i^2)^2\leqq 2\{(s_{ii}-1)^2+(b_i^2-a_i^2)^2\}$$

for every i. Therefore the proof of Lemma 2 (i) is complete.

Let us prove Lemma 2 (ii). When $b_i^2<s_{ii}$, we have from (16)

$$(s_{ii}-1)^2\leqq 2\{(v_i^2-u_i^2)^2+(b_i^2-a_i^2)^2\}\,. \tag{18}$$

When $b_i^2\geqq s_{ii}$ and $s_{ii}>1$, we have

$$(s_{ii}-1)^2\leqq(b_i^2-1)^2\leqq(b_i^2-a_i^2)^2\,. \tag{19}$$

When $b_i^2\geqq s_{ii}$ and $s_{ii}\leqq 1$, we have from (15) and (A.4)

$$(s_{ii}-1)^2\leqq 1<u_0^{-4}(v_i^2-u_i^2)^2\,. \tag{20}$$

It follows from (18)-(20) that

$$(s_{ii}-1)^2\leqq 2(b_i^2-a_i^2)^2+u_0^{-4}(v_i^2-u_i^2)^2$$

for every i, which proves Lemma 2 (ii).

Acknowledgement

The author is indebted to Professor Masashi Okamoto and the referees for careful reading of earlier drafts of this manuscript and many helpful comments and corrections. He also thanks Dr. Yasuyuki Toyooka for suggesting the framework of this paper and Mr. Masamori Ihara of Osaka University for his discussions.

Osaka University

References

[1] Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis, *Proc. Third Berkeley Sym. Math. Prob.*, **5**, 111-150.

[2] Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariance structures, *South Afr. Statist. J.*, **8**, 1-24.

[3] Fukutomi, K. (1973). On the adequacy of factor extractions, *TRU Math.*, **9**, 119-136.

[4] Kano, Y. (1983). Consistency of estimators in factor analysis, *J. Japan Statist. Soc.*, **13**, 137-144.

[5] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications* (2nd Ed.), Wiley, New York.

[6] Tumura, Y. and Fukutomi, K. (1970). On the improper solutions in factor analysis, *TRU Math.*, **6**, 63-71.

[7] Williams, J. S. (1978). A definition for the common-factor analysis model and the elimination of problems of factor score indeterminacy, *Psychometrika*, **43**, 293-306.