The Philippine Statistician
1992, Vol. 41,
Pages 9-17

SENSITIVITY ANALYSIS IN ALPHA FACTOR ANALYSIS

by

Zenaida F. Mateo
Graduate School of Natural Sciences and Technology
Okayama University

Yoshimasa Odaka
Information Processing Center
Okayama University of Science

Yutaka Tanaka
Department of Statistics
Okayama University

Abstract

There are several sensitivity analysis procedures which have been considered by Tanaka and Odaka (1989 a, b, c) to investigate the phenomena of how a small change of data affects the outcome of factor analysis. Among them are the principal factor analysis (PFA), maximum likelihood factors analysis (MLFA) and least square factor analysis (LSFA). This motivates us to show that a similar technique can also be utilized to develop the sensitivity analysis in alpha factor analysis (AFA). Some examples are explained to illustrate the present procedure and a comparison is made in particular with the cases of PFA and MLFA.

Key Words and Phrases: alpha factor analysis, generalizability, psychometric concept, influence functions, universe common factor, singular value decomposition.

1. Introduction

In this study, a similar method used by Tanaka and Odaka (1989 a, b, c) was applied to develop the sensitivity analysis procedure in alpha factor analysis (AFA, Kaiser and Caffrey, 1965) which is based upon the psychometric concept of generalizability. The basic idea of AFA is to determine the common factor f in such a way that they have maximum correlation with the corresponding universe common factors. Alpha factors, like ML factors, have the property of invariance for scale transformation, that is, the same factors are found regardless of the units of measurements of the observable variables and loadings are proportional to scaling constants.

Our main objective in this study is to evaluate the influence of a small change of data on the values of the unique variance matrix D and the matrix $T^{*}=LL^{T}$. It is interesting to note that the influence function $I(\underline{x};\sum_{s=1}^{q}\lambda_s \underline{v}_s \underline{v}_s^T)$ can be used effectively to develop a sensitivity analysis procedure in AFA as well as in PFA and MLFA. Comparison is made with PFA and MLFA with respect to their sensitivities to small changes in data.

AFA is one of the popular methods of factors analysis. In facts, it is often used by research workers including psychologists, and is implemented in major statistical packages such as SAS and SPSS. There are some fields in biometry such as psychiatry which have similar conditions as psychology to be natural to asssume underlying assumptions of AFA. Therefore, we consider that it is worth to study AFA and develop a procedure of sensitivity analysis in AFA, even if we already have such procedures in PFA, MLFA and LSFA.

2. Alpha Factor Analysis

Let us assume the ordinary factor analysis (FA) model for a p x 1 observation vector $\underline{x}$ given by

$$\underline{x} = \underline{\mu} + L\underline{f} + \underline{e} \tag{1}$$

where $\underline{\mu}$ is the mean vector, L is a p x q ($q < p$) factor loading matrix, $\underline{f}$ is a q x 1 common factor score vector, and $\underline{e}$ is a p x 1 unique factor score vector. And also, we assume

$$E(\underline{f}) = 0, \quad E(\underline{e}) = 0$$

$$E(\underline{f}\underline{f}^T) = I, \quad E(\underline{e}\underline{e}^T) = \Delta, \quad E(\underline{f}\underline{e}^T) = 0$$

Δ denoting a diagonal matrix. When the assumed model holds, the covariance matrix Σ of the random vector $\underline{x}$ is expressed in terms of the p x q loading matrix L and the unique matrix p x p diagonal matrix Δ as

$$\Sigma = LL^T + \Delta \tag{2}$$

which is known as the common factor decomposition. That is, the covariance matrix Σ of $\underline{x}$ is decomposed into two parts, one explained by common factors and the other explained by unique factors. A number of methods have been proposed for estimating L and Δ from the observed Σ. Among them are PFA, MLFA, and LSFA.

In AFA, the generalized Kuder-Richardson reliability coefficient

$$\alpha = \frac{p}{p-1}\left[1 - \frac{\underline{w}^T H \underline{w}}{\underline{w}^T(\Sigma - \Delta)\underline{w}}\right] \tag{3}$$

is maximized to establish a model which represents universe batteries, where H = diag(Σ-Δ) is called communalities. We can obtain this by solving the generalized eigenvalue problem such

$$[(\Sigma - \Delta) - \lambda H]\underline{w} = 0 \tag{4}$$

which amounts to obtaining eigenvalues and eigenvectors of the following ordinary eigenvalue problem

$$\left[H^{-1/2}(\Sigma - \Delta)H^{-1/2} - \lambda I\right]\underline{v} = 0, \quad \underline{v} = H^{-1/2}\underline{w} \tag{5}$$

where $V_1 = (\underline{v}_1, \ldots, \underline{v}_q)$ is a matrix of unit-length column eigenvectors associated with the q largest eigenvalues $l_1, \ldots, l_q$ and L_1 = diag($l_1, \ldots, l_q$) represents the diagonal matrix of these values. From (3), the relationship between the eigenvalue l_s and the coefficient of generalizability α_s is expressed by

$$\alpha_s = \frac{p}{p-1}\left[1 - \frac{1}{\lambda_s}\right] \tag{6}$$

If Δ or H is known, L and Δ are obtained directly using the above procedure. However, if it is unknown, we must apply an iterative procedure by taking some initial trial for H or Δ. In this way, the solution satisfies the following determining equations.

$$R^* = \Sigma - \Delta \tag{7}$$
$$H^{-1/2}R^*H^{-1/2} = V_1\Lambda_1V_1^T + V_2\Lambda_2V_2^T \text{: spectral decomposition} \tag{8}$$
$$L = H^{1/2}V_1\Lambda_1^{1/2},\ T^* = LL^T = H^{1/2}V_1\Lambda_1V_1^TH^{1/2} \tag{9}$$
$$\Delta = \mathrm{diag}(\Sigma - T^*) \tag{10}$$

Note that PFA and MLFA have similar determinating equations with $H^{-1/2}$ in (8) and (9) replaced by I and $\Delta^{-1/2}$, respectively. Thus, we can regard that these three procedures commonly try to decompose the variances and covariances due to common factors $\Sigma - \Delta$, but have different methods of scaling before decomposition. Though the variables are not scaled in PFA based on the covariance matrix, they are scaled by the square roots of their variances in PFA based on the correlation matrix, of only the common portion of their variances in AFA and only of the unique portion of the variances in MLFA. The three methods with the exception of PFA based on the covariance matrix have the property invariance for scale-transformation.

3. Influence functions for the common and unique variance matrices

Our concern here is to study the change of the outputs Δ and $T^* = LL^T$ when a small change Σ to $\Sigma + e\Sigma^{(1)}$ is introduced to those determining equations. To do this, we replace the unperturbed Σ, Δ, R^* and T^* in determining the equations by the perturbed counterparts $\Sigma + \varepsilon\Sigma^{(1)}$, $\Delta + \varepsilon\Delta^{(1)}$, $R^* + \varepsilon R^{*(1)}$ and $T^* + \varepsilon T^{*(1)}$, respectively and compare the first order terms of ε. Then we apply the lemma on the influence function for $\Sigma_{s=1}^{q}\lambda_s \underline{v}_s\underline{v}_s^T$ which was derived by Tanaka (1988), we obtained the following equations. When the perturbation $F \to F = (1-\varepsilon)F + \varepsilon d_x$ (F: unpertubed cdf of x; δ_x: the cdf of a unit point mass at $\underline{x}$) is introduced on the distribution function, the coefficients of ε, *i.e* $\Sigma^{(1)}$, $\Delta^{(1)}$ and $T^{*(1)}$, denote the influence functions or influence curves for Σ, Δ and T^* respectively. It is known that $\Sigma^{(1)}$ is expressed as $\Sigma^{(1)} = (\underline{x} - \underline{m})(\underline{x} - \underline{m})^T - \Sigma$.

$$R^{*(1)} = \Sigma^{(1)} - \Delta^{(1)} \tag{11}$$

$$\begin{aligned} T^{*(1)} = {} & \tfrac{1}{2}R_D^{*(1)}H^{-1/2}TH^{1/2} + \tfrac{1}{2}H^{1/2}TH^{-1/2}R_D^{*(1)} \\ & - \sum_{s=1}^{q}\sum_{r=q}^{p}\underline{v}_s^T\Big[\tfrac{1}{2}R_D^{*(1)}H^{-3/2}R^*H^{-1/2} + \tfrac{1}{2}H^{-1/2}R^*H^{-3/2}R_D^{*(1)} \\ & - H^{-1/2}R^{*(1)}H^{-1/2}\Big]\underline{v}_rH^{1/2}\underline{v}_s\underline{v}_r^TH^{1/2} \\ & - \sum_{s=1}^{q}\sum_{r=q+1}^{p}\lambda_s(\lambda_s - \lambda_r)^{-1}\underline{v}_s^T\Big[\tfrac{1}{2}R_D^{*(1)}H^{-3/2}R^*H^{-1/2} \\ & + \tfrac{1}{2}H^{-1/2}R^*H^{-3/2}R_D^{*(1)} - H^{-1/2}R^{*(1)}H^{-1/2}\Big]\underline{v}_r \\ & \left(H^{1/2}\underline{v}_s\underline{v}_r^T + H^{1/2}\underline{v}_r\underline{v}_s^TH^{1/2}\right) \end{aligned} \tag{12}$$

$$\Delta^{(1)} = \mathrm{diag}(\Sigma^{(1)} - T^{*(1)}) \tag{13}$$

where the subscript "D" indicates diagonal. Now considering that the relationship

$$T_{jj}^{*(1)} = R_{jj}^{*(1)}, j = 1,\ldots,p \tag{14}$$

holds using equations (11) and (13) we can evaluate the system of equations (11) through (13) in the following four general steps to derive the influence functions $\Delta^{(1)}$ and $T^{*(1)}$. Note that the elements of vectors and matrices are denoted by attaching the subscripts to the characters indicating the vectors and matrices.

Step 1. Calculate $\Sigma_{jk}^{(1)}$ and obtain $R_{jk}^{*(1)}$ for $j{\neq}k$ by $R_{jk}^{*(1)}{=}\Sigma_{jk}^{(1)}$, $j{\neq}k$

Step 2. Solve the simultaneous linear equations for $R_{jj}^{*(1)}$s as follows.

$$\left\{1-\left(H^{-\frac{1}{2}}TH^{\frac{1}{2}}\right)_{jj}\right\}R_{jj}^{*(1)}-\sum_{i=1}^{p}a_{(j)i}R_{ii}^{*(1)}=\sum_{\substack{i=1\\ i\neq i'}}^{p}\sum_{i'=i}^{p}b_{(j)ii'}R_{ii'}^{*(1)},\quad j=1,\dots,p$$

where

$$\begin{aligned}a_{(j)i} &= \sum_{r=1}^{q}\sum_{s=1}^{q}\left\{-\frac{1}{2}v_{is}(QV)_{ir}-\frac{1}{2}(QV)_{is}v_{ir}+H_{ii}^{-1}v_{is}v_{ir}\right\}H_{jj}v_{js}v_{jr}\\ &+2\sum_{s=1}^{q}\sum_{r=q+1}^{q}\lambda_s(\lambda_s-\lambda_r)^{-1}\left\{-\frac{1}{2}v_{is}(QV)_{ir}-\frac{1}{2}(QV)_{is}v_{ir}+H_{ii}^{-1}v_{is}v_{ir}\right\}H_{jj}v_{js}v_{jr}\\ b_{(j)ii'} &= \sum_{s=1}^{q}\sum_{r=1}^{q}H_{ii}^{-\frac{1}{2}}H_{i'i'}^{-\frac{1}{2}}v_{is}v_{i'r}v_{js}v_{jr}\\ &+2\sum_{s=1}^{q}\sum_{r=q+1}^{q}\lambda_s(\lambda_s-\lambda_r)^{-1}H_{ii}^{-\frac{1}{2}}H_{i'i'}^{-\frac{1}{2}}v_{is}v_{ir}v_{js}v_{jr}\quad,\\ Q &= H^{-\frac{1}{2}}R^{*}H^{-\frac{1}{2}}\end{aligned}$$

Step 3. Calculate $\Delta_{jj}^{(1)}$ by $\Delta_{jj}^{(1)}{=}R_{jj}^{*(1)}$ - $\Sigma_{jj}^{(1)}$ $j=1,\dots,p$.
Step 4. Calculate $T^{*(1)}$ for $j{\neq}k$ by equation (12) using $R_{jj}^{(1)}$ and $\Delta_{jj}^{(1)}$.

Now, one of the advantages of this method is that we only need to evaluate a system of linear equations with p unkowns (n + 1) times (once for considering the whole data and n times for the data with one observation omitted) compared to applying the usual factor analysis (n + 1) times. Furthermore, as cited by Tanaka and Castano-Tostado (1990) that it is sufficient to solve the system of equations $(1+p(p+1)/2)$ times instead of n + 1 times. Note that the above procedure, the theoretical influence functions are defined in terms of population parameters Δ and $T^* = LL^T$, although in actual analysis it is practical to use sample versions of empirical influence curves (EIC) $\hat{\Delta}^{(1)}$ and $T^{\circ(1)}$ which are calculated by substituting $S = n^{-1}\sum_{j=1}^{n}(\underline{x}_j-\bar{\underline{x}})(\underline{x}_j-\bar{\underline{x}})^T$ and $S^{(1)} = (\underline{x}_i-\bar{\underline{x}})(\underline{x}_i-\bar{\underline{x}})^T - S$ into Σ and $\Sigma^{(1)}$ in the above procedure.

4. Numerical Examples

Example 1. Stock Price data

As an illustration of our procedure we applied our method to the set of stock-price data (Johnson, R. and Wichern, D., 1988, Table 8.1). It consists of 100 weekly rates of return for five stocks (Allied Chemical, DuPont, Union Carbide, Exxon, and Texaco). The weekly rates of return are defined as (current Friday closing price - previous Friday closing price)/(previous Friday closing price) adjusted for stocks splits and dividends. The observations in 100 successive weeks appear to be independently distributed but the rates of return across stocks are correlated since as one expects, stocks tend to move together in response to general economic conditions. Table 1 shows the result of the factor analysis with AFA assuming a two factor model. The varimax rotated loadings are displayed in Table 1 along with the unique variances. The rotated loadings indicate that the chemical stocks (Allied Chemical, DuPont, and Union Carbide) load highly on the first factor, while the oil stocks (Exxon and Texaco) load highly on the second factor. The two rotated factors suggest the difference between the industries. Although it is very difficult to label these factors, we can say that factor 1 respresents unique economic forces that cause chemical stocks to move together while factor 2 appears to represent the economic condition affecting oil stocks.

Next, to investigate the influence of each individual, we calculated the empirical influence curves $\hat{\Delta}^{(1)}$, $\hat{T}^{*(1)}$ and scalar-valued measures $\|\hat{\Delta}^{(1)}\|$ and $\|\hat{T}^{*(1)}\|$ by using the proposed procedure. Similarly like in the studies of Tanaka and Odaka (1989 a, b) the Euclidean norm for $\Delta^{(1)}$ and $T^{*(1)}$ are used as influence measures to study the influence behavior of each observation. Figure 1 shows the index plots of these summarized scalar-valued influence measures indicating the influence of every observation. The visual inspection of the index plots of the 100 weekly ratess of return indicates that observation number 56 is the most influential followed by observation number 13 for both proposed measures.

On the other hand, Table 2 shows the result of the AFA for the data with 56th observation omitted. It is very clear that the influence of the omission of this observation is not small as revealed in the loadings of the first and second variables. The remarkable point is that the omission of this observation causes an improper solution as indicated by the unique variance of the second variable.

Table 1. Result of the AFA
(Stock-Price Data, n=100; two factor model)

Variable	Factor loading	Factor loading	Unique Var.
1	0.5919	0.3663	0.5154
2	0.8551	0.1622	0.2425
3	0.6427	0.3569	0.4596
4	0.3632	0.5116	0.6063
5	0.2073	0.8771	0.1877

(judgement of convergence = 0.00001; 48 iterations)

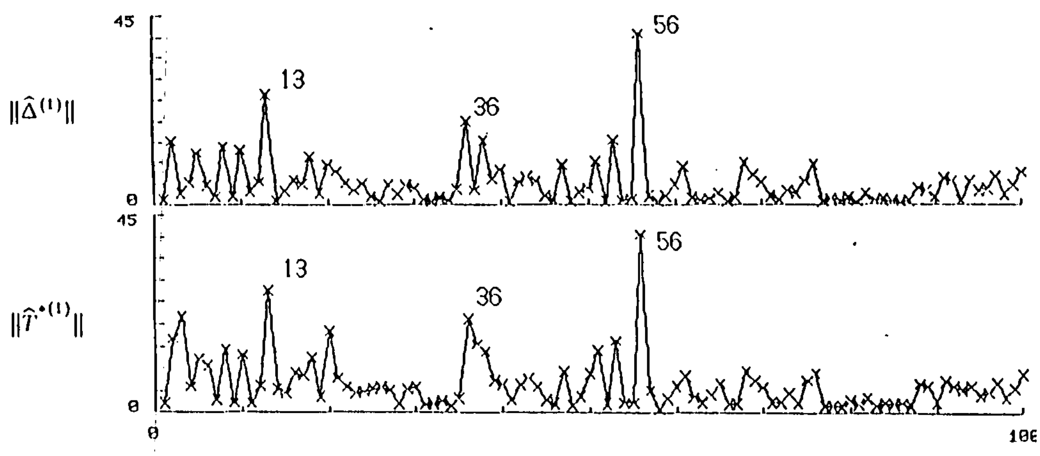

Fig. 1: Index plots of $\|\hat{\Delta}^{(1)}\|$ and $\|\hat{T}^{*(1)}\|$ (Stock -price data)

Example 2. Audiometric data (Jackson, J.E. 1991, Table 5.1)

The second example with *audiometry*, a study carried on within the Eastman Kodak Company involving the measurement of hearing loss. It was conducted to distinguish the differences between normal and induced hearing loss. The data consists of n = 100 males, age 39 who were made to measure their hearing level by means of an instrument called *audiometer* wherein the threshold measurements are calibrated in units called *decibel loss* in comparison to a reference standard for that particular instrument. Observations are registerd one at a time for a number of frequencies. The frequencies used are 500 Hz,

Table 2. Result of the AFA
(Stock-Price Data, n=99; observation 56 omitted)

Variable	Factor loading	Factor loading	Unique Var.
1	0.4987	0.4724	0.5281
2	0.9819	0.1921	0.0001
3	0.5539	0.4977	0.4455
4	0.2633	0.6098	0.5588
5	0.1632	0.7736	0.3748

(Note: In this case the iterative process went into the improper region, i.e $\hat{\Delta}_{22}<0$.Table 2 shows the converged values of the iterative process with fixed. $\hat{\Delta}_{22}$, i.e. $\hat{\Delta}_{22} = 0.0001$.) .

1000 Hz, 2000 Hz, and 4000 Hz which results in an eight-variable problem, considering two ears. The different factor loadings of the eight variables with their respective unique variances utilizing a two factor model are seen in Table 3. The result shows that variables 3, 4, 7 and 8 load heavily on the first factor. On the other hand, variables 1, 2, 5 and 6 load highly on the second factor. The interpretation of the factor loadings is the most straightforward if each variable loads highly on at most one factor, and if all the factor loadings are either large and positive or near zero, with few intermediate values. the variables then split into disjoint sets, each of which is associated with one factor, and perhaps some variables are left over. In this example, the two factors suggest a *contrast* between the high frequencies (2000 Hz and 4000 Hz) and the low frequencies (500 Hz and 1000 Hz) with respect to hearing level. We may regard the first factor as "high frequency" effect and the second factor as the "low frequency" effect. It is well known in the case of normal having that hearing loss as a function of age is first noticeable in the higher frequencies.

To study the influence pattern of every observation, we calculated the empirical influence curve $\hat{\Delta}^{(1)}$, $\hat{T}^{*(1)}$ and their corresponding scalar-valued measures. The index plots of $\|\hat{\Delta}^{(1)}\|$ and $\|\hat{T}^{*(1)}\|$ are shown in figure 2. Based on this figure, we are able to find the influential observations such as respondent number 75 exhibiting a relatively large influence on $\hat{\Delta}$ as compared to other observations. Table 4 shows the estimated L and Δ for the data set without the 75*th* observation, which is the most influential on $\hat{\Delta}^{(1)}$. As seen from the table, the estimates of the unique variances differ considerably from those for the whole data especially for variables 1, 2 and 5.

Table 3. Result of the AFA
(Audiometric Data, n=100; two factor model)

Variable	Factor loading	Factor loading	Unique Var.
1	0.1635	0.8494	0.2518
2	0.2848	0.8002	0.2786
3	0.6464	0.3792	0.4384
4	0.7171	0.1025	0.4796
5	0.0586	0.7378	0.4521
6	0.2718	0.7935	0.2965
7	0.5838	0.2554	0.5940
8	0.7240	0.0403	0.4742

(judgement of convergence = 0.00001; 22 iterations)

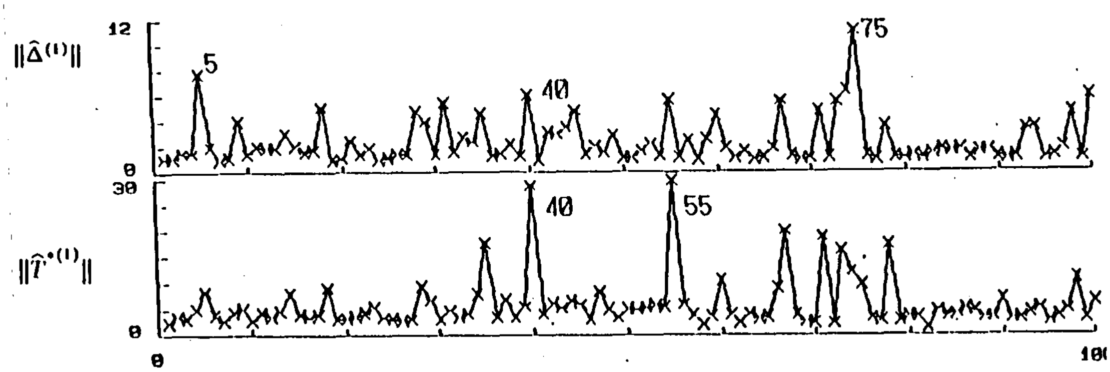

Fig. 2: Index plots of $||\hat{\Delta}^{(1)}||$ *and* $||\hat{T}^{*(1)}||$ *(Audiometric data)*

Table 4. Result of the AFA
(Audiometric Data, n=99; no. 75 omitted; two factor model)

Variable	Factor loading	Factor loading	Unique Var.
1	0.1722	0.8332	0.2761
2	0.3131	0.8148	0.2380
3	0.6592	0.3765	0.4238
4	0.7192	0.0935	0.4740
5	0.0223	0.7928	0.3709
6	0.2805	0.7826	0.3038
7	0.5831	0.2518	0.5965
8	0.7180	0.0349	0.4833

(judgement of convergence = 0.000001; 21 iterations)

5. Discussion

In this study, we have proposed a method for obtaining $\hat{\Delta}^{(1)}$ and $\hat{T}^{*(1)}$ representing the empirical influence functions for the two components $\Delta^{(1)}$, T^{*} of the common variance decomposition.

The second point of attention is the numerical verification of the validity of the quantities $\hat{\Delta}^{(1)}_{ij}$ or $EIC_i(\Delta_{ij})$ for both data. To accomplish this purpose, we disregard a single observation in turn, perform AFA n times and calculate $\hat{\Delta}_{(i)}$, which denotes the estimate of Δ based on the sample with the *i - th* observation deleted. Then the sample influence curve SIC_i's is evaluated by using the relation

$$SIC_i = -(n-1)(\hat{\Delta}_{(i)} - \hat{\Delta}), \quad i = 1,\ldots,n \tag{15}$$

and compare them with $EIC_i(\Delta)$'s based on the present procedure. Then, we draw the scatter diagrams of $EIC_i(\Delta)$ versus $SIC_i(\Delta)$ of the respective norms. As an illustration, Figure 3 gives us a pciture of the scatter diagrams of EIC versus SIC for the Euclidean norm using the stock price-data and audiometric data. as suggested form these figures, the correspondence is good enough for us to conclude that the quantities EIC_i's based on the present method can be used practically instead of SIC_i's. Besides, in terms

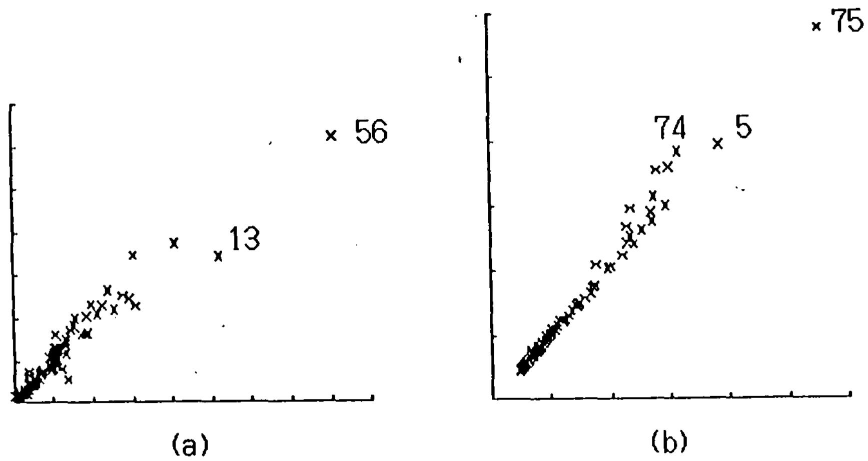

Fig. 3: Scatter diagram of EIC (horizontal) versus SIC (vertical) for the Euclidean norm: (a) Stock-price data; (b) Audiometric data

of computing time SIC_i requires much time to compute than EIC_i. In other words, we can deduce that EIC can be used effectively for detecting influential observations since it has a very high correlation with SIC which has the clear "leave-one-out" interpretation. In fact, we also observed a similar scenario as in the case of PFA and MLFA (Tanaka and Onada, 1989 a, b). Once EIC is obtained, it can be used for detecting influential subsets of observations as discussed by Tanaka, Castano-Tostado and Odaka (1990) and Moon, Yanagi and and Tanaka (1992).

Next a comparison of AFA is made with PFA and MLFA based on the scalar-valued measures $||\hat{\Delta}^{(1)}||$ and $||\hat{T}^{*(1)}||$ for both examples. In the case of stock data, Figure 4 shows the graph of the norm of $\hat{\Delta}^{(1)}$ using PFA, MLFA and AFA. As displayed from the figure, we can see that alpha factor solution is fairly close to principal factor and maximum likelihood analysis procedure based on the proposed measures. This example illustrates that sensitivities to smaller changes of data are almost equal in the three procedures. However, using the audiometric data set, the scatter diagram in Figure 5 gives us a different picture. As seen from the figure, AFA procedure is somewhat different compared to MLFA and PFA solution based on the norm of $\hat{\Delta}^{(1)}$. Notice also that the points are located almost at random for the three procedures.

Note that in the above four steps, the influence funtions $\Delta^{(1)}$ and $\hat{T}^{*(1)}$ are both linear functions of $\Sigma^{(1)}$. From this property, the influence of an arbitrary point can be decomposed into a finite number of components which will provide a more efficient tool for comparison of the different estimation methods. Although, we did not apply it in out present work, it is worth mentioning that this will be our next target of study.

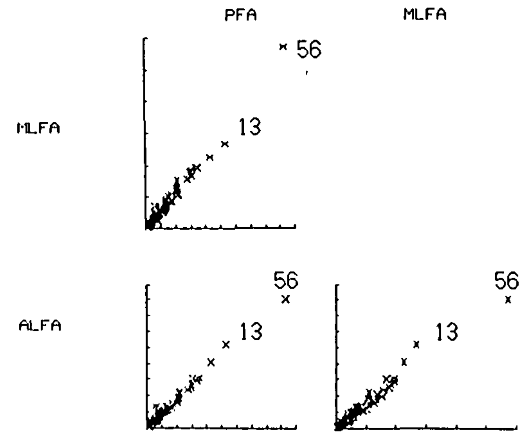

Fig. 4: Scatter diagram of MLFA versus PFA, AFA versus PFA and AFA versus MLFA based Euclidean norm: (Stock-price data)

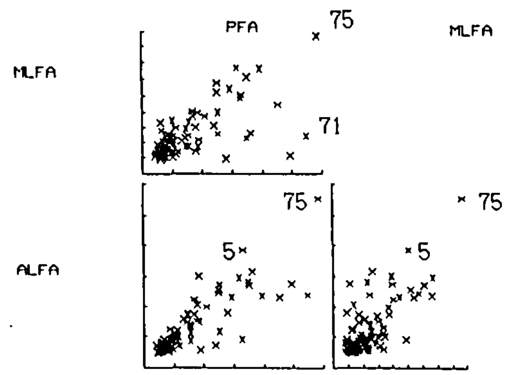

Fig. 5: Scatter diagram of MLFA versus PFA, AFA versus PFA and AFA versus MLFA based on Euclidean norm: (audiometric data)

References

Jackson, J. E. **A User's Guide to Principal Components.** John Wiley and Sons, Inc.,1991.

Johnson, R. A. and Wichern, D. W. **Applied Multivariate Statistical Analysis.** 2nd ed. Prentice-Hall, 1988.

Kaiser, H. F. and Caffrey, J. Alpha factor analysis. **Psychometrika**, 21, 1-14.

Mardia, K. V., Kent, J. T. and Bibby, J. M. **Multivariate Analysis**. Academic Press, 1979.

Moon, S. H., Yanagi, K. and Tanaka, Y. A Clustering Algorithm for Detecting Influential Subsets in Multivariate Methods. To be submitted for publication, 1992.

Tanaka, Y. Sensitivity analysis in principal component analysis: Influence on the subspace spanned by principal components. **Comm. Statist.**, A17, 3157-75.