Journal of the Korean
Statistical Society
Vol. 25, No. 1, 1996

Resistant Principal Factor Analysis†

Yong-Seok Choi and Ho-Seon Byun [1]

Abstract

Factor analysis is a multivariate technique for describing the interrelationship among many variables in terms of a few underlying but unobservable random variables called factors. There are various approaches for this factor analysis. In particular, principal factor analysis is one of the most popular methods. This follows the mathematical algorithm of the principal component analysis based on the singular value decomposition. But it is known that the singular value decomposition is not resistant, i.e., it is very sensitive to small changes in the input data. In this article, using the resistant singular value decomposition of Choi and Huh (1994), we derive a resistant principal factor analysis relatively little influenced by notable observations.

Key Words : Factor analysis; Principal component analysis; Singular value decomposition; Resistant.

---

†The present studies were supported (in part) by the Matching Fund Programs of Research Institute for Basic Sciences, Pusan National University, Korea, 1994, Project No. RIBS-PNU-94-103.

[1]Department of Statistics, Pusan National University, Pusan, 609-735, Korea.

1. INTRODUCTION

Factor analysis is a multivariate technique for describing the interrelationship among many variables in terms of a few underlying but unobservable random variables called factors.

The origin of factor analysis lies in the early twentieth-century attempts of Charles Spearman. Much of the early development of factor analysis was done by psychologists seeking a better understanding of the dimensions of human intelligence (Jobson, 1992, p. 388).

For this factor analysis, there are various alternative approaches. In particular, principal factor analysis is one of the most popular methods. The algorithm for this approach is mathematically equivalent to that of principal component analysis. It is well known that the main spirit of principal component analysis is dimension reduction. In fact, this can be done by the eigensystem or singular value decomposition. And there are reasons for preferring the use of the singular value decomposition which is one of the most useful methods in the areas of matrix computation (Choi and Huh, 1994).

In this paper, firstly in Section 2, we provide the principal factor analysis using the singular value decomposition. And we give the geometric interpretations of factor loadings plot. This plot geometrically gives the important interpretations of factor analysis. However Choi and Huh (1994) point out that the singular value decomposition of the data matrix is not resistant, i.e., it is very sensitive to small changes in the input data. Therefore, if there exist outliers in data matrix, the principal factor analysis using the singular value decomposition does not give the desirable results. And they developed the resistant singular value decomposition.

In Section 3, we provide a resistant version of principal factor analysis based on the resistant singular value decomposition. And we call this a resistant principal factor analysis. In addition, as the analogy with Section 2, we consider the geometric interpretations of the resistant factor loadings plot.

We give two numerical illustrations and discussions in Section 4. Finally, Section 5 gives the concluding remarks.

2. PRINCIPAL FACTOR ANALYSIS

2.1. The classical principal factor analysis using the singular value decomposition

Firstly, we review the factor analysis model. Let $\mathbf{x}' = (x_1, \ldots, x_p)$ be a $p \times 1$ random vector of p variables with mean vector $\mu' = (\mu_1, \ldots, \mu_p)$ and $p \times p$ covariance matrix Σ. Generally, the $m(\ll p)$-factor analysis model for $\mathbf{x}$ can be written in the matrix form

$$\mathbf{x} = \mu + \mathbf{Lf} + \varepsilon \tag{2.1}$$

where $\mathbf{L}$ is the $p \times m$ matrix of factor loadings l_{jk}, $j = 1, \ldots, p$; $k = 1, \ldots, m$, $\mathbf{f}$ is the $m \times 1$ vector of linearly independent common factors $f_k, k = 1, \ldots, m$ and ε is the $p \times 1$ vector of unique or specific factors $\varepsilon_j, j = 1, \ldots, p$.

The common factors are assumed to have mean 0 and variance 1. The specific factors are assumed to have mean 0 and specific variance ψ_j, $j = 1, \ldots, p$. In addition, it is assumed that all of the common factors are uncorrelated with the specific factors.

Given these assumptions, the covariance matrix Σ can be given by

$$\Sigma = \mathbf{LL}' + \Psi \tag{2.2}$$

where $\mathbf{E}[(\mathbf{x} - \mu)(\mathbf{x} - \mu)'] = \Sigma$, $\mathbf{E}[\varepsilon\varepsilon'] = \Psi$ and $\Psi = diag(\psi_1, \ldots, \psi_p)$ is a $p \times p$ specific variance matrix with diagonal elements ψ_j, $j = 1, \ldots, p$. Generally, the equation (2.2) is called the common factor decomposition of Σ (Tanaka, 1989).

In factor analysis model of (2.1), the general interesting problem is how to estimate $\mathbf{L}$ and Ψ (and hence $\Sigma = \mathbf{LL}' + \Psi$) from a p-variate sample variance-covariance matrix $\mathbf{S}$. Johnson and Wichern (1992, pp. 403-406) provided the algorithm for the classical principal factor analysis based on the eigensystem of $\mathbf{S}$. As noted in Section 1, the singular value decomposition is the more useful than the eigensystem in the areas of matrix computation. Now we provide the classical principal factor analysis using the singular value decomposition.

Consider the $n \times p$ variables-centered data matrix $\widetilde{\mathbf{X}} = (x_{ij} - \bar{x}_{.j})$, where $\bar{x}_{.j} = \sum_{i=1}^{n} x_{ij}/n$ and $i = 1, \ldots, n; j = 1, \ldots, p$. Naturally, it leads to a p-variate sample variance-covariance matrix

$$\mathbf{S} = \widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}/n. \tag{2.3}$$

And the singular value decomposition of matrix $\widetilde{\mathbf{X}}$ with rank r can be written

$$\widetilde{\mathbf{X}} = \mathbf{UD}_\lambda\mathbf{V}' = \sum_{k=1}^{r} \lambda_k \mathbf{u}_k \mathbf{v}_k{}' \tag{2.4}$$

where $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$ and $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_r)$ are $n \times r$, $p \times r$ matrices with orthogonal columns $\mathbf{u}_k$ and $\mathbf{v}_k$, $k = 1, \ldots, r$, respectively and $\mathbf{D}_\lambda = diag(\lambda_1, \ldots, \lambda_r)$ with singular values $\lambda_1 \geq \cdots \geq \lambda_r$.

Now as the analogous pattern with algorithm of Johnson and Wichern (1992, pp. 403-406), we provide an algorithm for the estimation of $\mathbf{L}$ and Ψ using the singular value decomposition of (2.4) as :

STEP 1: *We obtain the largest* $m (\ll r \leq p)$ *singular values,* $\lambda_1 \geq \cdots \geq \lambda_m$ *and the corresponding right singular vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_m$.

STEP 2: *We form a* $p \times m$ *matrix of estimated factor loadings* $\tilde{\mathbf{L}}_{(m)}$ *as*

$$\tilde{\mathbf{L}}_{(m)} = n^{-1/2}(\lambda_1 \mathbf{v}_1, \ldots, \lambda_m \mathbf{v}_m) = (\tilde{\mathbf{l}}_1, \ldots, \tilde{\mathbf{l}}_p)'$$

where $\tilde{\mathbf{l}}'_j = n^{-1/2}(\lambda_1 v_{j1}, \ldots, \lambda_m v_{jm})$, $j = 1, \ldots, p$.

STEP 3: *We take the estimated specific variances provided by the diagonal elements of the matrix* $\mathbf{S} - \tilde{\mathbf{L}}_{(m)}\tilde{\mathbf{L}}'_{(m)}$, *so*

$$\tilde{\Psi} = diag(\tilde{\psi}_1, \ldots, \tilde{\psi}_p)$$

where $\tilde{\psi}_j = s_{jj} - \tilde{\mathbf{l}}'_j\tilde{\mathbf{l}}_j$, s_{jj} *is the* j^{th} *diagonal element of* $\mathbf{S}$ *in* (2.3) *and* $\tilde{\mathbf{l}}'_j\tilde{\mathbf{l}}_j$ *is usually called the communality of the* j^{th} *variable,* $j = 1, \ldots, p$.

2.2 Geometric interpretations of the factor loadings plot

Now we consider the geometric interpretations of factor loadings plot. In fact, in STEP 2 of an algorithm for the classical principal factor analysis using the singular value decomposition, the rows $\tilde{\mathbf{l}}'_j$ $(j = 1, \ldots, p)$ of $\tilde{\mathbf{L}}_{(m)}$ give the coordinates of this factor loadings plot. Since $\tilde{\mathbf{L}}_{(m)}$ is algebraically equivalent to $\mathbf{H}$ of h-plot (Choi, 1995), we easily show that its geometric interpretations can be applied to those of the factor loadings plot.

Therefore the geometric interpretations of the factor loadings plot are as follows:

$$\begin{aligned} s_{jk} &\simeq \tilde{\mathbf{l}}'_j\tilde{\mathbf{l}}_k, \\ s_{jj} &\simeq \|\tilde{\mathbf{l}}_j\|^2, \\ r_{jk} &\simeq \cos(\theta). \end{aligned}$$

where "$\simeq$" denotes "approximation", s_{jk} is the sample covariance between the j^{th} and the k^{th} variables, s_{jj} is the sample variance of the j^{th} variable and is approximated by the squared length of $\tilde{\mathbf{l}}_j$ and r_{jk} is the correlation between the j^{th} and the k^{th} variables and θ is the angle between the rows $\tilde{\mathbf{l}}'_j$ and $\tilde{\mathbf{l}}'_k$.

Finally, we need a measure for the goodness of approximation of the factor loadings plot. For appreciating the goodness of approximation, there are a few heuristic methods (Johnson and Wichern, 1992, p. 406; Jobson, 1992, pp. 394-395).

Here as the analogous tool with Choi (1995), we use a measure for goodness of approximation

$$\begin{aligned} \rho_{(m)} &= 1 - \|\mathbf{S} - \tilde{\mathbf{L}}_{(m)}\tilde{\mathbf{L}}'_{(m)}\|^2/\|\mathbf{S}\|^2 \\ &= \sum_{k=1}^{m} \lambda_k^4 / \sum_{k=1}^{r} \lambda_k^4 \end{aligned} \tag{2.5}$$

In addition, we consider the residual matrix

$$\mathbf{S} - (\tilde{\mathbf{L}}_{(m)}\tilde{\mathbf{L}}'_{(m)} + \widetilde{\mathbf{\Psi}}) \tag{2.6}$$

resulting from the approximation of **S**. Generally, the diagonal elements shall be zero. And if the off-diagonal elements are small, we may subjectively take the optimal m-factor analysis model.

3. RESISTANT PRINCIPAL FACTOR ANALYSIS

3.1. The resistant principal factor analysis using the resistant singular value decomposition

In Section 2, we provided the principal factor analysis using the singular value decomposition. However Choi and Huh (1994) point out that the singular value decomposition of the $n \times p$ variables-centered data matrix $\widetilde{\mathbf{X}}$ is not resistant. Therefore if there exist notable observations in data matrix, classical principal factor analysis using the singular value decomposition does not give desirable results.

Choi and Huh (1994, Theorem) provide the resistant singular value decomposition of an $n \times p$ data matrix $\widetilde{\mathbf{X}}^*$ of rank r centered at a robust location estimate. Calculation of the resistant singular value decomposition can be done using the iterative procedure with Andrew's $\psi(\cdot)$ function given by

$$\psi(t) = \begin{cases} c\sin(t/c), & \text{for } 0 \le t < c\pi \\ 0, & \text{for } t \ge c\pi \end{cases}$$

where c is determined by $(c\pi)^2 = \chi^2_{0.95(p-m)}$, $\chi^2_{0.95(p-m)}$ is 95 percentile point of χ^2 distribution with $p - m$ degrees of freedom and m is the number of common factors.

We note that the resistant singular value decomposition of $\widetilde{\mathbf{X}}^*$ can be written as

$$\widetilde{\mathbf{X}}^* = \mathbf{U}\mathbf{D}_{\lambda^*}\mathbf{V}' \tag{3.1}$$

where $\mathbf{U}$ is an $n \times r$ matrix such that $\mathbf{U}'\mathbf{D}_\omega\mathbf{U} = \mathbf{I}_r$, $\mathbf{V}$ is a $p \times r$ matrix of eigenvectors of $\widetilde{\mathbf{X}}^{*\prime}\mathbf{D}_\omega\widetilde{\mathbf{X}}^*$ such that $\mathbf{V}'\mathbf{V} = \mathbf{I}_r$ and $\mathbf{D}_{\lambda^*} = diag(\lambda_1{}^*, \ldots, \lambda_r{}^*)$ with the k^{th} eigenvalue $\lambda_k{}^{*2}$ of $\widetilde{\mathbf{X}}^{*\prime}\mathbf{D}_\omega\widetilde{\mathbf{X}}^*$.

In fact, $\mathbf{D}_\omega = diag(\omega_1, \ldots, \omega_n)$ is an $n \times n$ diagonal matrix with the diagonal elements $\omega_i = \psi(\| \tilde{\mathbf{x}}_i^* - \hat{\mathbf{x}}_i^* \| /\hat{\sigma})/(\| \tilde{\mathbf{x}}_i^* - \hat{\mathbf{x}}_i^* \| /\hat{\sigma})$, $i = 1, \ldots, n$. Here, $\tilde{\mathbf{x}}_i^*$ denotes the i^{th} row of $\widetilde{\mathbf{X}}^*$ and can be viewed as n points in a p-dimensional space $\mathcal{R}^p$. And let $\hat{\mathbf{x}}_i^*$ in a subspace of dimension $m(1 \le m \le p)$ of $\mathcal{R}^p$ be the nearest point of an arbitrary point $\tilde{\mathbf{x}}_i^*$ in $\mathcal{R}^p$. We use the median scale estimator as $\hat{\sigma} = [\text{med}_i(\|\tilde{\mathbf{x}}_i^* - \hat{\mathbf{x}}_i^*\|^2)/\chi^2_{0.50(p-m)}]^{1/2}$, $i = 1, \ldots, n$, where $\chi^2_{0.50(p-m)}$ is 50 percentile point of χ^2 distribution with $p - m$ degrees of freedom. For more details, see Choi and Huh (1994).

Thus we obtain the weighted sample variance-covariance matrix from the resistant singular value decomposition of (3.1)

$$n^*\mathbf{S}^* = \widetilde{\mathbf{X}}^{*\prime}\mathbf{D}_\omega\widetilde{\mathbf{X}}^* \tag{3.2}$$

where $n^* = \sum_{i=1}^n w_i = \mathbf{1}_n'\mathbf{D}_\omega\mathbf{1}_n$. Of course, the algorithm based on the eigensystem can be applied to $\mathbf{S}^*$ in (3.2).

However we provide an algorithm for the resistant estimation of $\mathbf{L}$ and Ψ using the resistant singular value decomposition (3.1) as a strict analogy with an algorithm proposed in Subsection 2.1. This proceeds as follows.

STEP 1: *We obtain the largest $m(\ll r \le p)$ resistant singular values, $\lambda_1^* \ge \cdots \ge \lambda_m^*$ and the corresponding resistant right singular vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$.*

STEP 2: *We form a $p \times m$ matrix of estimated resistant factor loadings $\widetilde{\mathbf{L}}^*_{(m)}$ as*

$$\widetilde{\mathbf{L}}^*_{(m)} = n^{*-1/2}(\lambda^*_1 \mathbf{v}_1, \ldots, \lambda^*_m \mathbf{v}_m) = (\tilde{\mathbf{l}}^*_1, \ldots, \tilde{\mathbf{l}}^*_p)'$$

where $\tilde{\mathbf{l}}^{*\prime}_j = n^{*-1/2}(\lambda^*_1 v_{j1}, \ldots, \lambda^*_m v_{jm}),\ j = 1, \ldots, p.$

STEP 3: *We take the estimated resistant specific variances provided by the diagonal elements of the matrix $\mathbf{S}^* - \widetilde{\mathbf{L}}^*_{(m)}\widetilde{\mathbf{L}}^{*\prime}_{(m)}$, so*

$$\widetilde{\Psi}^* = diag(\tilde{\psi}^*_1, \ \ldots \ , \tilde{\psi}^*_p)$$

where $\tilde{\psi}^*_j = s^*_{jj} - \tilde{\mathbf{l}}^{*\prime}_j \tilde{\mathbf{l}}^*_j$, s^*_{jj} is the j^{th} diagonal element of $\mathbf{S}^*$ in (3.2) *and $\tilde{\mathbf{l}}^{*\prime}_j \tilde{\mathbf{l}}^*_j$ is called the resistant communality of the j^{th} variable, $j = 1, \ldots, p$.*

3.2 Geometric interpretations of the resistant factor loadings plot

Now consider the geometric interpretations of the resistant factor loadings plot.

As the analogy with Subsection 2.2, this plot is given by the rows $\tilde{\mathbf{l}}^*_j (j = 1, \ldots, p)$ of $\widetilde{\mathbf{L}}^*_{(m)}$ in STEP 2 of an algorithm for the resistant principal factor analysis. And the geometric interpretations of the resistant factor loadings plot are as follows:

$$\begin{aligned} s^*_{jk} &\simeq \tilde{\mathbf{l}}^{*\prime}_j \tilde{\mathbf{l}}^*_k, \\ s^*_{jj} &\simeq \|\tilde{\mathbf{l}}^*_j\|^2, \\ r^*_{jk} &\simeq \cos(\theta). \end{aligned}$$

Therefore we note that the geometric interpretations of the factor loadings plot discussed in Subsection 2.2 also can be applied to the resistant factor loadings plot.

Finally, with using $\mathbf{S}^*$ and $\widetilde{\mathbf{L}}^*_{(m)}$ instead of $\mathbf{S}$ and $\widetilde{\mathbf{L}}_{(m)}$ in (2.6), we use a measure for goodness of resistant approximation

$$\begin{aligned} \rho^*_{(m)} &= 1 - \|\mathbf{S}^* - \widetilde{\mathbf{L}}^*_{(m)}\widetilde{\mathbf{L}}^{*\prime}_{(m)}\|^2 / \|\mathbf{S}^*\|^2 \\ &= \sum_{k=1}^{m} \lambda^{*4}_k \Big/ \sum_{k=1}^{r} \lambda^{*4}_k \end{aligned} \tag{3.3}$$

Also we consider the residual matrix

$$\mathbf{S}^* - (\widetilde{\mathbf{L}}^*_{(m)}\widetilde{\mathbf{L}}^{*\prime}_{(m)} + \widetilde{\Psi}^*) \tag{3.4}$$

resulting from the approximation of $\mathbf{S}^*$. As noted in Subsection 2.2, if the off-diagonal elements are small, we may take the optimal resistant m-factor analysis model.

4. NUMERICAL ILLUSTRATIONS

Example 1 : The preferences data is the scores of eighteen students on their preferences for the following subjects : Mathematics, Physics, English, Natural Sciences, Foreign Language and History (Jambu, 1991, Table A.5, p. 443).

The 2-dimensional factor loadings plot is given in Fig. 1 with goodness of approximation 99.73%. We note that PHY(physics) and NSC(natural sciences) have a similar characteristic. So, their correlation is high and thus the angle among them must be small by the geometric interpretations of factor loadings plot given in Subsection 2.2. Also since FOR(foreign language) and HIS(history) are similar in characteristic, we expect that the angle among them becomes small. But these interpretations are not clear in Fig. 1.

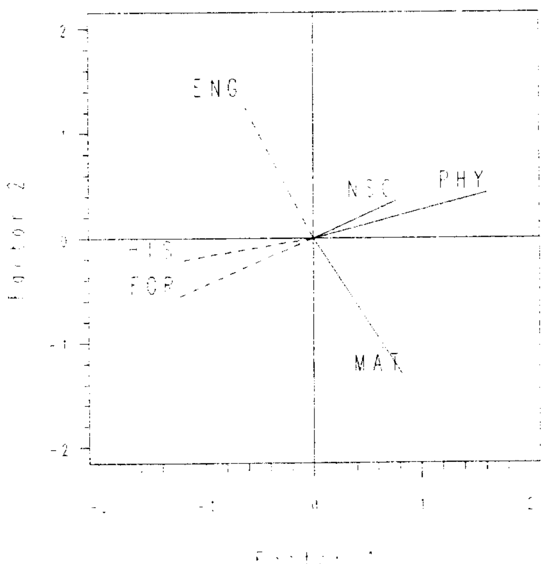
Now consider the resistant version of principal factor analysis. As noted in Subsection 3.1, we use the Andrew's $\psi(\cdot)$ function with c=0.98, where c is determined by $(c\pi)^2 = \chi^2_{0.95(4)}$. And we use 1.83 for the median scale estimate.

The final weights used in computing resistant singular value decomposition are in the diagonal matrix

$$\mathbf{D}_\omega = \text{diag}(0.998\ 0.998\ 0.000\ 1.000\ 0.998\ 0.000\\ 0.000\ 0.000\ 0.998\ 0.000\ 0.000\ 0.000\\ 0.992\ 0.995\ 0.188\ 0.992\ 0.992\ 0.000).$$

The 2-dimensional factor loadings plot for resistant principal factor analysis is shown in Fig. 2 with goodness of approximation 99.99%. By reducing the influence of the notable observations, Fig. 2 gives somewhat lucid interpretations of principal factor analysis. That is, the angle among FOR(foreign language) and HIS(history) is much smaller than that of Fig. 1. But the patterns of PHY(physics) and NSC(natural science) are not changed. We note that half the factor loadings are positive and half the factor loadings are negative on the first factor. A factor with this pattern of the factor loadings is called a bipolar factor (Johnson and Wichern, 1992, pp. 420-422).

Table 1 shows the estimated factor loadings, specific variances and cumulative proportion of total variance for the classical and resistant principal factor analyses. We note that specific variances are the portion of variances of each variable not explained by the first two common factors. In such a view, the resistant principal factor analysis gives much better results than those in classical principal factor analysis. That is, specific variances are small for the resistant principal factor analysis. Also note that cumulative proportion of total variance explained by the first two common factors is large for the resistant principal factor analysis. Thus we may call the first factor a "science-nonscience" factor.



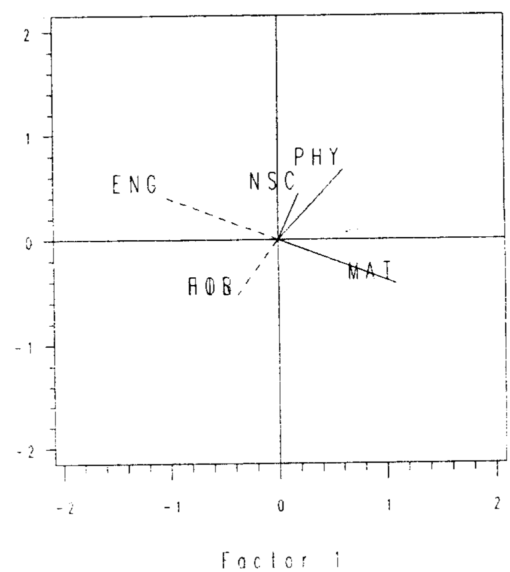


Fig. 1 Classical factor loadings plot **Fig. 2** Resistant factor loadings plot

for the preferences data

Table 1. Estimated factor loadings, specific variances and cumulative proportion of total variance given by classical and resistant methods

Variables	Classical			Resistant		
	l_{ij}		ψ_i	l^*_{ij}		ψ^*_i
MAT	0.83	-1.31	0.02	1.08	-0.43	0.00
PHY	1.61	0.43	0.12	0.60	0.67	0.00
ENG	-0.66	1.32	0.03	-1.08	0.43	0.00
NSC	0.76	0.35	0.28	0.19	0.45	0.00
FOR	-1.26	-0.57	0.09	-0.40	-0.56	0.00
HIS	-1.27	-0.22	0.09	-0.40	-0.56	0.00
Cumulative proportion	0.61	0.95		0.65	0.99	

Finally, we compare the residual matrices (2.6) and (3.4) as noted in Subsection 2.2 and Subsection 3.2, respectively. Note that the elements of the residual matrix in resistant principal factor analysis are much smaller than those of the residual matrix in classical principal factor analysis. The residual matrices are as follows :

$$\mathbf{S} - (\tilde{\mathbf{L}}\tilde{\mathbf{L}}' + \tilde{\Psi}) = \begin{pmatrix} 0 & -0.01 & 0.02 & -0.02 & -0.02 & 0.00 \\ & 0 & 0.02 & -0.17 & 0.04 & -0.01 \\ & & 0 & -0.05 & 0.01 & -0.02 \\ & & & 0 & -0.05 & 0.01 \\ & & & & 0 & -0.08 \\ & & & & & 0 \end{pmatrix}$$

$$\mathbf{S}^* - (\tilde{\mathbf{L}}^*\tilde{\mathbf{L}}^{*\prime} + \tilde{\Psi}^*) = \begin{pmatrix} 0 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ & 0 & 0.00 & 0.00 & 0.00 & 0.00 \\ & & 0 & 0.00 & 0.00 & 0.00 \\ & & & 0 & 0.00 & 0.00 \\ & & & & 0 & 0.00 \\ & & & & & 0 \end{pmatrix}$$

Example 2: The census-tract data (Johnson and Wichern, 1992, Table 8.2, p. 392) consists of fourteen tract informations on five socio-economic variables.

With goodness of approximation 99.60%, the 2-dimensional factor loadings plot is shown in Fig. 3. Since MSY(median school year) and MVH(median value home) have a similar pattern, we expect that they can be described as the same factor more clearly. Also POP(total population), TOE(total employment) and HSE(health services employment) have the same characteristic and so they must be described as the same factor more naturally. But these interpretations are not clear in Fig. 3.

Therefore, we consider the resistant principal factor analysis. As defined in Subsection 3.1, we use the Andrew's $\psi(\cdot)$ function with c=0.89, where c is determined by $(c\pi)^2 = \chi^2_{0.95(3)}$. And we use 1.54 for the median scale estimate.

We have the diagonal matrix used in computing the resistant singular value decomposition

$$\mathbf{D}_\omega = \text{diag}(0.000\ 0.000\ 0.905\ 0.427\ 0.756\ 0.808\ 1.000\\ 0.000\ 0.907\ 0.885\ 0.772\ 0.000\ 0.000\ 0.000).$$

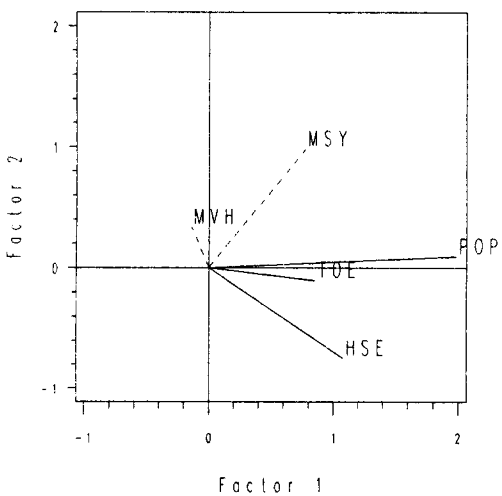
Now by reducing the influence of the notable observations, we obtain the 2-dimensional factor loadings plot in Fig. 4 with goodness of resistant approximation 99.99%. We note that the resistant principal factor loadings plot gives the more sufficient results. That is, the angle between MSY(median school year) and MVH(median value home) becomes small. So their patterns being described as factor 2 are more natural than those in Fig. 3. Also with respect to the geometric interpretations in Subsection 3.2, POP(total population), TOE(total employment) and HSE(health services employment)can be more clearly described as factor 1.

Table 2 shows the estimated factor loadings, specific variances and cumulative proportion of total variance for the classical and resistant principal factor analyses. We note that specific variances are smaller for the resistant principal factor analysis. And cumulative proportion of total variance explained by the first two common factors is large for the resistant principal factor analysis. Therefore, we note that resistant principal factor analysis gives the more satisfactory results.

Finally, as in Example 1, we can compare both methods(classical and resistant principal factor analyses) by residual matrices. The results show that the elements of the residual matrix in resistant principal factor analysis are much smaller than those in the classical principal factor analysis.

Table 2. Estimated factor loadings, specific variances and cumulative proportion of total variance given by classical and resistant methods

Variables	Classical			Resistant		
	l_{ij}		ψ_i	l^*_{ij}		ψ^*_i
POP	1.98	0.09	0.06	1.22	0.11	0.01
MSY	0.78	0.98	0.07	0.41	0.65	0.00
TOE	0.85	-0.11	0.01	0.57	-0.01	0.00
HSE	1.08	-0.75	0.10	1.00	-0.42	0.01
MVH	-0.14	0.34	0.34	0.08	0.17	0.01
Cumulative proportion	0.74	0.93		0.82	0.99	



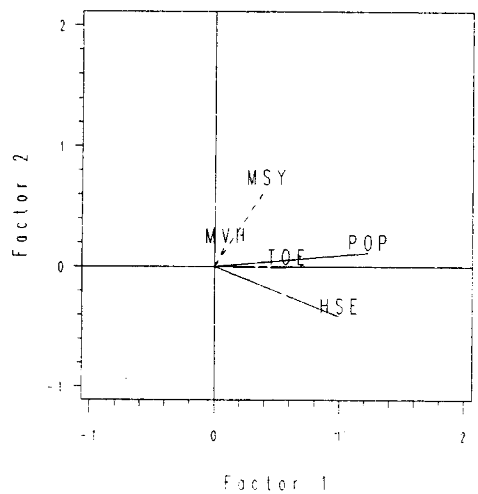


Fig. 3 Classical factor loadings plot **Fig. 4** Resistant factor loadings plot for the census-tract data

The residual matrices are as follows:

$$\mathbf{S} - (\tilde{\mathbf{L}}\tilde{\mathbf{L}}' + \tilde{\Psi}) = \begin{pmatrix} 0 & -0.06 & 0.00 & -0.07 & 0.01 \\ & 0 & -0.01 & 0.06 & -0.06 \\ & & 0 & -0.01 & 0.01 \\ & & & 0 & 0.07 \\ & & & & 0 \end{pmatrix}$$

$$\mathbf{S}^* - (\tilde{\mathbf{L}}^*\tilde{\mathbf{L}}^{*\prime} + \tilde{\Psi}^*) = \begin{pmatrix} 0 & -0.01 & 0.00 & -0.01 & 0.00 \\ & 0 & 0.00 & 0.00 & 0.00 \\ & & 0 & 0.00 & 0.00 \\ & & & 0 & 0.00 \\ & & & & 0 \end{pmatrix}$$

And in general, we note that since the data in this example involve measurements on different scales, they neeed to be pre-standardized.

5. CONCLUDING REMARKS

We provide a resistant version of principal factor analysis using the resistant singular value decomposition. We call this a resistant principal factor analysis. This approach seems to be more desirable with respect to the geometric interpretations of factor analysis, residual matrix and goodness of approximation.

In this paper, we limit ourselves to reducing the influence of outliers and obtaining the resistant version of principal factor analysis. Our goal is not to detect outliers in factor analysis. Finally, Johnson and Wichern (1992, pp, 409-411) discussed a modification of the principal factor analysis. This approach deserves a good further research area.

ACKNOWLEDGEMENTS

The authors are grateful to two anonymous referees for their kind comments and suggestions.

REFERENCES

(1) Choi, Y. S. and Huh, M. H. (1994), Resistant singular value decomposition and its applications. *Unpublished.*

(2) Choi, Y. S. (1995), Resistant h-plot for a sample variance-covariance matrix. *Journal of the Korean Statistical Society,* **24**, 407-417.

(3) Jambu, M. (1991), *Exploratory and Multivariate Data Analysis.* Academic Press, Inc.

(4) Jobson, J. D. (1992), *Applied Multivariate Data Analysis* Vol. II : *categorical and Multivariate Methods.* Springer-Verlag, New-York.

(5) Johnson, R. A. and Wichern, D. W. (1992), *Applied Multivariate Statistical Analysis.* Prentice-Hall, New Jersey.

(6) Tanaka, Y. (1989), Influential observations in principal factor analysis. *Psychometrika*, **3**, 475-485.