

SOME ROBUST METHODS OF ESTIMATION IN FACTOR ANALYSIS IN THE PRESENCE OF OUTLIERS

G. M. Oyeyemi and R. A. Ipinyomi

Department of Statistics
University of Ilorin
Ilorin, Nigeria

ABSTRACT

A robust method of estimating covariance matrix in multivariate data set is proposed. The goal is to compare the proposed method with the most widely used robust methods (Minimum Volume Ellipsoid and Minimum Covariance Determinant) and the classical method (MLE) in area of Factor Analysis. MVE- MCE- and Proposed-factor analyses use robust covariance matrix in estimating the factor loadings while the classical factor analysis estimates the factor loadings using the MLE of the covariance matrix. While classical factor analysis is found to be more reliable when there are few or no outliers in the data, robust factor analysis will be preferred in the presence of multiple outliers. The Proposed Robust Factor analysis is also found to compete favourably well with existing robust methods.

KEYWORDS: Covariance Matrix, Communalities, Factor Loadings, Minimum Covariance Determinant, Minimum Volume Ellipsoid, Mean Square Error, Optimality criteria.

INTRODUCTION

Let $X = \{x_1, x_2, x_3, \dots, x_m\}$ be a set of m points in $\mathbb{R}^p$, where $x_i, i = 1, 2, 3, \dots, m$ are independent and identically distributed multivariate normal $N_p(\mu, \Sigma)$. The usual Maximum Likelihood Estimate (MLE) method of estimating μ and Σ are the sample mean vector and sample covariance matrix $\bar{x}$ and S respectively as

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p]'$$

$$S = \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})(x_i - \bar{x})', \text{ i.e. } S = \begin{bmatrix} S_{11} & S_{12} & S_{1j} & \dots & S_{1p} \\ S_{21} & S_{22} & S_{2j} & \dots & S_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & S_{pj} & \dots & S_{pp} \end{bmatrix}$$

Covariance matrix plays a prominent role in multivariate data analysis. It measures the spread of the individual variables as well as level of inter-relationship (intercorrelation) that may exist between pairs of the variables in the multivariate data. The statistic is used in many multivariate techniques as a measure of spread and intercorrelation among variables. Such techniques include multivariate regression model, Principal Component and Factor Analyses, Canonical Correlation, and Linear Discriminant and even in Multivariate Statistical Process Control (MSPC).

It is well known that the Maximum Likelihood Estimate (MLE) method can be very sensitive to deviations from the assumptions made on the data, in particular, to unexpected outliers in the data (Vandev and Neykov 2000). To overcome this problem, many robust alternatives to Maximum Likelihood Estimator (MLE) have been developed in recent years. All the methods converged on tackling the problem of robust estimation by finding a sufficiently large subset of uncontaminated (free of outliers) of the data. Such subset will be mainly elements of the true population and estimation is then based on this subset.

When the estimate of the covariance matrix of a multivariate data is not robust (Biased) it often times renders the technique or the analysis it is used for to be inconsistent and in most cases found to be unreliable. For instance, in Multivariate Control chart like Hotelling T^2 chart, where the presence of multiple outliers can affect the estimation of the covariance matrix using Maximum Likelihood Estimation (MLE) method to the extent that all the outliers will go undetected by the chart (Vargas 2003). The same effect can be experienced in other analyses like Principal Component and Factor analysis and other multivariate techniques.

Outliers can heavily influence the estimation of the covariance matrix Σ and subsequently the parameters or statistics that are needed to be derived from it. Hence a robust estimate of the covariance matrix that will not be affected by outliers is required to obtain valid and reliable results (Hubert and Engelen, 2007). There have been many robust methods of estimating the covariance matrix of a multivariate data. Such methods include Minimum Volume Ellipsoid (MVE), Minimum Covariance Determinant (MCD), S-Estimator, M-Estimator and Orthogonalized Gnanadesikan-Kettering (OGK) methods.

1. ROBUST METHODS

1.1. MINIMUM VOLUME ELLIPSOID (MVE)

The Minimum Volume Ellipsoid (MVE) estimator is first proposed by Rousseeuw (1984). It has been studied extensively for non-control chart settings and frequently used in detection of multivariate outliers. The estimation seeks to find the ellipsoid of minimum volume that covers a subset of at least h data points. The subset of size h is called halfset

because h is often chosen to be just more than half of the m data points. The location estimator is the geometrical center of the ellipsoid and the estimator of the variance-covariance matrix defining the ellipsoid itself multiplied by an appropriate constant to ensure consistency (Rousseeuw and van Zomeren (1990), Rousseeuw and van Zomeren (1991), and Rocke and Woodruff (1998)).

Assuming that we have a multivariate data set containing m samples, $\{x_i \in \mathbb{R}^p\}_{i=1}^m$. In order to solve the MVE problem we need to obtain a $p \times p$ positive definite matrix $C \in \mathbb{R}^{p \times p}$ and the centre of the ellipsoid t so as to maximize $\det(C^{-1})$ subject to $(x_i - t)^T C^{-1} (x_i - t) \leq p$ (Titterton, 1975).

The MVE for the data set $\{x_i\}_{i=1}^m$ must go through at least $p + 1$ and at most h support vectors. Thus the MVE estimates of the location and dispersion do not correspond to the sample mean vector and sample variance-covariance matrix of a particular halfset. For more detailed discussion on MVE one may see Daves (1987), Lopuhaa and Rousseeuw (1991), Titterton (1975) and Agullo, 1996).

1.2. MINIMUM COVARIANCE DETERMINANT (MCD)

An alternative high breakdown estimation procedure to the MVE is an estimator based on the Minimum Covariance Determinant (MCD), which was first proposed by Rousseeuw (1984). It is obtained by finding the halfset of multivariate data points that gives the minimum value of the determinant of the covariance matrix. The resulting estimator of location is the sample mean vector of the points that is the halfset and the estimator of the dispersion is the sample covariance matrix of the points multiplied by an appropriate constant to ensure consistency just as was done for MVE.

The MCD estimators are intuitively appealing because a small value of the determinant corresponds to near linear dependencies of the data in the p -dimensional space. That is because a small determinant corresponds to a small eigenvalue which suggests a near linear dependency that suggests that there is a group of points that are similar to each other (Jensen *et al.* 2002).

Let $p < m/2$, let $X = \{x_1, x_2, x_3, \dots, x_p\}$ be a set of n points in $\mathbb{R}^p$. Let h be a natural number, $m/2 < h < m$. The Minimum Covariance Determinant problem for X and h , MCD for short, is the problem to find an h -elements set $X^h = \{x_{i_1}, x_{i_2}, \dots, x_{i_h}\} \in X$ such that $\det(X^h)$ is minimal overall h -element sets. The empirical covariance matrix $C(X^h)$, with minimal determinant yields a robust estimate S of the scatter matrix, with $S = S(X^h) = C_0 C(X^h)$, where C_0 is a suitably chosen constant to achieve consistency. The estimate of the location parameter is given as;

$$t = t(X^h) = \frac{1}{h} \sum_{x \in X^h} x$$
 of all the h points in the set X^h . The pair (t, S) is called the MCD-estimate with respect to X .

2. THE PROPOSED ROBUST METHOD (PRM)

Given a p -dimensional multivariate normal data $X_{p \times m}$ with m observations $\{x_i\}_{i=1}^m$. Our interest is to obtain a subset of $\{x_i\}_{i=1}^m$ of size $k = p + 1$ that will satisfy some optimality criteria. Therefore we sample without replacement a sample of size k from m , this will give C_{p+1}^m possible subsets of size $p + 1$, $x_{j1}, x_{j2}, \dots, x_{jp+1}, \{x_j\}_{j=1}^{p+1}$. If we denote each subset by J_j , $j = 1, 2, \dots, C_{p+1}^m$.

For each J_j , we estimate the variance-covariance matrix $C_j = \frac{1}{p+1} (x_j - \bar{x}_j)(x_j - \bar{x}_j)^T$ and for each of the $p \times p$ matrix C_j , the characteristic roots or eigenvalues $e_{j1}, e_{j2}, \dots, e_{jp}$ are obtained and from such eigenvalues we calculate the following optimality criteria;

E_A = The Minimum of the minimum eigenvalues. $E_A = \min\{\min e_i\}$

E_P = The Minimum of the product of the eigenvalues.
$$E_P = \min \left\{ \prod_{i=1}^p e_i \right\}$$

E_H = The Minimum of the harmonic mean of the eigenvalues.
$$E_H = \min \left\{ \sum_{i=1}^p \frac{1}{e_i} \right\}^{-1}$$

The objective is to obtain data points $(p + 1)$ such that its variance-covariance matrix will satisfy all the three optimality criteria. Such covariance matrix will be inflated or deflated to accommodate good data points among the observed data. The resulting variance-covariance matrix is then multiplied by a constant for consistency.

2.1. THE ALGORITHM FOR OBTAINING THE PROPOSED METHOD

Let $X = \{x_1, x_2, x_3, \dots, x_m\}$ be a set of m points in $\mathcal{R}^p$. Let h be a natural number such that $m/2 < h < m$.

1. We select $p + 1$ data points $\{x_1, x_2, \dots, x_{p+1}\}$ that satisfy the three optimality criteria.
2. Use such $p + 1$ data points to obtain the center and variance-covariance matrix

$$\bar{x}_* = \frac{1}{p+1} \sum_{i=1}^p x_i, \text{ and } S_* = \frac{1}{p+1} (x_* - \bar{x}_*)(x_* - \bar{x}_*)^T \text{ respectively.}$$

3. Calculate the Mahalanobis distances for all the m observations using $\bar{x}_*$ and S_* as

$$d_i^2 = (x_i - \bar{x}_*) S_*^{-1} (x_i - \bar{x}_*)^T$$

4. The d_i^2 ($i = 1, 2, 3, \dots, m$) is arranged in order of magnitude from the least to the highest.

5. The first $p+j$ ($j = 2, 3, 4, \dots, h-p-1$) distances are selected and their corresponding sample units (points) are used to compute the next $\bar{x}_*$ and S_* as follows;

$$\bar{x}_* = \frac{1}{p+j} \sum_{i=1}^{p+j} x_i, \text{ and } S_* = \frac{1}{p+j} (x_* - \bar{x}_*)(x_* - \bar{x}_*)^T \text{ respectively.}$$

6. The new set of $\bar{x}_*$ and S_* are then used to obtain the Mahalanobis distances for all the observations.

7. Steps 4, 5, and 6 are repeated until the number of units selected is $h = \frac{m+p+1}{2}$.

8. The proposed robust estimators are then given as $\bar{x}_{prop} = \frac{1}{h} \sum_{i=1}^h x_i$ and $S_{prop} = \frac{1}{h} (x_* - \bar{x}_*)(x_* - \bar{x}_*)^T$ where $x_* = \{x_1, x_2, \dots, x_h\}$.

3. FACTOR ANALYSIS

Factor analysis is another area of application of covariance matrix in multivariate data analysis. Factor Analysis is a standard technique in multivariate analysis which is routinely used in social and behavioural sciences and most especially when dealing with high dimensional data (Kramar, 1972). The main goal of factor analysis is to summarize or explain the variable structure in the observed variables in terms of the underline latent variables. These observable variables are believed to be highly correlated (Timm, 1985).

Given a p -variate multivariate normal random variables $X_1, X_2, X_3, \dots, X_p, X_i \in \mathcal{R}^p$, that is $X_p \sim N(\mu, \Sigma)$ factor analysis will obtain a $k < p$ unobservable or latent variables $f_1, f_2, f_3, \dots, f_k$ which are called factors and are linked with the original variables through the equation

$$X = \Lambda F + \varepsilon \quad (3.1)$$

X is as defined earlier, Λ is a $(p \times k)$ matrix of factor loadings, both F and ε are vectors of factors (latent or unobservable variables) and error terms respectively (See Timm, 1987).

Equation (3.1) can be written in an explicit form as follows;

$$\begin{aligned} X_1 &= \lambda_{11}f_1 + \lambda_{12}f_2 + \lambda_{13}f_3 + \dots + \lambda_{1k}f_k + \varepsilon_1 \\ X_2 &= \lambda_{21}f_1 + \lambda_{22}f_2 + \lambda_{23}f_3 + \dots + \lambda_{2k}f_k + \varepsilon_2 \\ &\vdots \\ X_p &= \lambda_{p1}f_1 + \lambda_{p2}f_2 + \lambda_{p3}f_3 + \dots + \lambda_{pk}f_k + \varepsilon_p \end{aligned} \quad (3.2)$$

The error terms $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_p$ are independent but with specific variances $\psi_1, \psi_2, \psi_3, \dots, \psi_p$ respectively. Other assumptions of the model are that $E(F) = E(\varepsilon) = 0$ and $Cov(\varepsilon) = \psi$, where ψ is a diagonal matrix containing on its diagonal the specific variances. Also ε and F are independent.

3.1. ESTIMATION OF PARAMETERS

From equation (3.1), one needs to estimate both Λ and Ψ , which can be obtained by decomposing the variance-covariance matrix Σ as follows:

$$\Sigma = \Lambda\Lambda^1 + \Psi \quad (3.3)$$

If the random variables $X_1, X_2, X_3, \dots, X_p$ are standardized, then the correlation matrix R is used in place of Σ . In factor analysis, the objective is to find the estimates $\hat{\Lambda}$ and $\hat{\Psi}$ satisfying the condition that $\hat{\Lambda}^1\Psi^{-1}\hat{\Lambda}$ is diagonalised. That is,

$$S = \hat{\Lambda}\hat{\Lambda}^1 + \hat{\Psi} \quad (3.4)$$

Given an estimate of Λ , it is natural to set

$$\hat{\psi}_{ii} = s_{ii} - \sum_{j=1}^k \lambda_{ij}^2, \quad i = 1, 2, 3, \dots, p; \quad j = 1, 2, 3, \dots, k \quad (3.5)$$

so that the diagonal matrix $\hat{\Lambda}^1\Psi^{-1}\hat{\Lambda}$ holds exactly. Setting $\Sigma = S = \hat{\Lambda}\hat{\Lambda}^1 + \hat{\Psi}$, we get

$$\hat{\sigma}_{ii} = s_{ii} - \sum_{j=1}^k \hat{\lambda}_{ij}^2 + \hat{\psi}_{ii} \quad (3.6)$$

If the correlation matrix is used, then the model is written as $R = \hat{\Lambda}\hat{\Lambda}^1 + \hat{\Psi}$ and equation (3.6) becomes $\hat{\psi}_{ii} = 1 - \sum_{j=1}^k \hat{\lambda}_{ij}^2$, since the diagonal elements of R is 1.

There are various methods of estimating the parameters Λ and Ψ , these methods include principal component, maximum likelihood, principal factoring methods among others. (Kramer, 1972; Norman, 1987). Principal component method is used in this work.

3.2. PRINCIPAL COMPONENT METHOD

The method makes use of principal component analysis approach in extracting the factors. Suppose the variables have been standardized such that the scatter matrix is estimated by the correlation matrix R . To obtain the estimates of Λ and Ψ , that is the matrices of factor loadings and specific variances respectively based on a sample of n observations of the p -variate multivariate normal distribution. We assume that k , the number of factors and a reasonable initial estimate $\hat{\Psi}$ of Ψ are available. Then we let

$$R - \hat{\Psi}_0 = \hat{\Lambda}_0\hat{\Lambda}_0^1 \quad (3.7)$$

Equation (3.7) is called reduced correlation matrix because the 1's in the diagonal matrix have been replaced by the estimated communalities. Replacing $R - \hat{\Psi}_0$ by R_r known as reduced correlation matrix, equation (3.7) suggests that a matrix $\hat{\Lambda}_0$ be found using principal component analysis such that $R_r = \hat{\Lambda}_0\hat{\Lambda}_0^1$. Matrix $\hat{\Lambda}_0$ is found by solving the determinantal equation

$$|R_0 - \delta I| = 0 \text{ which yields } \hat{\Lambda}_0 = P\Delta^{\frac{1}{2}}$$

where P contains the first k eigenvectors of R_r such that $P_i P_j^1 = 1$ for $i = j$ and 0 for $i \neq j$, and Δ is a diagonal matrix of the eigenvalues of R_r . Repeated principal factoring employs an iterative process. Having used $\hat{\Psi}_0$ to calculate $\hat{\Lambda}_0$, we compute new $\hat{\Psi}_1 = \text{Diag}(R - \hat{\Lambda}_0 \hat{\Lambda}_0^1)$ which is a new estimate for Ψ_1 . Then equation (1.07) becomes $R - \hat{\Psi}_1 = \hat{\Lambda}_1 \hat{\Lambda}_1^1$. The principal component analysis is applied repeatedly until the estimates for both of Λ and Ψ converge.

In all classical factor analysis methods, the variance-covariance matrix Σ (or the correlation matrix R) is estimated by the sample covariance matrix S . It is this estimate that is decomposed to obtain the estimates of Λ and Ψ . It is however known that outliers can heavily influence the estimation of the variance-covariance matrix Σ and consequently the estimates of Λ and Ψ . In order to remove the influence of outliers in the factor analysis, a robust method of estimating the variance-covariance matrix is inevitable.

Two of the most widely used robust methods of estimating the variance-covariance matrix of multivariate data which are Minimum Volume Ellipsoid and Minimum Covariance Determinant are compared with the proposed and classical methods when employed in estimating the variance-covariance matrix in factor analysis.

3.3. SIMULATION AND COMPARISON

Assuming we have a multivariate data set with number of observations $n = 100$ and with number of variables $p = 5$ such that $X_5 \sim N(\mu, \Sigma)$, where $\mu = [0 \ 0 \ - \ - \ 0]^1$ and Σ is variance-covariance matrix. The data are then contaminated with outliers at different percentages and magnitude. At each setting 1000 samples are simulated. Factor analysis is then used to extract $k = 2$ factors based on the variance-covariance matrix estimated by;

- Classical method.
- Minimum Volume Ellipsoid
- Minimum Covariance Determinant
- Proposed Robust Method

The resulting factor analyses are then called MVE-factor analysis, MCD-factor analysis, Classical factor analysis, and Proposed factor analysis respectively. If we denote $L \in \mathbb{R}^{5 \times 2}$ and $U \in \mathbb{R}^5$ as the factor loadings and unique variances of the factor analysis respectively, then $L_n^{(0)} \in \mathbb{R}^{5 \times 2}$ and $U_n^{(0)} \in \mathbb{R}^5$ are the factor loadings and unique variances of the uncontaminated data. While $L_n^{D1}, L_n^{D2}, L_n^{D3}$ and $L_n^{D4} \in \mathbb{R}^{5 \times 2}$ are the factor loadings and $U_n^{D1}, U_n^{D2}, U_n^{D3}$ and $U_n^{D4} \in \mathbb{R}^5$ are the unique variances of the disturbed (contaminated) data using MVE, MCD, Proposed and Classical factor analyses respectively.

Since the loadings matrix is only determined up to an orthogonal matrix, we consider the $p \times p$ matrix $A_n \in L_n L_n'$ where $A_n^{(0)} \in L_n^{(0)} (L_n^{(0)})'$ will be for undisturbed data and

$A_n^D \in L_n^D(L_n^{D'})'$ for the contaminated data. We shall compare the elements a_{ij}^D of A_n^D with the corresponding a_{ij}^0 of the matrix A_n^0 of the undisturbed data by computing the mean squared error (MSE) of the estimates;

$$MSE(a_{ij}) = \frac{1}{m} \sum_{i=1}^m (a_{ij}^{D(m)} - a_{ij}^0)^2 \text{ for } ij = 1, 2, 3, -, -, -, p$$

$$\text{The average MSE is therefore } MSE(A) = \frac{1}{p^2} \sum_{i=1}^p \sum_{j=1}^p MSE(a_{ij})$$

Similarly, for the square root of the unique variances U_j , we compute

$$MSE(U_j) = \frac{1}{m} \sum_{i=1}^m \left(\sqrt{U_j^{D(m)}} - \sqrt{U_j^{D(m)}} \right)^2$$

$$\text{And the average MSE is } MSE(U) = \frac{1}{p} \sum_{j=1}^p MSE(U_j), \text{ where } j = 1, 2, 3, -, -, -, p$$

Since we have set the number of observations to be 100 ($n = 100$), we introduce into the data 5, 10, 15, 20, 25 and 30 representing 5%, 10%, 15%, 20%, 25% and 30% outliers respectively. The levels or magnitudes of outliers used are $\sqrt{\chi_{p,\alpha}^2}$, $\chi_{p,\alpha}^2$ and $2\chi_{p,\alpha}^2$ where p is the number of variables and $\alpha = 0.95$. The results for varying magnitudes are shown in Tables 1 to 6.

4. DISCUSSION OF RESULTS AND CONCLUSIONS

The mean square errors (MSE) of factor loadings and unique variances of classical factor analysis are found to be smaller than those of any other three approaches to factor analysis used when the magnitude of outlier introduced to the data is small

($\sqrt{\chi_{p,0.95}^2} = 3.33$). For varying level of contamination, the classical approached will be preferred for small magnitude of outlier, because of its relatively smaller mean square error (Figures 1 and 4; Tables 1 and 4). For larger values of outliers ($\chi_{p,0.95}^2 = 11.09$ or $2\chi_{p,0.95}^2 = 22.18$), the other three methods; MVE- and MCD- and Proposed-factor analyses have the mean square error of their factor loadings and unique variances smaller than the classical factor analysis at various levels (percentages) of contamination (see Figures 2, 3, 5 & 6 and Tables 2, 3, 5 & 6).

It can be inferred from the graphs that the Proposed method competes favourably well with the existing robust methods (MCD and MVE) when there are multiple outliers and even for

large sizes of outliers. Though, the classical factor analysis is the best for fewer and smaller size of outliers, the proposed method performed better than the other two robust methods (MVE and MCD) because its mean square errors for both factor loadings and unique variances are less than any of the two methods.

Outliers in %	METHODS			
	Classical	MVE	MVD	Proposed
5%	0.0015	0.0050	0.0080	0.0045
10%	0.0049	0.0075	0.0100	0.0060
15%	0.0075	0.0115	0.0135	0.0085
20%	0.0120	0.0155	0.0175	0.0130
25%	0.0155	0.0200	0.0220	0.0165
30%	0.0185	0.0225	0.0265	0.0195

Table 1. The Mean Square Error of Factor Loadings when the centre is $\sqrt{\chi^2_{p,0.95}}$

Outliers in %	METHODS			
	Classical	MVE	MVD	Proposed
5%	0.0600	0.0195	0.0100	0.0315
10%	0.1000	0.0225	0.0185	0.0345
15%	0.1350	0.0235	0.0200	0.0365
20%	0.1650	0.0300	0.0245	0.0455
25%	0.1850	0.0345	0.0288	0.0485
30%	0.1990	0.0375	0.0340	0.0500

Table 2. The Mean Square Error of Factor Loadings when the centre is $\chi^2_{p,0.95}$

Outliers in %	METHODS			
	Classical	MVE	MVD	Proposed
5%	0.0600	0.0100	0.0100	0.0150
10%	0.1000	0.0185	0.0185	0.0200
15%	0.1350	0.0200	0.0200	0.0250
20%	0.1650	0.0245	0.0245	0.0285
25%	0.1850	0.0288	0.0288	0.0345
30%	0.1990	0.0340	0.0340	0.0395

Table 3. The Mean Square Error of Factor Loadings when the centre is $2\chi^2_{p,0.95}$

Outliers in %	METHODS			
	Classical	MVE	MVD	Proposed
5%	0.0010	0.0075	0.0065	0.0040
10%	0.0015	0.0095	0.0075	0.0060
15%	0.0025	0.0100	0.0085	0.0070
20%	0.0044	0.0115	0.0100	0.0085
25%	0.0052	0.0125	0.0120	0.0100
30%	0.0062	0.0135	0.0130	0.0125

Table 4. The Mean Square Error of Unique Variance when the centre is $\sqrt{\chi^2_{p,0.95}}$

Outliers in %	METHODS			
	Classical	MVE	MVD	Proposed
5%	0.0250	0.0005	0.0010	0.0040
10%	0.0580	0.0009	0.0030	0.0045
15%	0.0810	0.0020	0.0055	0.0068
20%	0.0954	0.0050	0.0085	0.0105
25%	0.1150	0.0140	0.0100	0.0125
30%	0.1200	0.0155	0.0150	0.0175

Table 5. The Mean Square Error of Unique Variance when the centre is $\chi^2_{p,0.95}$

Outliers in %	METHODS			
	Classical	MVE	MVD	Proposed
5%	0.1100	0.0050	0.0010	0.0030
10%	0.1650	0.0095	0.0030	0.0045
15%	0.1800	0.0025	0.0055	0.0088
20%	0.1950	0.0055	0.0085	0.0105
25%	0.2000	0.0145	0.0100	0.0155
30%	0.2090	0.0160	0.0150	0.0205

Table 6. The Mean Square Error of Unique Variance when the centre is $2\chi^2_{p,0.95}$

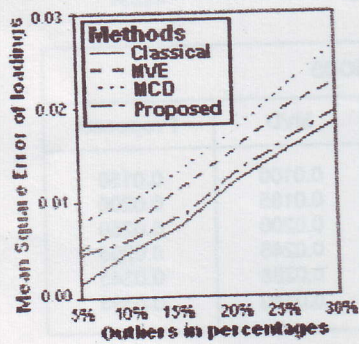


Figure 1. MSE of Loadings when centre = $\sqrt{\chi^2_{p,0.95}}$

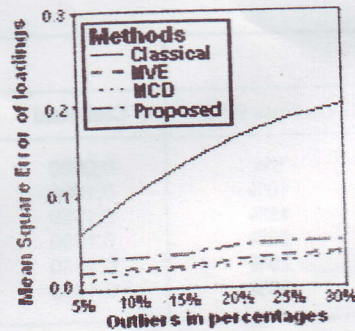


Figure 2. MSE of Loadings when centre = $\chi^2_{p,0.95}$

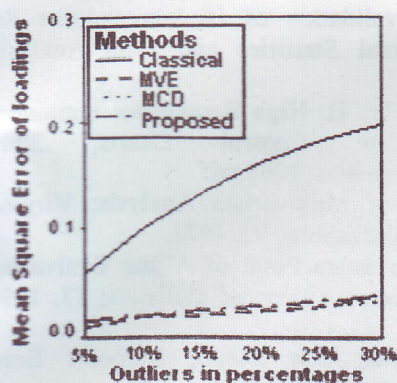


Figure 3. MSE of Loadings
when centre = $2\chi^2_{p,0.95}$

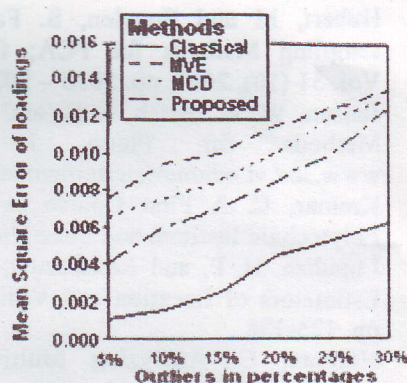


Figure 4. MSE of Unique Variance
when centre = $\sqrt{\chi^2_{p,0.95}}$

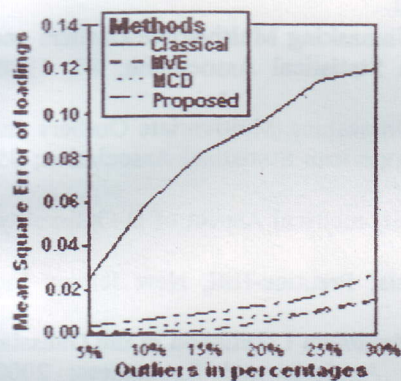


Figure 5. MSE of Unique Variance
when centre = $\chi^2_{p,0.95}$

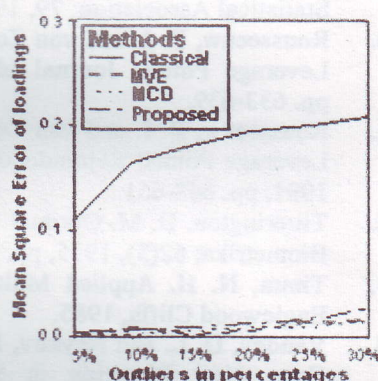


Figure 6. MSE of Unique Variance
when centre = $2\chi^2_{p,0.95}$

REFERENCES

1. Agullo, J. Exact Iterative Computation of the Multivariate Minimum Volume Ellipsoid with Branch and Bound Algorithm; In Proceedings of Computational Statistics; 1996, pp. 175-180.
2. Daves, P. L. Asymptotic Behavior of S-Estimates of Multivariate Location Parameters and Dispersion Matrices; Annals of Statistics; 15, 1987, pp. 1269-1292

3. Hubert, M and Engelen, S. Fast Cross-validation of High-breakdown Re-sampling Methods for PCA; *Computational Statistics and Data Analysis*; Vol. 51 (10), 2007, pp. 5013 – 5024.
4. Jensen, W. A., Birch, J. B. and Woodall, W. H. High Breakdown Estimation Methods for Phase I Multivariate Control Charts, 2002. www.stat.vt.edu/newsletter/annual_report_statistics_2008.pdf
5. Kramar, C. A First Course in Methods of Multivariate Analysis; Virginia Polytechnic Institute and State University; Blacksburg, VI, 1972.
6. Lopuhaa, H. P. and Rousseeuw, P. J. Breakdown Point of Affine Equivariant Estimators of Location and Variance Matrices; *Annals of Statistics*; 27, 1991, pp. 125-138.
7. Norman, C. Analyzing Multivariate Data; San Diego, Harcourt Brace Jovanovich, 1987.
8. Rocke, D. M. and Woodruff, D. L. Robust Estimation of Multivariate Location and Shape; *Journal of Statistical Planning and Inference*; 57, 1998, pp. 245-255.
9. Rousseeuw, P. J. Least Median of Squares Regression; *Journal of the American Statistical Association*; 79, 1984, pp. 871-880.
10. Rousseeuw, P. J. and van Zomeren, B. C. Unmasking Multivariate Outliers and Leverage Points; *Journal of the American Statistical Association*; 85, 1990, pp. 633-639.
11. Rousseeuw, P. J. and van Zomeren, B. C. Unmasking Multivariate Outliers and Leverage Points; *Rejoinder-Journal of the American Statistical Association*; 85, 1991, pp. 648-651
12. Titterton, D. M. Optimal Design; Some Geometrical Aspect of D-Optimality; *Biometrika*; 62(3), 1975, pp. 313 – 320.
13. Timm, N. H. Applied Multivariate Analysis; Prentice-Hill, New Jersey, Inc. Englewood Cliffs, 1985.
14. Vandev, D. L. and Neykov, N. M. Robust Maximum Likelihood in the Gaussian Case. New Direction in Statistical Data Analysis and Robustness; 2000. www.fmi.uni-sofia.bg/fmi/statisti/Personal/Vandev/robust.html.
15. Vargas, J. A. Robust Estimation in Multivariate Control Charts for Individual Observations; *Journal of Quality Technology*; 35(4), 2003, pp. 367-376.