A Comparison of Various Robust Estimators in Mahalanobis Depth

Surabhi S Nair

KNM Govt. Arts and Science College, Kanjiramkulam, Thiruvananthapuram, Kerala, India
Email: *surabhinair93[at]gmail.com*

Abstract: ***One of the fundamental ideas in multidimensional data analysis is data depth. The phrase data depth refers to the depth of a specific point within a large multivariate data cloud. The sample points can be ranked from the center outward rather than the usual smallest to biggest rank. This is one method for identifying a good representation across the entire data. The Mahalanobis Depth (MD) approach, which is one of the most used depth methods, is based on the standard mean vector and covariance matrix. If certain presumptions are valid, conventional MD should function quite well; nevertheless, while a few of these assertions are false, traditional MD may not be reliable. Anomalies are very likely to occur in both the sample mean vector and covariance matrix. Due to this, the classic Mahalanobis depth is unable to produce accurate findings when the data contains abnormalities. As a result, this work proposed a set of robust Mahalanobis depths for location estimation namely Robust MD based on M-estimator (RMD - M), MM estimator (RMD - MM), and Minimum Regularized Covariance Determinant estimator (RRMD- Robust Regularized Mahalanobis Depth). All the proposed depth functions work well and give reliable location when the data is not high-dimensional, the variable number p is less than sample size n. But in the high dimensional data set, where variable number p is greater than sample size n, some of the proposed MD cannot be determined. Even with high-dimensional and corrupted data, one of the proposed depth functions RRMD produces credible findings when compared to existing approaches and other proposed depth functions in this study. In comparison to other depth functions that have been suggested, this study demonstrates that RRMD is successful in locating a central point even in high-dimensional data sets with real data study and simulation study up to a specific level of contamination.***

Keywords: Mahalanobis depth – Outliers – Robust distance – Robust estimators

1. Introduction

The nonparametric approach to multidimensional data analysis revolves around the idea of data depth. The idea of data depth was initially put forward by [17] to visualize two-dimensional data sets, and it has since been expanded by [3] to include multidimensional data sets. Instead of using density or linear ordering to represent data distribution, a statistical technique called data depth uses center-outward ordering [11]. According to [6], the idea of data depth is used for statistical analysis with several variables because it offers a nonparametric way. Numerous researchers have developed various depth preliminary notions that have been documented in the literature.

To solve the problem of sensitivity issues in complex data analysis, researchers are seeking for answers in robust depth processes. In addition, robust estimators like M-estimators [4], MM-estimators [18] Minimum Covariance Determinants (MCD) estimators [5], Minimum Volume Ellipsoids (MVE) estimators [16], and Minimum Regularized Covariance Determinants (MRCD) [1] can be used to replace the classical estimators in robust depth procedures to deal with the presence of outliers.

An assortment of RMDs is suggested in this article for determining the measure of location, particularly in high dimensions. They are based on taking into account various combinations of reliable location and covariance matrix estimators, namely the M estimator, MM estimator, and MRCD estimator. There is a requirement for regularization in the estimators because some of the recommended techniques, with the exception of RRMD, are unable to find Mahalanobis depth in multidimensional data. It is investigated that the suggested RRMD technique for location estimation in data with high dimensions gives reliable results regardless of whether the data is corrupted using different types of simulations and two real datasets.

The basic structure of the paper is as follows. Section 2 provides descriptions of conventional Mahalanobis depth, the employed robust estimator, and the proposed strategy. Section 3 will present the findings and discussions constructed on the real data and simulated study. The final portion will include the conclusion.

2. Methods

Huge volumes of data are produced and tainted by noise in many industries nowadays. A crucial topic is the creation of training methods that are resilient to data inconsistencies and disturbances. In this section, the principles of controlled learning, including the conventional Mahalanobis Depth, robust estimators used, and the suggested depth process, are discussed.

2.1 Mahalanobis Depth

From Mahalanobis distance, [6] initially explained Mahalanobis depth (MD). Generalized distance is a statistical concept introduced by [7] that may be estimated using a traditional mean vector and covariance matrix. The Mahalanobis distance is used to calculate the Mahalanobis depth of an observation. Mahalanobis depth is the contrary of the inverse of Mahalanobis distance. For an observation about a d-dimensional data with the form $y \in S \subset R^d$

The Mahalanobis depth (MD) is given by

$$MD(Y, \bar{Y}, S) = [1 + D(Y, \bar{Y}, S)]^{-1}$$

and $D(Y,\bar{Y},S)=(Y-\bar{Y})'S^{-1}(Y-\bar{Y})$ is the squared Mahalanobis distance (D), where $\bar{Y}$ and S are the mean vector and sample covariance matrix. Since it is reliant on non-robust parameters like the mean and dispersion matrix, this algorithm lacks to be reliable.

2.2 M-Estimator

M estimator was proposed by [4], which expands maximum likelihood estimation to minimization of the function

$$\sum_{i=1}^{n}\rho\,(y_i,\theta)=min,\theta\in\Theta \quad (1)$$

where ρ is a function with specific properties and the solution

$$\hat{\theta}=\arg min_{\theta}\sum_{i=1}^{n}\rho\,(y_i,\theta),\theta\in\Theta$$

are called M-estimators (M for maximum likelihood type). Therefore, maximum likelihood estimators are a particular kind of M estimator. Different robust estimators are produced by choosing various functions, ρ, and each robust estimator's robustness is measured by how unaffected it is by outliers or deviations from the presumptive statistical model.

For location and scale factors, the M estimator can be created in both univariate and multivariate contexts. The maximal likelihood estimator of a parametric model, where $\rho=\{y,\theta\}$, is the density function of P_θ, is included in the class of M-estimators. MLE is the solution of minimization

$$\sum_{i=1}^{n}(-\log f(Y_i,\theta))=min,\theta\in\Theta$$

The M-estimator, say T, is a root or roots of the equation if ρ in (1) is differentiable with continuous derivative $\varphi(.,\theta)$.

$$\sum_{i=1}^{n}\varphi(Y_i,\theta)=min,\theta\in\Theta$$

From now

$$\frac{1}{n}\sum_{i=1}^{n}\varphi(Y_i,T)=E[((Y_i,T)]=0,T\in\Theta \quad (2)$$

From (1) and (2) as a resolution of the minimization problem, the M function which corresponds to T is defined

$$\int\rho(y,I(P)dP(y)=E[(Y,I(P)]=min,T(P)\in\Theta \quad (3)$$

The solution of the equation becomes

$$\int\varphi(y,I(P)dP(y)=E[(Y,I(P)]=min,I(P)\in\Theta$$

where the function I(P) is Fisher constant.

The model with shift parameter θ is a significant specific example, in which $y_1,y_2,\dots,y_n$ are independent observations with a similar distribution function $F(Y-\theta)$, $\theta\in\Theta$, and the distribution function F is often unknown. M-estimator of location parameter T is defined as a result of minimization

$$\sum_{i=1}^{n}\rho\,(y_i-\theta)=min$$

And if $\rho(.)$ is differentiable with absolutely derivative $\varphi(.)$, then T solves the equation

$$\sum_{i=1}^{n}\varphi\,(y_i-\theta)=min$$

Then the solution of the equation is

$\int\varphi((y_i-\theta)dP(y)=min$ have unique solution $\theta=0$

2.3 MM-Estimator

The class of MM estimator is an estimator with a high breakdown value, was invented by [18]. It is the extension of the S estimator and was proposed [13]. The determinant of the matrix S is defined as the S estimator of location and covariance such that it is minimized under the constraints.

$$\frac{1}{n}\sum_{i=1}^{n}\phi\sqrt{(X_i-\mu)'S^{-1}(X_i-\mu)}=c \quad (4)$$

Where c is a constant and $\phi(X)$ is the loss function.

According to [15], Tukey's bi-weight function is a frequent choice of loss.

$$\phi(X)=\begin{cases}\frac{k^2}{6}\left[\left[1-(1-\frac{x}{k})^2\right]\right]^3, |x|\le k\\ \frac{k^2}{6}, |x|>k\end{cases} \quad (5)$$

As suggested by [8], the following stages should be taken into account while estimating the MM estimator.

Define a loss function ϕ to compute the S estimator of location $\hat{\mu}$ and covariance $\hat{\Sigma}$.

Compute $\hat{\sigma}=|\Sigma|^{1/2\phi}$

Determine the MM estimator of the location parameter $\hat{\mu}$ and shape parameter $\hat{\Gamma}$, which minimizes

$\frac{1}{n}\sum_{i=1}^{n}\phi_1((X_i-\mu)'\Gamma^{-1}(X_i-\mu))^{1/2}/\hat{\sigma}$

Compute the MM estimator of the covariance matrix

$\hat{\Sigma}=\hat{\sigma}\hat{\Gamma}$

2.3 Minimum Covariance Determinant MCD Estimator

The Minimum Covariance Determinant (MCD) Estimator was established by [5], which can calculate the mean vector and covariance matrix as well as identify outliers. The smallest determinant-containing sample of h observations is searched with respect to the covariance matrix. The scatter estimate is a multiple of the scatter matrix, and the location estimate is the mean value of that group according to this estimator.

$M_X(H)=h^{-1}X_H'I_h$ (6)

$$S_X(H)=(h-1)^{-1}(X_H-M_X(H))'(X_H-M_X(H)) \quad (7)$$

After that, the MCD method seeks to minimize the determinant of $S_X(H)$ for all $H \in \mathcal{H}_h$.

$$H_{MCD} = \underset{H \in \mathcal{H}_h}{argmin}\left(det(S_X(H))^{1/p}\right) \quad (8)$$

The MCD estimation of the location M_{MCD} is defined by the mean of the h-subset, whereas the estimated MCD scatter estimate S_{MCD} is specified as a multiplier of the samples scatter matrix and is obtained by

$$M_{MCD} = M_X(H_{MCD}) \quad (9)$$

$$S_{MCD} = C_\alpha S_X(H_{MCD}) \quad (10)$$

where C_α is a consistency factor similar to that offered by [2] that is dependent upon the trimming percentage = $(n-h)/n$. Its major disadvantage is that when the dimension is larger than the size of the subset, it produces incorrect results. MCD must be modified for high dimensions because the current MCD techniques are inefficient and inadequate.

2.4 Minimum Regularized Covariance Determinant (MRCD) Estimator

The MRCD estimator is a modified version of MCD estimators for high-dimensional data and was proposed by [1]. To guarantee that the MRCD scatter estimator is scale equivariant and location unvarying, as is common in the literature, first, standardize the variables. The use of a trustworthy univariate location and scale estimate is required for standardization. For this, the median of each subset is calculated and placed in a location vector called m_x. Additionally, each variable's scale using the Qn estimator is calculated [13], then insert these scales into d_x, the diagonal matrix.

$$Z_i = d_x^{-1}(x_i - m_x) \quad (11)$$

The regularized scatter matrix of the standardized observation is

$$\mathcal{S}(H) = \rho T + (1-\rho) C_\alpha S_Z(H)$$

where $S_Z(H)$ is defined in (7), however, in the case of Z, C_α is the same consistency parameter as in (10).

Assume that the orthogonal matrix Q provides the pertinent eigenvectors and that A is the diagonal matrix holding the eigenvalues of T. It will be useful to use the spectral decomposition $T = QAQ'$.

Then,

$$\mathcal{S}(H) = QA^{1/2}[\rho I + (1-\rho) C_\alpha S_W(H)] AA^{1/2}Q' \quad (12)$$

where W is the $n \times p$ matrix consisting the transformed standardized observations

$w_i = A^{-1/2}Q'Z_i$, and $S_W(H) = A^{-1/2}Q'S_Z QA^{-1/2}$

The MRCD subset is given by

$$H_{MRCD} = \underset{H \in \mathcal{H}_h}{argmin}\left(det(\rho I + (1-\rho) C_\alpha S_W(H))^{1/p}\right) \quad (13)$$

The MRCD location, M_{MRCD} and scatter, S_{MRCD} estimations of the initial data matrix X are defined as follows

$M_{MRCD} = m_X + d_x M_Z(H_{MRCD})$

$S_{MRCD} = C_\alpha S_X(H_{MCD})$

2.5 Mahalanobis Depth based on Various Robust Estimators

The classical Mahalanobis Depth (MD) is defined in the equation (1). Since it is reliant on non-robust parameters like the mean and dispersion matrix, this algorithm lacks to be reliable. To get a reliable result, MD is calculated using robust location vector and covariance matrix using various robust estimators. MD using MCD estimator was proposed by [10]. This paper proposed a collection of Mahalanobis Depth for location estimation using various robust estimators namely, MD using M-estimator (MD-M), MD using MM-estimator (MD-MM), and MD using MRCD-estimator (RRMD-Robust Regularized Mahalanobis Depth).

The computational Robust depth procedure for MD is as follows.

Robust Mahalanobis Depth is obtained by replacing the classical location and scatter matrix by robust location and scatter matrix.

Let μ_{Robust}, and Σ_{Robust} be the robust location and scatter matrix. The Robust Mahalanobis Depth (RMD) obtained from the squared Mahalanobis distance.

$D(Y, \mu_{Robust}, \Sigma_{Robust}) = (Y - \mu)'{\Sigma_{Robust}}^{-1}(Y - \mu_{Robust})$
and is given by
$RMD = [1 + D(Y, \mu_{Robust}, \Sigma_{Robust})]^{-1}$

Let $Y = (Y_1, Y_2, \dots, Y_d)$ be a d dimensional multivariate data set and y be a numerical vector whose depth is to be calculated.

1) By using the dataset calculate robust location (μ_{Robust}) and scatter estimators (Σ_{Robust}).
2) The Squared Mahalanobis distance, can be calculated from (i)
   a) e., $D(Y, \mu_{Robust}, \Sigma_{Robust}) = (Y - \mu)'{\Sigma_{Robust}}^{-1}(Y - \mu_{Robust})$
      - S_D be the sorted distance given in (ii)
      - M_{S_D} be the median from the distance from (iii),

   b) $M_{S_D} = Median(S_D)$
3) D_y be the difference between Squared Mahalanobis distance value from (ii) and median from (iv), ie., $D_y = D(Y, \mu_{Robust}, \Sigma_{Robust}) - M_{S_D}$
4) $Abs(D_y)$ be the absolute value of the difference in (v)
5) Now, the proposed depth procedure, Robust Mahalanobis Depth can be calculated by $RMD = [1 + Abs(D_y)]^{-1}$
6) Putting (μ_M, Σ_M), (μ_{MM}, Σ_{MM}), $(\mu_{MCD}, \Sigma_{MCD})$, $(\mu_{MRCD}, \Sigma_{MRCD})$, in (ii) get MD-M, MD-MM, MD-MCD, and RRMD respectively.
7) where (μ_M, Σ_M), (μ_{MM}, Σ_{MM}), $(\mu_{MCD}, \Sigma_{MCD})$, $(\mu_{MRCD}, \Sigma_{MRCD})$, be the location and scatter matrix using

M, MM, MCD and MRCD respectively.

3. Experimental Result

In this section, it was compared to the usual Mahalanobis depth procedure in order to analyze the performance of the suggested robust depth procedures. So as to compute the location measurement corresponding to the deepest point in multivariate and high-dimensional data, experiments were conducted in actual and simulated environments.

3.1 Real Data Study

In this section, Delivery Time Data from Montgomery and Peck (1982), which is multivariate real data, is taken into consideration to study the effectiveness of the mentioned depth methods. The data set contains 25 observations and 3 variables namely the number of products, the distance walked by the route driver, and Delivery time. Found that there are 6 outliers using distance-distance plot. Also, classical and proposed Mahalanobis depth function, which can be used to find the location parameter in high-dimensional datasets. Here the NCI60 data which is obtained from the "ISLR" package of R software. The data contains expression levels on 6830 genes from 64 cancer cell lines. Cancer type is also recorded. Due to the enormous dimensions of these datasets, this study only used the first p $(p = 3 * n)$ variables for convenience. The results obtained from real data study is concluded in the form of Table 1 and Table 2.

Table 1: Deepest point and observation under various Mahalanobis Depth for n > p

Methods	MD	RMD - M	RMD - MM	RMD - MCD	RRMD
Deepest Point	0.932 (0.836)	0.855 (0.855)	0.989 (0.825)	0.853 (0.790)	0.888 (0.756)
Observation Number	15 (6)	17 (17)	6 (17)	17 (17)	17 (17)

(.) without outliers

Table 2: Deepest point and observation under various Mahalanobis Depth for n < p

Methods	MD	RRMD
Deepest Point	0.097 (0.086)	0.031 (0.029)
Observation Number	23 (5)	49 (49)

(.) without outliers

Classical MD and the suggested algorithm, Robust Mahalanobis Depth (RMD) is obtained based on various robust estimators and then the location parameter is calculated using the depth values for the two real data sets. In the first case except the classical method, the suggested methods perform well and give same location measurement under with and without outlier condition. But in high-dimensional data, the classical methods produce various results in both conditions. The suggested methods MD-M, MD-MM, and MM-MCD did not produce any depth values, therefore the location measurement can't be calculated. Only the suggested depth procedure RRMD gives the reliable result in both studies.

3.2 Simulation Study

Two simulation models with various contamination levels are carried out to evaluate the performance of the suggested algorithms compared with the existing method. The experiments were carried out by computing the location measure that correspond to maximum depth values.

First generated data with dimension 100×6, with mean vector $\mu = (0,0,\dots,0)_{1\times 6}$and covariance matrix$\Sigma = I_6$. Here n=100, and p=6. Further same experiments were performed under various levels of contaminations, such as ε= 5%, 10%, 20%, 30% (For Location$\mu = (1.5,1.5,\dots,1.5)$ and $\Sigma = I_6$, Scale $\mu = (0,0,\dots,0)_{1\times 300}$, and $\Sigma = 1.5I_6$, Location, and Scale, $\mu = (1,1,\dots,1)$ and $\Sigma = 2I_6$) are taken into account.

For the second simulation study, generate data with dimension 100×300, with mean vector $\mu = (0,0,\dots,0)_{1\times 300}$and covariance matrix$\Sigma = I_{300}$. Here n=100, and p=300. Further same experiments were performed under various levels of contaminations, such as ε= 5%, 10%, 20%, 30% (For Location$\mu = (1.5,1.5,\dots,1.5)$ and $\Sigma = I_{300}$, Scale $\mu = (0,0,\dots,0)_{1\times 300}$, and $\Sigma = 1.5I_{300}$, Location, and Scale, $\mu = (1.5,1.5,\dots,1.5)$ and $\Sigma = 1.5I_{300}$) are taken into account.

From the simulation study it is concluded that the suggested robust depth procedures performs well when the number of observations is higher than the variables. But in high dimensional case only the RRMD can tolerate certain levels of contamination and gives the same deepest point up to a certain level of contamination. The results obtained from simulation studies are given in Table 3 and Table 4.

Table 3: Deepest point and observation under various contamination models with n>p

Location Contamination					
e	**MD**	**RMD - M**	**RMD - MM**	**RMD - MCD**	**RRMD**
0.05	0.603 (83)	0.634 (83)	0.616 (83)	0.566 (83)	0.702 (83)
0.10	0.550 (83)	0.567 (83)	0.575 (83)	0.584 (83)	0.665 (83)
0.20	0.505 (83)	0.557 (80)	0.534 (85)	0.567 (83)	0.652 (83)
0.30	0.473 (50)	0.610 (50)	0.607 (50)	0.609 (50)	0.633 (50)
Scale Contamination					
e	**MD**	**RMD - M**	**RMD - MM**	**RMD - MCD**	**RRMD**
0.05	0.613 (83)	0.623 (83)	0.642 (83)	0.664 (83)	0.721 (83)
0.10	0.573 (83)	0.583 (83)	0.633 (83)	0.637 (83)	0.711 (83)
0.20	0.518 (83)	0.581 (83)	0.602 (83)	0.524 (83)	0.686 (83)
0.30	0.500 (50)	0.666 (50)	0.627 (50)	0.632 (50)	0.650 (50)

Location - Scale Contamination					
e	MD	RMD - M	RMD - MM	RMD - MCD	RRMD
0.05	0.622 (83)	0.614 (83)	0.626 (83)	0.649 (83)	0.718 (83)
0.10	0.579 (83)	0.607 (83)	0.611 (83)	0.589 (83)	0.725 (83)
0.20	0.529 (50)	0.569 (83)	0.565 (83)	0.559 (83)	0.716 (83)
0.30	0.520 (50)	0.692 (50)	0.670 (50)	0.669 (50)	0.606 (50)

Table 4: Deepest point and observation under various contamination models with $n < p$

	Location Contamination		Scale Contamination		Location-Scale Contamination	
e	MD	RRMD	MD	RRMD	MD	RRMD
0.05	0.178(34)	0.019 (33)	0.987(20)	0.018 (33)	0.057(5)	0.019 (33)
0.10	0.099 (96)	0.019 (33)	0.870 (41)	0.018 (33)	0.044 (93)	0.019 (33)
0.20	0.126 (10)	0.019 (33)	0.108 (77)	0.018 (33)	0.032 (88)	0.019 (33)
0.30	0.014 (38)	0.019 (28)	0.082 (30)	0.018 (22)	0.046 (90)	0.019 (22)

4. Conclusion

Conventional methods should work reasonably well if certain assumptions are true, however, they may not be trustworthy if one or more of these assumptions are erroneous. Both the sample mean vector and covariance matrix are extremely susceptible to anomalies. As a result, when the data contains anomalies, the traditional Mahalanobis depth fails to generate reliable results. For non-normal conditions, a robust alternative is required to improve accuracy even when the data somewhat depart from the model assumptions. This study suggested various notions of Mahalanobis depth based on robust estimators namely MD-M, MD-MM, MD-MCD, and RRMD to find the location parameters. The proposed methods are compared with the existing procedure and give reliable results up to certain levels of contamination when the variable number is less than the number of observations. But in high dimensional data (the variable number is greater than the number of observations) RRMD only gives the reliable results. The study demonstrated that, even with high-dimensional data, the suggested depth approach, RRMD, outperforms the existing and other suggested method for robust and affine equivariant location. The research groups can locate the best location with more accuracy when employing these methods by locating the deepest point in a set of data rather than depending on a more traditional method to figure out location.

References

[1] Boudt K, Rousseeuw P J and Vanduffel S, The minimum Regularized covariance determinant estimator, Statistics and Computing, Vol.30, pp.113-128, 2020

[2] Croux C and Dehon C, Robust linear discriminant analysis using S-estimators, *Canadian Journal of Statistics*, Vol.29, pp.473-493, 2001.

[3] Donoho D L and Gasko M, Breakdown properties of location estimates based on halfspace depth and projected outlyingness, *Annals of Statistics*, Vol.2, pp.1803–1827, 1992.

[4] Huber P J, Robust Estimation of a Location Parameter, *Annals of Mathematical Statistics*, Vol.35, pp.73-101, 1964.

[5] Hubert M, and Debruyne M, Minimum covariance determinant, *WIREs Computational Statistics*, Vol.2, pp.36-43, 2009.

[6] Liu R Y, Parelius J M and Singh K, "Multivariate analysis by data depth: Descriptive Statistics, Graphics and Inference", *The Annals of Statistics*, Vol.27, pp.783-858, 1993.

[7] Mahalanobis P, "On the generalized distance in statistics", *Proceedings of the National Academy India*, Vol.12, pp.49–55, 1936.

[8] Maronna R A, Douglas Martin R, Yohai V J, Robust Statistics: Theory and Methods, Wiley Series in Probability and Statistics, 2006.

[9] Modarres R, Data Depth. In: Lovric, M. (eds) International Encyclopaedia of Statistical Science. Springer, Berlin, Heidelberg, pp.334-336, 2014.

[10] Muthukrishnan R and Poonkuzhali G, Computing Robust Measure of Multivariate Location – Data Depth Approach, Global Journal of Pure and Applied Mathematics, Vol.13, pp. 691-700, 2017.

[11] Rousseeuw P J, Least median of squares regression. *Journal of the American Statistical Association*, Vol.79, pp.871–880, 1984.

[12] Rousseeuw P J and Leroy A M, Robust Regression and Outlier Detection, Wiley Interscience, New York (Series in Applied Probability and Statistics), 1987.

[13] Rousseeuw P J and Croux C, Alternatives to the median absolute deviation, *Journal of the American Statistical Association*, Vol.88, pp.1273–1283, 1993.

[14] Rousseeuw P J and Hubert M, High-Breakdown Estimators of Multivariate Location and Scatter. In: Becker C, Fried R, Kuhnt S (eds) Robustness and Complex Data Structures. Springer, Berlin, Heidelberg, 2013.

[15] SenGupta A, Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions, *Journal of Multivariate Analysis*, Vol.23, pp.209–219, 1987.

[16] Stefan V A and Rousseeuw P J, Minimum volume ellipsoid, *WIREs Computational Statistics*, Vol.1, pp.71-82, 2009.

[17] Tukey J W, "Mathematics and the picturing of data". *In: Proceeding of the International Congress of Mathematicians*, Vancouver, pp.523–531, 1975.

[18] Yohai V J, High Breakdown-Point and High Efficiency Robust Estimates for Regression, *The Annals of Statistics,* Vol.15, pp. 642-656, 1987.