Journal of Computer and Mathematical Sciences, *Vol.9(12),1907-1914 December 2018* *ISSN 0976-5727 (Print)*
(An International Research Journal), *www.compmath-journal.org* *ISSN 2319-8133 (Online)*

Robust Quadratic Discriminant Analysis using Kurtosis Method

Lakshmi. R and T. A. Sajesh

Department of Statistics,
St. Thomas College Autonomous, Thrissur, INDIA.
email: lakshmi.nss19@yahoo.com, sajesh.t.abraham@gmail.com

(Received on: November 29, 2018)

ABSTRACT

The objective of the Discriminant Analysis is to classify the objects into mutually exclusive classes on the basis of some apriori information. But the presence of outliers in the data set lead to wrong assumptions. It is always necessary to have some consistency and objectivity in the treatment of outliers. In this paper we propose a robust quadratic discriminant rule based on kurtosis method and a comparative study is conducted.

Keywords: Discriminant Analysis, Robust Estiamtors, Kurtosis Method.

1. INTRODUCTION

Let there be k populations $\Pi_1, \Pi_2, \ldots\ldots, \Pi_k$. Assume that their exact distributions are known. Let $x_{10}, x_{20}, \ldots\ldots, x_{p0}$ be a set of values of p variables of an object of any of the k populations. The objective of the discriminant analysis is to identify the population of this object on the basis of the observed values of p variables. The identification of the population is done by three different methods. They are: (i) Maximum Likelihood Discriminant Rule. (ii) Bayes Discriminant Rule (iii) Fishers Linear Discriminant Function.

These conventional classification rules are affected to a great extent if multivariate normality of data and homogeneity of group covariance matrices are not valid. However, it is suggested Quadratic Classification Rule if multivariate normality of data is justified but not the homogeneity of group covariance matrices. The heterogeneity of group covariance matrices may arise if multivariate normality is not fixed.

The Quadratic Discriminant function is given by

$$Q(x) = \frac{1}{2}\ln\frac{|\Sigma_2|}{|\Sigma_1|} - \frac{1}{2}(x-\mu_1)'\,\Sigma^{-1}{}_1\,(x-\mu_1) + \frac{1}{2}\,(x-\mu_2)'\,\Sigma^{-1}{}_2(\,x-\mu_2) \qquad (1)$$

For sample data this quadratic composite is

$$q(x) = \frac{1}{2}\ln\frac{|S_2|}{|fS_1|} - \frac{1}{2}\,(x-\overline{X}_1)'\,S_1{}^{-1}(x-\overline{X}_1) + \frac{1}{2}(x-\overline{X}_2)'\,S_2{}^{-1}(x-\overline{X}_2) \qquad (2)$$

The object is allocated to $\Pi_1$if $q(x) > 0$, otherwise to Π_2. In contaminated site studies it is common to find that the data contain some surprisingly high values. It is well known that the classical estimates of population mean and dispersion are very sensitive to outlying observations. Consequently, the CQDR based on the classical methods is inappropriate at contaminated data sets. To overcome this it is always better to use some robust alternative methods based on robust estimators of location and scatter. Hubert and Van Driessen (2004) proposed a robust discriminant method using reweighted Minimum Covariance Determinant (MCD) estimators of location and scatter (Rosseeuw, 1983, 1984). Hubert and Van Driessen (2004) shows that their method allows to discriminate between several populations, with equal or unequal covariance structure, and with equal or unequal membership probabilities. However, Sajesh and Srinivasan (2011) showed that the MCD estimators are failed to detect outliers from moderately large contaminated situations and it is getting worse with increasing dimension. This study proposes a robust discriminant rule based on Kurtosis estimation of location and scatter. Performance of the proposed method is compared with that of CQDR and RQDR based on MCD using simulation.

Pena and Prieto in 2001 present a procedure to detect outliers based on the analysis of the projections of the sample points onto a certain set of $2p$ directions, where p is the dimension of the sample space. These directions are obtained by maximizing and minimizing the kurtosis coefficient of the projections Consider univariate case in which different types of outliers produce different effects on the kurtosis coefficient. Outliers generated by usual symmetric contaminated model increase the kurtosis coefficient of the observed data. A small proportion of outliers generated by an asymmetric contaminated model also increase the kurtosis coefficient of the observed data. These two results suggest that for multivariate data outliers may be revealed on univariate projections onto directions obtained by maximizing the kurtosis coefficient of the projected data. However, a large proportion of outliers generated by an asymmetric contamination model can make the kurtosis coefficient of the data very small, close to its minimum possible value. This result suggests searching for outliers also using directions obtained by minimizing the kurtosis of the projection. Therefore, a procedure that would search for outliers by projecting the data onto the directions that maximize or minimize the kurtosis of the projected point is used in kurtosis method.

The computation of directions maximizing the kurtosis coefficient is affine equivariant. Also the kurtosis method possesses high breakdown value and bounded influence function. Next section of this paper propose Robust Quadratic Discriminant Rule based on the above mentioned KURTOSIS estimators and its efficiency is checked by a simulation study in which comparison of the proposed rule against classical rule and robust MCD method is done.

2. ROBUST QUADRATIC DISCRIMINANT RULE

Discriminant analysis tries to obtain rules that describe the separation between the observations. These rules then allow classifying new observations into one of the population.

Here we will focus on the Bayesian discriminant rule which is a generalization of the maximum likelihood rule. We assume that we can describe experiment in each population π_j by a p −dimensional random variable X_j with density f_j.Denote p_j as the prior probability of the populationπ_j that is the probability of an observation to come fromπ_j. Then the Bayesian discriminant rule assigns an observation $x \in \mathbb{R}^p$ to that population π_k for which $\ln(p_j f_j(x))$ is maximal over all $j = 1,2, \dots, l$. If f_j is the density of the multivariate normal distribution$N_p(\mu_j, \Sigma_j)$, it is easilybe derived that this discriminant rule corresponds with maximizing the quadratic discriminant scores $d_j^Q(x)$, defined as

$$d_j^Q(x) = -\frac{1}{2}\ln|\Sigma_j| - \frac{1}{2}(x-\mu_j)^t \Sigma_j^{-1}(x-\mu_j) + \ln(p_j) \quad (3)$$

When all the covariance matrices are assumed to be equal. The quadratic scores (3) can be simplified to $d_j^L(x) = \mu_j^t \Sigma^{-1} x - \frac{1}{2}\mu_j^t \Sigma^{-1}\mu_j + \ln(p_j)$

As μ_j, Σ_j, and p_j are in practice unknown, they have to be estimated from the sample data. To estimate μ_j, Σ_j, one usually uses the group mean $\bar{x}_j$ and the group empirical covariance matrix S_j, yielding the Classical Quadratic Discriminant Rule(CQDR):

Allocate x to π_k if $\hat{d}_k^{CQ}(x) > \hat{d}_j^{CQ}(x)$ for all $j = 1,2, \dots, l, j \neq k$ with

$$\hat{d}_j^{CQ} = -\frac{1}{2}\ln|S_j| - \frac{1}{2}(x-\bar{x}_j)^t S_j^{-1}(x-\bar{x}_j) + \ln(\hat{p}_j^C). \quad (4)$$

For the estimates of the membership probabilities $\hat{p}_j^C$ in (4) two popular choices are mentioned here. First choice is that the $\hat{p}_j^C$ are considered being constant overall populations, yielding $\hat{p}_j^C = 1/k$ for each j and the second one is they are estimated as the relative frequencies of the observations in each group, thus $\hat{p}_j^C = n_j/n$. It is however known that the classical estimates $\bar{\mathbf{x}}_j$ and $\mathbf{S}_j$ are very sensitive to outlying observations, making the CQDR rule inappropriate at contaminated datasets.

In the first step of robust quadratic discriminant rule, group-wise kurtosis estimates of location and scatter are calculated. Let $\mathbf{X}^{(j)}$ be the j[th] population matrix with rows $\mathbf{x}_i^{(j)\mathrm{T}}$ (i = 1, 2, …, n_j) and columns $\mathbf{X}_t^{(j)}$ (t = 1,2, …, p). Then the robust estimates of mean and dispersion of j[th] group are $\hat{\boldsymbol{\mu}}_{j,kurtosis}$ and $\hat{\boldsymbol{\Sigma}}_{j,kurtosis}$ respectively. These robust estimators of location and scatter now allow to flag the outliers in the data and to obtain more robust estimates of the membership probabilities.Let $\tilde{n}_j$ denote the number of non-outliers in group j, and $\tilde{n} = \sum_{j=1}^{k} \tilde{n}_j$, then the robust estimate of the membership probabilities is $\hat{p}_j^R = \frac{\tilde{n}_j}{\tilde{n}}$.

The Robust Quadratic Discriminant Rule $RQDR_{kurtosis}$ thus becomes:

Allocate $\mathbf{x}$ to π_l if $\hat{d}_l^{RQ}(\mathbf{x}) > \hat{d}_j^{RQ}(\mathbf{x})$ for all j = 1, 2, …, k, $j \neq l$ with

$$\hat{d}_j^{RQ}(\mathbf{x}) = -\frac{1}{2}\ln\left|\hat{\boldsymbol{\Sigma}}_{j,kurtosis}\right| - \frac{1}{2}\left(\mathbf{x}-\hat{\boldsymbol{\mu}}_{j,kurtosis}\right)^T \hat{\boldsymbol{\Sigma}}_{j,kurtosis}^{-1}\left(\mathbf{x}-\hat{\boldsymbol{\mu}}_{j,kurtosis}\right) + \ln(\hat{p}_j^R) \quad (5)$$

The performance of this rule is investigated through a simulation study and the results are compared with RQDR based on MCD and CQDR. Throughout this study $RQDR_{(MCD)}$ and $RQDR_{(Kurtosis)}$ are used to denote robust quadratic discriminant rules based on MCD and kurtosis respectively.

3. SIMULATION STUDY

To evaluate the discriminant rules it is necessary to have an estimate of the associated probability of misclassification. To estimate the classification error consists of, splitting the observations randomly into a training set which is used to compose the discriminant rule, and a validation set is used to estimate the misclassification error. In this study, the training set consists of 60% of the observation and the validation set is formed by the remaining 40% of the observations. Because it can happen that this validation set also contains outlying observations which should not be taken in to account, the misclassification probability of group j is estimated by the proportion of non outliers from the validation set that belong to group j and that are badly misclassified. An overall misclassification estimate (MP) is then given by the weighted mean of the misclassification probabilities of all the groups, with weights equal to the the estimated membership probabilities,

$$MP = \sum_{j=1}^{k} p^{\wedge}{}_j^R MP_j$$

Where MP_j is the misclassification probability of group j. Source code for the discriminant $RQDR_{(kurtosis)}$ and the misclassification probability have been written in MATLAB and the simulation is conducted as that of Hubert and Van Driessen (2004).

Hubert and Van Driessen (2004) considered five contaminated situations and one uncontaminated case with varying amounts of contamination. In this study same situations are considered for varying dimensions. Situation A_p considers the uncontaminated situation with dimension p where training dataset is obtained by drawing 500 observations from each population, which denoted by

A_p. $\pi_1 : 500N_p(\boldsymbol{\mu}_{1p}, \boldsymbol{\Sigma}_{1p})$,
$\pi_2 : 500N_p(\boldsymbol{\mu}_{2p}, \boldsymbol{\Sigma}_{2p})$,
$\pi_3 : 500N_p(\boldsymbol{\mu}_{3p}, \boldsymbol{\Sigma}_{3p})$.

Training data sets which also contain outliers are sampled from another distribution. These cases are given below:

B_p. $\pi_1 : 400N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{1,p}) + 100N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_2 : 400N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{2,p}) + 100N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{4,p})$,
π: $400N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{3,p}) + 100N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{4,p})$
C_p. $\pi_1 : 600N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{1,p}) + 150N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{4,p})$,

$\pi_2 : 400N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{2p}) + 100N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_3 : 200N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{3,p}) + 50N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{4,p})$
D_p. $\pi_1 : 600N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{1,p}) + 150N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_2 : 80N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{2,p}) + 20N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_3 : 400N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{3,p}) + 100N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{4,p})$
E_p. $\pi_1 : 400N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{1,p}) + 100N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_2 : 450N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{2p}) + 50N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_3 : 350N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{3,p}) + 150N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{4,p})$
F_p. $\pi_1 : 160N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{1,p}) + 40N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_2 : 160N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{2,p}) + 40N_p(\boldsymbol{\mu}_{1,p}, \boldsymbol{\Sigma}_{4,p})$,
$\pi_3 : 160N_p(\boldsymbol{\mu}_{3,p}, \boldsymbol{\Sigma}_{3,p}) + 40N_p(\boldsymbol{\mu}_{2,p}, \boldsymbol{\Sigma}_{4,p})$.
where
$\boldsymbol{\mu}_{1,5} = (1,0,0,0,0)$ $\boldsymbol{\mu}_{2,5} = (0,1,0,0,0)$ $\boldsymbol{\mu}_{3,5} = (0,0,1,0,0)$
$\boldsymbol{\mu}_{1,10} = (1,0,0,0,0,0,0,0,0,0)$
$\boldsymbol{\mu}_{2,10} = (0,1,0,0,0,0,0,0,0,0)$
$\boldsymbol{\mu}_{3,10} = (0,0,1,0,0,0,0,0,0,0)$
$\boldsymbol{\mu}_{1,20} = (1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)$
$\boldsymbol{\mu}_{2,20} = (0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)$
$\boldsymbol{\mu}_{3,20} = (0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)$
and
$\boldsymbol{\Sigma}_{1,5} = diag(0.4,0.4,0.4,0.4,0.4)^2$
$\boldsymbol{\Sigma}_{2,5} = diag(0.25,0.75,0.75,0.25,0.75)^2$
$\boldsymbol{\Sigma}_{3,5} = diag(0.9,0.6,0.3,0.9,0.6)^2$
$\boldsymbol{\Sigma}_{1,10} = diag(0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4)^2$
$\boldsymbol{\Sigma}_{2,10} = diag(0.25,0.75,0.75,0.25,0.75,0.25,0.75,0.75,0.25,0.75)^2$
$\boldsymbol{\Sigma}_{3,10} = diag(0.9,0.6,0.3,0.9,0.6,0.9,0.6,0.3,0.9,0.6)^2$
$\boldsymbol{\Sigma}_{1,20} = diag(0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4,0.4)^2$
$\boldsymbol{\Sigma}_{2,20} = diag(0.25,0.75,0.75,0.25,0.75,0.25,0.75,0.75,0.25,0.750.25,0.75,0.75,0.25,0.75,$
$0.25,0.75,0.75,0.25,0.75)^2$
$\boldsymbol{\Sigma}_{3,20} = diag(0.9,0.6,0.3,0.9,0.6,0.9,0.6,0.3,0.9,0.6,0.9,0.6,0.3,0.9,0.6,0.9,0.6,0.3,0.9,0.6)^2$

Note that the situationsB_p, C_p, D_p and E_p generate 25% outliers that change the covariance structure of the populations. In setting B the three groups contains an equal number of observations, whereas settings C_p and D_p have unequal group sizes, the most unbalanced situation being considered inD_p. In simulation E_p, the percentage of outliers in the groups varying between 10% and 30%. Finally in setting F_p radical outliers are considered. Per case 100 Monte Carlo simulations are performed. Available MATLAB coding has been used for $RQDR_{(MCD)}$ and CQDR.

To evaluate the discriminant rule a validation set of 1000 observations are generated from each uncontaminated population. As explained in the previous Section the data points that were not flagged as outliers are considered as validation set. Denote $V_1, V_2, and\ V_3$ as those subsets of the validation set in each population. Note that these subsets changed in every trial, and that their size was close to 1000 because the validation set was sampled from the uncontaminated distributions. The misclassification probability of the robust discriminant rule was then estimated in each group as the proportion of badly classified observations from $V_1, V_2, and\ V_3$ using $RQDR_{(MCD)}, RQDR_{(kurtosis)}$ and CQDR. Both the robust and the classical rules were thus evaluated through the same validation sets. Moreover, for each sample the total MP of CQDR is computed as weighted mean of MP1, MP2 and MP3 with weights equal to the robust membership probabilities, in order to make a fair comparison with the robust RQDR methods.

Table 1: The mean of the misclassification probability estimates for RQDR and CQDR methods based on 400 Monte Carlo samples with dimensions p = 5,10,15

Method		p=5				p=10				p=20			
		MP1	**MP2**	**MP3**	**MP**	**MP1**	**MP2**	**MP3**	**MP**	**MP1**	**MP2**	**MP3**	**MP**
A_p	RQDR	0.064	0.074	0.086	0.074	0.041	0.036	0.043	0.040	0.021	0.010	0.014	0.015
	MCD	0.043	0.167	0.194	0.127	0.052	0.198	0.194	0.140	0.058	0.165	0.301	0.143
	CQDR	0.046	0.173	0.163	0.122	0.042	0.209	0.161	0.132	0.038	0.201	0.243	0.138
B_p	RQDR	0.063	0.074	0.082	0.072	0.048	0.037	0.043	0.042	0.023	0.013	0.012	0.016
	MCD	0.061	0.204	0.145	0.135	0.101	0.243	0.803	0.332	0.348	0.523	0.996	0.522
	CQDR	0.684	0.653	0.779	0.703	0.300	0.690	0.871	0.574	0.456	0.442	1.000	0.541
C_p	RQDR	0.040	0.065	0.144	0.065	0.028	0.035	0.073	0.037	0.015	0.012	0.023	0.015
	MCD	0.027	0.144	0.287	0.100	0.043	0.051	0.995	0.090	0.007	0.768	1.000	0.299
	CQDR	0.035	0.614	1.000	0.352	0.005	0.739	1.000	0.276	0.013	0.721	1.000	0.289
D_p	RQDR	0.038	0.218	0.056	0.057	0.027	0.137	0.036	0.038	0.011	0.159	0.016	0.023
	MCD	0.046	0.434	0.062	0.077	0.044	0.890	0.510	0.215	0.007	1.000	0.627	0.163
	CQDR	0.007	1.000	0.672	0.297	0.001	1.000	0.894	0.288	0.001	1.000	0.808	0.197
E_p	RQDR	0.066	0.071	0.092	0.075	0.045	0.029	0.048	0.039	0.025	0.011	0.016	0.017
	MCD	0.097	0.270	0.255	0.208	0.100	0.242	0.442	0.247	0.528	0.383	0.972	0.547
	CQDR	0.652	0.468	0.792	0.632	0.631	0.383	0.897	0.613	0.658	0.292	0.991	0.564
F_p	RQDR	0.070	0.086	0.080	0.078	0.055	0.036	0.044	0.045	0.041	0.018	0.012	0.023
	MCD	0.052	0.215	0.156	0.138	0.074	0.193	0.166	0.141	0.060	0.222	0.184	0.150
	CQDR	0.027	0.254	0.228	0.164	0.440	0.218	0.193	0.147	0.028	0.242	0.290	0.175

From Table 1, the results of simulated situation A_p shows that the misclassification estimates are very comparable at uncontaminated data. In simulation C_p, if we look at the misclassification probability of group three, for p=5, 10, 20, RQDR attains a rather larger misclassification probability. This is however a very small group compared to the sizes of the other two and consequently results in low robust discriminant scores. A similar situation hits in simulationD_p, the $RQDR_{(kurtosis)}$ misclassification probability of group two is larger, the corresponding CQDR misclassification probability of group three and group two, for p=5, 10, 20, in both the simulations C_p and $D_p$attains one. This implies that $RQDR_{(kurtosis)}$ is the best and CQDR is badly misclassified all the observations. In simulation F_p, from the table we get that the CQDR misclassification probability for group one is extremely small for and much lower than$RQDR_{(kurtosis)}$ misclassification probability for group one.

It is clear that in larger dimension cases also, $RQDR_{(kurtosis)}$ with respect to MP. It can also be seen that the misclassification probabilities of $RQDR_{(kurtosis)}$ decreases with increasing dimension. For example, in the case of C_p the misclassification probability (MP) of $RQDR_{(kurtosis)}$ are decreasing from 0.06577 for p=5 to 0.037087 for p=10 to 0.015039 to 0.015 for p=20. This shows that the $RQDR_{(kurtosis)}$ method is suitable for larger dimensional sets.

4. CONCLUSION

In this study a Robust Quadratic Discriminant Rule based on kurtosis estimators has been proposed. The robust quadratic discriminant scores are obtained by using the kurtosis estimates for each group in to the generalized maximum likelihood discriminant rules. Also the membership probabilities are estimated in a robust way by taking only the non outliers in to account.

Simulation study clearly showed how the robust approach was not affected by outliers, unlike the classical rules and examined the performance of discriminant rules in varying situations. $RQDR_{(kurtosis)}$ is compared with $RQDR_{(MCD)}$ and CQDR with respect to the misclassification probabilities. The results showed that the proposed Robust Quadratic Discriminant Rule is much less affected by the presence of outliers compared to other methods and hence can be applied in real life data sets.

ACKNOWLEDGEMENT

First of all, I would like to extend my sincere gratitude to the almighty. I would like to express my wholehearted thanks to my guide Dr.Sajesh, St. Thomas College, for his valuable, sincere and inspiring guidance throughout the course of my work and in the preparation of this project report.

I thank all my teachers for their support and encouragement. And would like to thank my friends for their suggestions, last but not least I express my sincere thanks to my parents for their support and prayers.

REFERENCES

1. Anderson, T.W. An Introduction to Multivariate Statistical Analysis, 2nd edition, John Wiley and Sons Inc., New York (1984).
2. Daniel Pena, Prieto, P., "Multivariate Outlier Detection and Robust Covariance Matrix Estimation," *Journal of the American Statistical Association and the American Society for Quality Technometrics,*VOL.43, No.3 (2001).
3. Hubert, M and Van Driesen, K. "Fast and Robust Discriminant Analysis," *Computational Statistics and Data Analysis*, 45, 301 – 320 (2004).
4. Maronna, R., Zamar, R., "Robust estimators of location and dispersion for high dimensional data sets," *Technometrics,* 44 (4), 307-317 (2002).
5. Rousseeuw, P. J., "Least median of squares regression," Journal *of the American Statistical Association*, 79 (388), 871-880 (1984).
6. Rousseeuw, P.J., "Multivariate Estimation with High Breakdown Point," *Fourth Pannonian Symposium on Mathematical Statistics and Probability.* (1983).
7. Rousseuw, P., Multivariate estimation with high breakdown point. In: Grossman, W., Pflug, G., Vincze, I., Wertz, W. (Eds.), *Mathematical Statistics and Applications*. Vol. B. Riedel Publishing Company, pp. 283-297 (1985).
8. Rousseuw, P., Van Driessen, K., "A fast algorithm for the minimum covariance determinant estimator," *Technometrics,* 41, 212-223 (1999).
9. Sajesh, T.A. and Srinivasan, M.R. "Outlier Detection for High Dimensional Data using Comedian Approach*," iFirst Journal of statistical computation and simulation*. (2011).