International Journal of Research in Advent Technology, Vol.6, No.5, May 2018
E-ISSN: 2321-9637
Available online at www.ijrat.org

Projection Based Data Depth Procedure with Application in Discriminant Analysis

Muthukrishnan R[1], Vadivel M[2], Ramkumar N[3]
Department of Statistics, Bharathiar University, Coimbatore, Tamil Nadu, India[1,2,3]
Email: muthukrishnan1970@gmail.com[1] ,chinnastatttp@gmail.com[2] ,ram222u@gmail.com[3]

Abstract- Projection depth and its associated estimators, namely, Stahel-Donoho (S-D) estimator, Projection Trimmed Mean (PTM), Projection depth Contours (PC) and Projection Median (PM) have been studied in bivariate data. An attempt has been made to compute projection depth and its associated estimators by using the pair of location and scale estimator (Mean, Standard Deviation (SD)), (Median, Median Absolute Deviation (MAD)), and (Median, Q_n).The efficiency of these estimators is carried out by computing average misclassification error in discriminant analysis by using the projection depth based Stahel-Donohoestimator under real and simulating environment. The study concluded that (Median, MAD) and (Median, Q_n) based projection depth estimators performs well when compared with (Mean, SD).

Index Terms- Projection depth and its associated estimators; Robust discrimination analysis.

1. INTRODUCTION

Data depth is a concept which plays an important role in many notable fields of statistics,namely; data exploration, ordering, asymptotic distributions and robust estimation(Liu et al. 1999). The essence of the depth function in multivariate analyses is to measure degree of centrality of a point relative to a data set or to a probability distribution. Many robust procedures have been developed to compute the data depth. The data depth based approach has been received much attention now-a-days. Numerous depth notations have been proposed during the last few decades, namely,half space depth (Tukey 1975), simplicial depth (Liu 1990), regression depth (Rousseeuw and Hubert 1999) and projection depth (Liu 1992; Zuo and Serfling 2000; Zuo 2003).

The Projection Depth(PD) is very favorable to the robust statistics when compared with the other depth notations.It is due to the reason that all the desirable properties of the general statistical depth function defined in Zuo and Serfling (2000), namely, affine invariance, maximality at center, monotonicity relative to deepest point, and vanishing at infinity are satisfied by the PD.

The main objective of this paper is to estimate the associated estimators such as Stahel-Donohoestimator, projection trimmed mean, projection depth contours and projection median for bivariate data based on various pair of projection depth procedures.Further, the performance of the pairs has been studied under various levels of contaminations with the help of Stahel-Donohoestimator, by computing average misclassification probabilities in the context of robust linear discriminant analysis in Hubert and Van Driessen (2014).

The rest of the paper is organized as follows. Section 2 describes the methodology of projection depth and its associated estimators. Section 3 discussesrobust linear discriminant analysis. Section 4 examines the performance and critically compares the three pairs of projection depth procedures. Section 5 presents results obtained in real and simulation studyin the context of robust discriminate analysis.The paper ends with conclusion in the last section.

2. PROJECTION DEPTH AND ITS ASSOCIATED ESTIMATORS

Zuo (2003) introduced a Projection-based depth functions, which has the highest breakdown point among all the existing affine equivariant multivariate location estimators and associated medians. It can induce a lot of favorable estimators, such as Stahel-Donohoestimator and depth weighted means for multivariate data (Zuo et al. 2004; Zuo 2006).Further, Zuo (2006) studied multidimensional trimming based on projection depth. Exact computation of bivariate projection depth and Stahel-Donoho estimator, with a proper choice of (μ,σ) are formulated and studied by Zuo and Lai (2011). Liu and Zuo (2014) studied computational aspects of projection depth and its associated estimators.The brief description of theory of projection depth is as follows.

Let μ (.) and σ (.)be univariate location and scale measures, respectively. Then the outlyingness of a point $x \in R^{P}$ with respect to the distribution function F of X defined by (Liu 1992, Zuo 2003).

$$PD(x,F) = \frac{1}{1+O(x,F)},$$

where,

$$O(x,F) = \sup_{\|u\|=1} |Q(u,x,F)|, \tag{1}$$

where, $Q(u,x,F) = \left(u^{T}x - \mu(F_u)\right)/\sigma(F_u)$ and F_u is the distribution of $u^{T}x$. If $u^{T}x - \mu(F_u) = \sigma(F_u) = 0$, then define $Q(u,x,F) = 0$, which denotes the projection of x onto the unit vector u. Note that the most popular outlying function has the robust choice of μ and σ be the Median and MAD. Here, the pair (*med*, Q_n), where med and Q_n is considered as location and scale estimator of (μ(F), σ(F)) for a given sample $X^n = \{X_1, X_2, \ldots, X_n\}$ from X. Let F_n be the corresponding distribution, then the projection depth and its associated estimators depend on the robust choice of (*Med*, Q_n), $Q(x,u,X^n)$ in (1) with respect to u.

$$Q(x, X^{n}) = \sup_{\|u\|=1} Q(u,x,X^{n}), \tag{2}$$

The outlying function is defined as

$$Q(u,x,X^{n}) = \left|\frac{u^{T}x - Med(u^{T}X^{n})}{Q_n(u^{T}X^{n})}\right|, \tag{3}$$

Where u^T denotes the projection of x onto the unit vector u and $u^{T}X^{n} = \{u^{T}X_1, u^{T}X_2, \ldots, u^{T}X_n\}$. Let $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ denote the order statistics corresponding to the univariate random variables Z_n.

$$Med(X^{n}) = \frac{X(\lfloor (n+1)/2 \rfloor) + X(\lfloor (n+2)/2 \rfloor)}{2},$$

$$Q_n(X^{n}) = d\left\{|X_i - X_j|; i<j\right\}_{(k)},$$

where d is a constant factor and $k = \binom{h}{2} \approx \binom{n}{2}/4$, $h = \lfloor n/2 \rfloor + 1$ is roughly half the number of observations. That is, $\binom{n}{2}$ is the interpoint distances of k[th] order statistics.

The main function of the projection depth is to be responsible for a center-outward ordering for the bivariate data. Based on this ordering, one can make the projection depth contours, which can provide us with a bivariate data of the quantile of an underlying distribution (Halin et al. 2010).

It's defined as

$$PR(\alpha,F) = \{x \in R^{P} : PD(x,F) \ge \alpha\}, \tag{4}$$

where $0 \le \alpha \le \alpha^{*} = \sup_{x \in R^{P}} PD(x,F)$ with α^{th} Projection depth Region (*DR*). Then the corresponding α^{th} Projection Depth Contour can be distinct as the boundary of PDR (α,*F*) under some conditions (Zuo 2003) is given by

$$PC(\alpha,F) = \{x \in R^{P} : PD(x,F) = \alpha\}. \tag{5}$$

The innermost depth contour, which is a singleton in many situations, is the Projection depth Median (*PM*) of Zuo (2003)

$$PM(F) = PC(\alpha^{*}, F).$$

Based on the projection depth region PR (α, F), one can define the α^{th} Projection depth Trimmed Mean (*PTM*), (Zuo (2006)) as

$$PTM(\alpha,F) = \frac{\int_{PR(\alpha,F)} x\, w_1(PD(x,F)F(dx))}{\int_{PR(\alpha,F)} w_1(PD(x,F)F(dx))}, \tag{6}$$

where $w_1(.)$ is a suitable (bound) weight function on [0, 1]. *PTM* is highly robustness and efficiency α=0 and the famous degenerates PTM into the Stahel-Donoho location estimators (Stahel 1981; Donoho and Gasko 1992), i.e. the Projection Weighted Mean (PWM) and Projection Weighted Scatter (PWS)

$$PWM(F) = \frac{\int x\, w_1(PD(x,F))F(dx)}{\int w_1(PD(x,F))F(dx)}, \tag{7}$$

$$PWS(F) = \frac{\int (x - PWM(F))(x - PWM(F))^{T} w_2(PD(x,F))F(dx)}{\int w_2(PD(x,F))F(dx)}, \tag{8}$$

where $PWM(F)$ and $PWS(F)$ is the aforementioned Stahel-Donoho location and scatter estimator, $w_2(.)$ denotes the weight function on [0, 1] based on projection depth outlying function (μ(*F*), $Q_n(F)$) as respectively. Note that the projection depth and its associated estimators such as*PTM (F), PWM (F)*and*PWS (F)*to be well defined, certain monotony conditions are required as follows:

$$\int w_i\big(PD(x,F)\big)F(dx) > 0,$$

$$\int \|x\|^i w_i\big(PD(x,F)\big)F(dx) < \infty, i = 1,2.$$

with a finite sample $X^n = \{X_1, X_2, ..., X_n\}$ from X andF_n be the corresponding empirical distribution of F based onX^n. By simply replacing F by$F_n$in projection depth and its related estimators can obtain their sample version.

3. ROBUST DISCRIMINANT ANALYSIS

Let p be the variable with *n*observations that are sampled from*l* different populations $\pi_1,\ldots,\pi_l$. The discriminant analysis settingis in the membership of each observation with respect to the populations, i.e., the data points into*l*groups with $n_1, n_2, \ldots, n_l$ observations. Trivially, $\sum_{j=1}^{l} n_j = n$. Therefore, then the observations by $\{x_{ij}; j = 1,...,l; i = 1,...,n_j\}$. Based on the initial estimates $\mu_{j,0}$ and $S_{j,0}$ are computed for each observation $x_{ij}$of group *j*and its (preliminary) robust distance is given by

$$RD^0_{ij} = \sqrt{\left(x_{ij} - \mu_{j,0}\right)^t S^{-1}_j \left(x_{ij} - \mu_{j,0}\right)}.$$

The assign weight 1 to $x_i$if

$$RD^0_{ij} > \sqrt{\chi^2_{P,0.975}}.$$

The reweighted projection depth estimator for group *j* is then obtained as the median$PWM_j$and the scatter matrix$PWS_j$of those observations of group *j* with weight 1. It is shown that this reweighting step increases the finite-sample efficiency of the projection depth estimator considerably, whereas the breakdown value remains the same. These robust estimates of location and scatter now allow us to flag the outliers in the data, and to obtain more robust estimates of the membership probabilities. First compute the robust distance for each observation $x_{ij}$from group *j*,

$$RD_{ij} = \sqrt{\left(x_{ij} - PWM_{j,0}\right)^t PWS^{-1}_{j,0} \left(x_{ij} - PWM_{j,0}\right)}.$$

One can consideran observation $x_{ij}$is an outlier if and only if

$$RD_{ij} > \sqrt{\chi^2_{p,0.975}}.$$

Further, the projection depth estimates$PWM_j$and $PWS_j$are obtained for each group, and then the individual covariance matrixes are pooled together for further computation.

Let$n_j$denote the number of non-outliers in group *j*, and $n = \sum_{j=1}^{l} n_j$, then the robustly estimate the membership probabilities as

$$P_j^R = \frac{n_j}{n}.$$

Note that the usual estimates implicitly assume that all the observations have been correctly assigned to their group. It is however also possible that typographical or other error has occurred when the group numbers were recorded. The observations that are accidently put in the wrong group will then probably show up as outliers in that group, and so they will not influence the estimates of the membership probabilities. Of course, if one is sure that this kind of error is not present in the data, one can still use the relative frequencies based on all the observations.

4. RESULTS AND DISCUSSION

4.1. *Simulation (Computing Projection Depth values)*

A simulation study is performed to compare the efficiency of the various notions of projection depth procedures. To illustrate this 25 sample points are simulated from multivariate normal distribution with the mean vector $\mu = (1,1)$ and the covariance matrix $\Sigma = I_2$

The obtained finite number of optimal direction vectors under the exact projection depth values of the sample points with respect to the data cloud χ^n is reported in Table 1 and is given in Appendix. For the sake of comparison, it is also computed the approximate projection depth values based on 5×10^4 random direction vectors. It is observed from the table, the exact projection depth values are almost greater than the random projection depth values by considering all the pairs. Further it is noted that the exact projection depth values is greater than the random projection depth values produced by the pair (*Mean*, *SD*). It is concluded that the pairs (*Median*, *MAD*) and (*Median*, Q_n) produces similar depth values under exact and random projections.

The projection depth size plots under exact and random projections are displayed in the figure 1.It is noted that the size of the plotted points is increases when the depth values increases. That is, the plotted pointsare in bigger size when the depth value is large. Again, thedepth central points are largerelative to those on the skirts. This is a confirmation that the projection depth provides a center-outward ordering for the given data cloud.

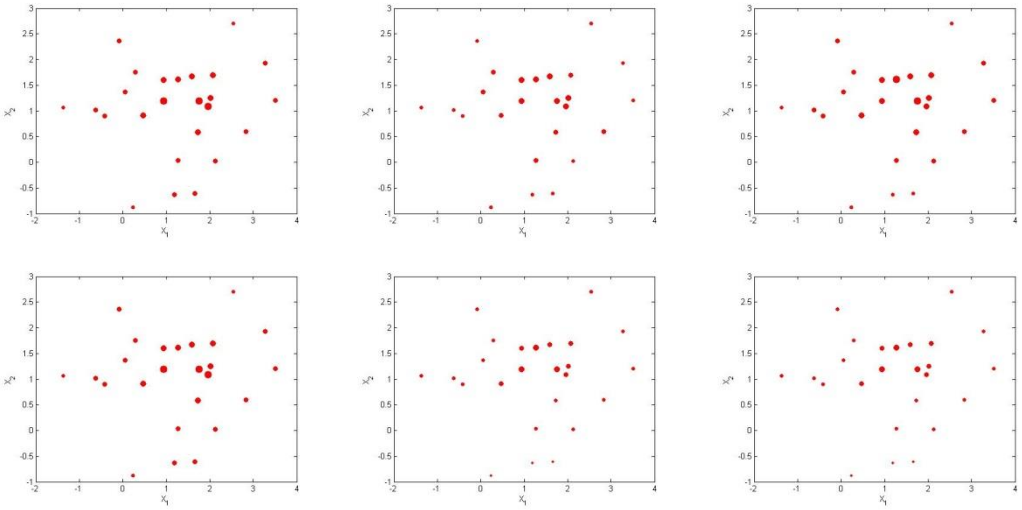


Figure 1 Projection Depth-Size Plots

4.2. *Simulation (Computing Projection-based depth and its associated estimators)*

In order to compare the projection-based depth and its associated estimators, 100 datapoints were generated from the normal distribution with the mean vector μ=(1, 1) and covariance matrix $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Further, the location and scale estimates for the generated data is computed which are as follows:μ = (1.1271,1.0392),Med = (1.1476,1.0706) and $\Sigma = \begin{pmatrix} 1.0481 & -0.1097 \\ -0.1097 & 1.1041 \end{pmatrix}$ which are mean, median and covariance respectively. The estimated Projection based Median, Weighted Median and Trimmed Median under the three pairs (Mean, SD), (Median, MAD) and (Median, Q_n) are summarized in table 2 and 3.

Further, the study was extended with contamination. The data generated with μ = (-4, -4), ∑ = $4I_P$, and the level of contamination 5%, 10% and 15% were considered, and then the same experiment was performed. For the contaminated data, the computed location and scatter values mean, median and covariance are μ = (-0.5293, -0.7022), med = (-0.0333,-0.3286) and $\Sigma = \begin{pmatrix} 4.1224 & 2.5105 \\ 2.5105 & 4.2617 \end{pmatrix}$ respectively.The estimated Projection based Median, Weighted Median and Trimmed Median under the three pairs (Mean, SD), (Median, MAD) and (Median, Q_n) are also summarized in table 2 and 3.

Table2 Estimated Projection Depth Location Estimators (with/without contamination)

Error	Estimators	Projection Depth Procedures		
		(Mean, SD)	(Median, MAD)	(Median, Q_n)
0.00	PM	(1.1271, 1.0392)	(1.0807, 1.0724)	(1.0830, 1.0661)
	PWM	**(1.1326, 1.0458)**	**(1.1402, 1.0538)**	**(1.1326, 1.0458)**
	PTM	(1.1463, 1.0520)	(1.1329, 1.0589)	(1.1463, 1.0520)
0.05	PM	(0.9082,0.6899)	(1.0198,0.8114)	(1.0211,0.8138)
	PWM	**(1.0291,0.8099)**	**(1.0458,0.8360)**	**(1.0501,0.8346)**
	PTM	(1.0893,0.8661)	(1.1041,0.8980)	(1.2106,0.8844)
0.10	PM	(1.1271,1.0392)	(1.0825,1.0593)	(1.0830,1.0661)
	PWM	**(1.1359,1.0366)**	**(1.1322,1.0412)**	**(1.1325,1.0457)**
	PTM	(1.1499,1.0432)	(1.1261,1.0388)	(1.1464,1.0519)
0.15	PM	(0.3484,0.2410)	(0.7592,0.5882)	(0.8110,0.6337)
	PWM	**(0.6903,0.4843)**	**(0.7928,0.5825)**	**(0.8904,0.6602)**
	PTM	(0.8037,0.5591)	(1.0345,0.7809)	(1.1385,0.8555)

Table3 Estimated Projection depth weighted Scatter Estimators (with/withoutcontamination)

Error	Projection Depth Procedures		
	(Mean, SD)	(Median, MAD)	(Median, Q_n)
0.00	$\begin{pmatrix} 0.9501 & -0.0627 \\ -0.0627 & 0.9424 \end{pmatrix}$	$\begin{pmatrix} 0.9221 & -0.0268 \\ -0.0268 & 0.8483 \end{pmatrix}$	$\begin{pmatrix} 0.9501 & -0.0627 \\ -0.0627 & 0.9424 \end{pmatrix}$
0.05	$\begin{pmatrix} 1.1877 & 0.3062 \\ 0.3062 & 1.3157 \end{pmatrix}$	$\begin{pmatrix} 1.0857 & 0.2550 \\ 0.2550 & 1.1774 \end{pmatrix}$	$\begin{pmatrix} 1.1133 & 0.2219 \\ 0.2219 & 1.2143 \end{pmatrix}$
0.10	$\begin{pmatrix} 0.9424 & -0.0730 \\ -0.0730 & 0.9526 \end{pmatrix}$	$\begin{pmatrix} 0.9324 & -0.0653 \\ -0.0653 & 0.9279 \end{pmatrix}$	$\begin{pmatrix} 0.9507 & -0.0630 \\ -0.0630 & 0.9431 \end{pmatrix}$
0.15	$\begin{pmatrix} 2.9177 & 1.6544 \\ 1.6544 & 2.4875 \end{pmatrix}$	$\begin{pmatrix} 2.1899 & 1.1662 \\ 1.1662 & 1.9967 \end{pmatrix}$	$\begin{pmatrix} 1.7404 & 0.7348 \\ 0.7348 & 1.6890 \end{pmatrix}$

It is observed from the above tables, the estimated location and scatter values are close to the actual value under the three pair of estimators when there is no contamination. Further, it is noted that the pair (Median, Q_n) can tolerate certain amount of contamination, specifically, one can see that the contamination level is 15%, the results get affected under the pairs (Mean, SD) and (Median, MAD) but not in the case of (Median, Q_n). It is concluded that, the impact of the outliers on*(Median, Qn)* are very limited.The estimated location points under the three pairs along with data points with/without contaminations are displayed in the form of scatter plots in Figure.2. Figures reveal that the ordinary

mean is placed outside the bulk of the data points by a few outliers; while other projection depth based location estimators are positioned among the majority of the data.

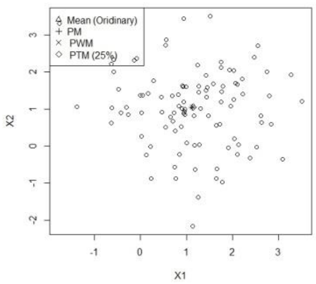
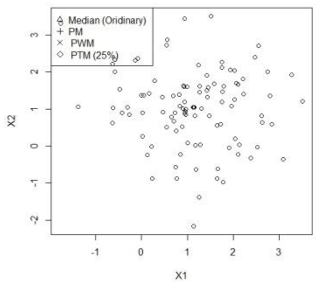
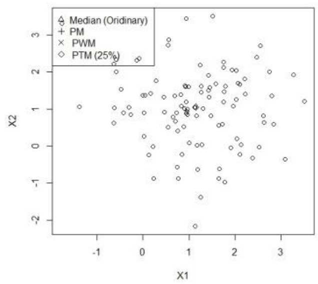
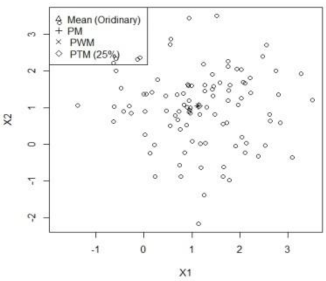
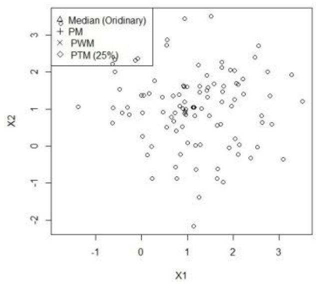
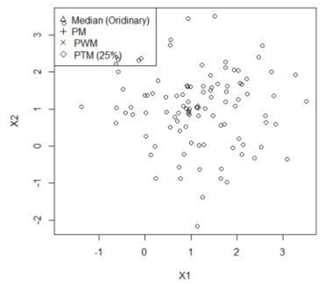
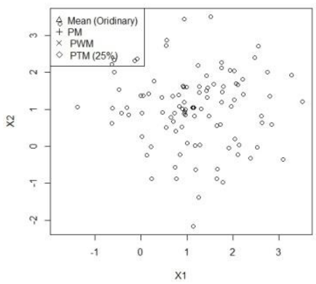
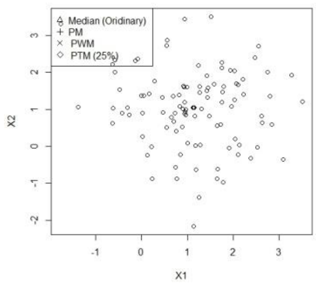
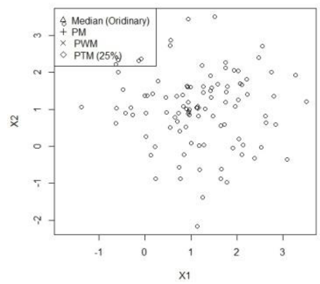
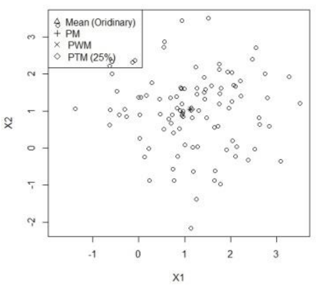
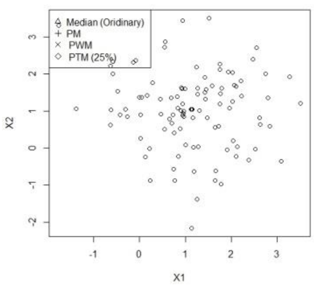
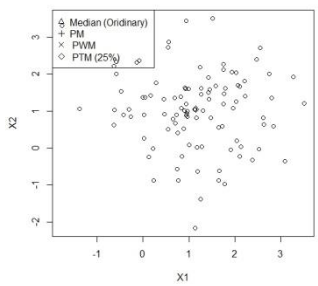
	(Mean,SD)	(Median, MAD)	(Median, Q_n)
e=0.00			
e=0.05			
e=0.10			
e=0.15			

Figure 2 Projection Depth-Size Plots (PWM, PWS)

The location and scatter estimators confirm high robustness of projection depth and its associated estimators (Zuo 2003, 2006).It is worthy to note that, during the computation of *PWM*, *PTM* and *PWS*; weight functions$w_i(.)$, i =1,2,used here as suggested by Zuo and Cui (2005). Further the projection depth contours applied to various projection-based depth procedures under various level of contaminations are display in figure 3 and is given in Appendix. The contours indicate a similarity in the structures of the projection depth procedures (Median, MAD) and (Median, Q_n) which are both unlike the results of the procedure (Mean, SD).

5. APPLICATIONS IN DISCRIMINANT ANALYSIS

5.1. *Real data*

This section presents the performance of projection depth based SDE in robust linear discriminant analysis by computing misclassification probabilities with three pairs of location and scatter approach. It is considered a data set with two groups (Johnson and Wichern (2009)). The data description is as follows: Two different groups: π_1 is ridingmover owners and π_2 is without riding movers to identify the best sales prospects. The owners or non-owners on the basis ofthe variables x_1(income), x_2(lot size), random sampleofsize n_1(=12) current owners and n_2(=12) current non-owners respectively. Discriminant analysis for these two groups is performed and computed misclassification probabilities under various projection depths based approaches and is given in table 4.

Table 4 Computed misclassification probabilities under various projection depth

Procedures	Misclassification Probabilities		
	π_1	π_2	Average
(Mean, SD)	0.1667	0.1667	0.1667
(Median, MAD)	0.2083	0.2083	0.2083
(Median, Q_n)	0.1667	0.1667	0.1667

The estimated average misclassification probabilities are almost same except the procedure(Median, MAD).

5.2. *Simulation study*

This section presents the results obtained under various projection depth based approaches under simulating environment with/without contaminations (Location and Scale). In this context, two groups (g=2) with two variables (p=2) are considered to simulate the data. The data were generated under thenormal distribution which hasthe covariance matrices $\sum_1=I_P$ and $\sum_2= 2I_P$ with means μ_1= (1, 1) and μ_2= (5, 5)under sample sizes of 50 and 100. The various levels of contaminations such as 5%, 10%, 15%, 20%, 25%, 30%, 35% and 40% were considered in all cases. The obtained results with the contamination levels 0%, 5%, 10% and 15% are same and the results based on the remaining contaminations are displayed in the table 5.

Table5 Computed misclassification probabilities under various projection depths with contaminations

n_1=n_2=50				n_1=n_2=100		
Eror	(Mean, SD)	(Median, MAD)	(Median, Q_n)	(Mean, SD)	(Median, MAD)	(Median, Q_n)
0.20	0.0337	0.0227	0.0118	0.0168	0.0113	0.0058
0.25	0.2083	0.1429	0.1250	0.0824	0.0718	0.0058
0.30	0.4667	0.4731	0.4302	0.1746	0.1649	0.0120
0.35	0.4845	0.4792	0.4681	0.3109	0.3109	0.1236
0.40	0.4896	0.5000	0.4792	0.3827	0.3827	0.2065

On comparing the average probability of misclassification values in the above table, it is evident that the procedures (Median, MAD) and (Median, Q_n) produces less when compared with (Mean,SD). Also, it is observed that when sample size increases the misclassification probabilities decreased under all the procedures. It is concluded that the procedure (Median, Q_n) performs better than the other two procedures. It shows that it is superior to the other two procedures when the level of contamination increases.

6. CONCLUSION

Location and scatter estimator play vital role in almost all statistical data analyses. The conventional estimates, sample mean vector and covariance matrix are very sensitive when the outlying observations in the data. In order to obtain the reliable location and scatter estimate, data depth approaches attract the researchers now-a-days. This paper proposes a projection based data depth approach to compute location and scatter estimate, namely (Median, Q_n). Further the superiority of the proposed estimator has been studied under real and simulation by applying it in discirminant analysis by computing the misclassification probabilities with various other projection based depth approaches (Mean, SD) and

(Median, MAD). The simulation study shows that the projection depth based on the mean and standard deviation fails to produce reliable results when compared with the other projection depth procedures. It is noted that (Median, MAD) and (Median, Q_n) performs well over the (Mean, SD). The study concluded that the proposed projection depth procedure (Median,Q_n)shows that its superiority over the other procedures (Mean, SD) and (Median, MAD), in the context of tolerance level of contaminations and misclassification rate.

Acknowledgement

This research work was funded by the Rajiv Gandhi National Fellowship (RGNF) programme, UGC, New Delhi and carried out at Department of Statistics, Bharathiar University, Tamil Nadu and India.

REFERENCES

1. Donoho, D. L., and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. . *The Annals of Statistics*. 20, 1803–1827.
2. Halin, M., Paindaveine, D., and Siman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: From L1 optimization to halfspace depth. *The Annals of Statistics*, 38(2), 635-669.
3. Hu, Y., Li, Q., Wang, Y., and Wu, Y. (2012). Rayleigh projection depth. *Computational Statistics*, 27, 523-530.
4. Hubert, M. and Van Driessen, K. (2004). Fast and Robust Discriminant Analysis. *Computaional Statistics & Data Analysis*, 45,301,320.
5. Johnson, R.A., and Wichern, D.W (2009). Applied multivariate analysis, 5th ed. Prentice Hall, Englewood Cliffs, New Jersey.
6. Liu, R.Y. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*, 18, 191-219.
7. Liu, R.Y. (1992). Data depth and multivariate rank test. In: L1- Statistical Analysis and Related Methods, 279-294. North-Holland, Amsterdam.
8. Liu, R.Y., Parelius, J.M., and Singh, K. (1999). Multivariate analysis by data depth: Descriptive Statistics, Graphics and Inference. *The Annals of Statistics*, 27, 783-858.
9. Liu, X., and Zuo, Y. (2014). Computing projection depth and its associated estimators. *Stat.Comput.,* 24, 51-63.
10. Liu, X.H., Zuo, Y.J., Wang, Z.Z. (2011). Exactly computing bivariate projection depth contours and median. *Preprint*.
11. Rousseeuw, P.J., and Hubert, M. (1999). Regression depth. *Journal of the American Statistical Association*, 94, 388-433.
12. Stahel, W.A. (1981). Breakdown of covariance estimators. *Research Report* 31, 1029-1036.
13. Tukey, J.W. (1975). Mathematics and the picturing of data. In: *Proceedings of the International Congress of Mathematicians*, pp. 523-531. *Canadian Mathematical Congress*, Montreal.
14. Zuo, Y. (2003). Projection-based depth functions and associated medians, *The Annals of Statistics*, 31, 1460-1490.
15. Zuo, Y., and Serfling, R. (2000). General notions of statistical depth function. *The Annals of Statistics*, 28, 461–482.
16. Zuo, Y.J. (2006). Multidimensional trimming based on projection depth. *The Annals of Statistics*, 34, 2211-2251.
17. Zuo, Y.J., and Lai, S.Y. (2011). Exact computation of bivariate projection depth and Stahel-Donoho estimator. *Computational Statistics & Data Analysis*, 55(3), 1173-1179.
18. Zuo, Y.J., and Cui, H.J. (2005). Depth weighted scatter estimators. *The Annals of Statistics*, 33, 381-413.
19. Zuo, Y.J., Cui, H.J., He, X.M. (2004). On the Stahel-Donoho estimators and depth – weighted means for multivariate data. *The Annals of Statistics*, 32(1), 189-218.

Appendix A.

Table 1 Computed Exact and Random Projection Depth Values

Index	ExPDV			RndPDV (50000)		
	(Mean, SD)	(Median, MAD)	(Median, Q_n)	(Mean, SD)	(Median, MAD)	(Median, Q_n)
1	2.157455	0.192624	0.295180	0.316700	0.192652	0.192629
2	0.325019	0.500000	0.624752	0.754699	0.500000	0.500000
3	1.381532	0.207411	0.366216	0.419877	0.207424	0.207452
4	1.427049	0.269415	0.413776	0.411965	0.269420	0.269431
5	0.756	0.448	0.593	0.569	0.448	0.448

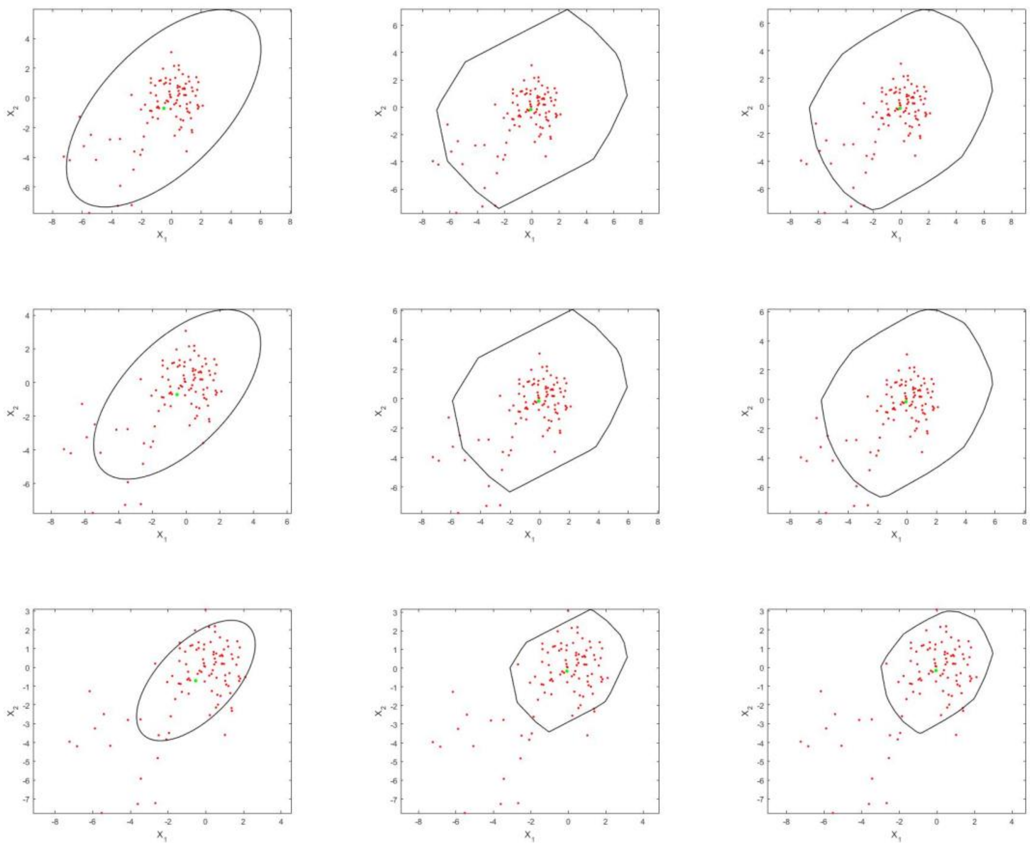
	764	481	490	181	493	500
6	1.110 552	0.232 646	0.403 110	0.473 809	0.232 665	0.232 688
7	0.578 716	0.468 108	0.627 823	0.633 426	0.468 113	0.468 110
8	2.088 440	0.210 192	0.330 685	0.323 738	0.210 192	0.210 194
9	1.365 669	0.266 899	0.389 780	0.422 712	0.266 935	0.266 904
10	1.849 254	0.231 068	0.370 514	0.350 968	0.231 075	0.231 072
11	0.711 800	0.441 931	0.577 139	0.584 118	0.441 932	0.441 936
12	0.968 409	0.375 329	0.507 785	0.508 010	0.375 332	0.375 342
13	0.664 505	0.431 879	0.595 967	0.600 772	0.431 887	0.431 883
14	0.654 543	0.323 455	0.512 907	0.604 381	0.323 470	0.323 503
15	2.226 154	0.162 738	0.287 439	0.309 962	0.162 756	0.162 757
16	1.894 023	0.159 594	0.296 955	0.345 536	0.159 607	0.159 621
17	1.177 729	0.316 469	0.442 848	0.459 100	0.316 471	0.316 484
18	0.438 135	0.523 140	0.677 843	0.695 193	0.523 145	0.523 143
19	1.925 905	0.211 008	0.364 037	0.341 707	0.211 022	0.211 023
20	1.534 453	0.245 126	0.363 464	0.394 543	0.245 160	0.245 131
21	1.860 163	0.165 077	0.306 217	0.349 630	0.165 093	0.165 103
22	1.083 207	0.302 623	0.431 504	0.480 021	0.302 626	0.302 646
23	0.647 058	0.395 693	0.534 016	0.607 130	0.395 696	0.395 718
24	1.874 344	0.228 289	0.355 446	0.347 905	0.228 299	0.228 294
25	0.657 895	0.500 000	0.637 665	0.603 174	0.500 000	0.500 000

(Mean,SD) (Median, MAD) (Median, Q_n)

$\alpha = 0.05$

$\alpha = 0.10$

$\alpha = 0.20$



$\alpha = 0.30$

$\alpha = 0.40$

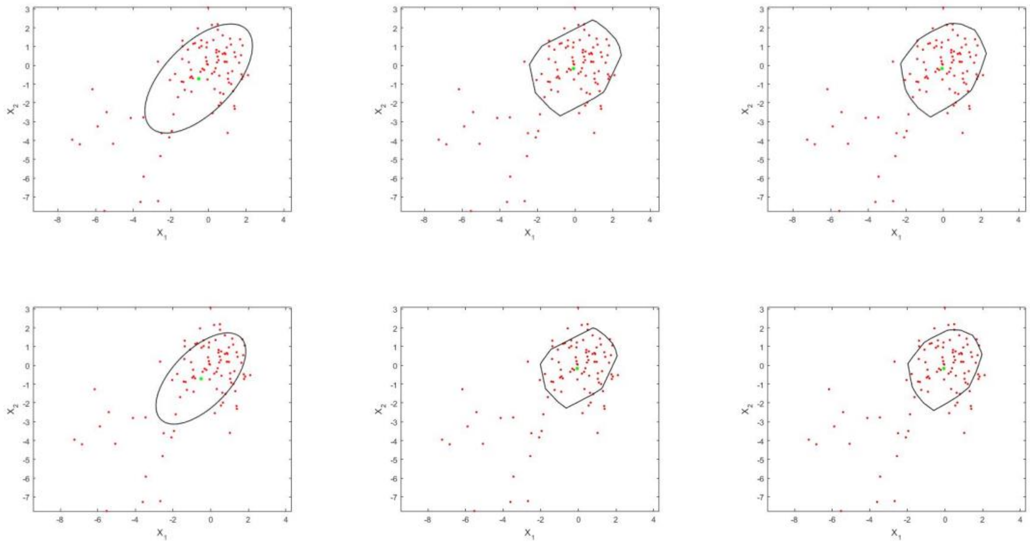


Figure 3 Projection Depth Contour Plots under various procedures with level of contaminations