Far East Journal of Mathematical Sciences (FJMS)
Volume 74, Number 2, 2013, Pages 249-268
Published Online: March 2013
Available online at http://pphmj.com/journals/fjms.htm
Published by Pushpa Publishing House, Allahabad, INDIA

ON ROBUST MAHALANOBIS DISTANCE ISSUED FROM MINIMUM VECTOR VARIANCE

Hazlina Ali and Sharipah Soaad Syed Yahaya

School of Quantitative Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok, Kedah, Malaysia
e-mail: hazlina@uum.edu.my
sharipah@uum.edu.my

Abstract

Detecting outliers in high dimension datasets remains a challenging task. Under this circumstance, robust location and scale estimators are usually proposed in place of the classical estimators. Recently, a new robust estimator for multivariate data known as minimum variance vector (MVV) was introduced. Besides inheriting the nice properties of the famous MCD estimator, MVV is computationally more efficient. This paper proposes MVV to detect outliers via Mahalanobis squared distance (MSD). The results revealed that MVV is more effective in detecting outliers and in controlling Type I error compared with MCD.

© 2013 Pushpa Publishing House

2010 Mathematics Subject Classification: 65C60.

Keywords and phrases: robust estimation, Mahalanobis distance, minimum vector variance, outliers detection.

Submitted by K. K. Azad

Received October 16, 2012

1. Introduction

Datasets that are multivariate in nature are common in industries. With the enhancement in computer technology, researchers have no problem working with enormous multivariate datasets. Nevertheless, as the dimensions of the data increase, the presence of outliers in the datasets will also increase. The presence of outliers can hardly be detected using naked eyes when the dimension is more than 2. This is the risk the researchers have to be cautious about when working with large datasets of high dimensions. Thus, a reliable method is needed to detect outliers especially for this sort of datasets. Fortunately, researchers in this area are relentlessly searching for such methods. Mahalanobis square distance (MSD) is a prominent method for outlier detection. For multivariate normally distributed data, MSDs are approximately chi-square distributed with p degrees of freedom (χ_p). An outlier would then be defined as an observation having larger distance value than the critical value, i.e., $\chi_{p,\alpha}$ (Mardia et al. [15]) and (Serfling [24]). However, the performance of MSD suffers from masking and swamping effect due to the non-robustness (sensitiveness) of the classical estimators. These estimators are sensitive to outliers and will be greatly influenced by their presence. Consequently, it is unlikely to use the MSD to find outliers since MSD itself is sensitive to outliers. Under this circumstance, robust estimators are needed to prevent these errors from influencing the computation of MSD, which will result in spurious detection of outliers.

A wide range of robust estimators of multivariate location and scatter are available, see Maronna and Zamar [17] and Maronna et al. [16] for a review. Nonetheless, the Minimum Covariance Determinant (MCD) estimator introduced by Rousseeuw [20] has received a considerable attention by scientific community and widely used in practice. This is due to the fact that MCD gives the exact solution (Hadi [9] and Hubert et al. [14]), and it also has good theoretical properties with affine equivariance, high breakdown value, bounded influence function and also has a better convergence rate (Butler et al. [4]; Croux and Haesbroeck [5]). Due to computational complexity of the exact MCD estimators especially for large n, several improved versions of MCD algorithm are available. For example, feasible solution algorithm (FSA) in Hawkins [10] and Hawkins and Olive [11], MULTOUT in Woodruff and Rocke [26], Fast MCD algorithm in Rousseeuw and van Driessen [22], block adaptive computationally-efficient outlier nominators (BACON) in Billor et al. [2] and improved Fast MCD algorithm in Hubert et al. [14]. However, the main contribution in this domain is the Fast MCD algorithm which has been available in many computer packages such as Matlab, R, SAS, and S-Plus. Furthermore, its applications can be found in a very wide spectrum area, for example, multivariate statistical process control, multivariate process capability analysis, information sciences, data depth, data mining, etc. Thus, this shows that Fast MCD is very well accepted. Nevertheless, Fast MCD is not without limitation. The use of “minimizing covariance determinant” as the objective function in data concentration process can be computationally laborious especially when the dataset is of high dimension. On the other hand, as Angiulli and Pizzuti [1] have pointed out, the computational efficiency is as important as effectiveness. To overcome the weaknesses of Fast MCD algorithm, Herwindiati [12] proposed minimum vector variance (MVV) as an alternative measure of multivariate data concentration.

The use of vector variance in place of covariance determinant as the objective function of the stopping rule is comprehensively discussed in Djauhari [7]. Herwindiati et al. [13] revealed that MVV was successfully used as an objective function in Fast MCD algorithm to substitute the MCD criterion. The findings showed that MVV is computationally more efficient than Fast MCD. Motivated by their findings and to fill some of the gaps in the research, this paper investigates on the robustness and effectiveness of MVV estimator in detecting outliers via MSD covering both correlated and uncorrelated variables. The performance evaluation is measured in terms of Type I error and probability of outlier(s) detection, respectively. Section 2 recalls on minimum vector variance (MVV) estimator. Section 3 will report on the results of simulation experiments which strongly indicate the advantage of MVV. The concluding remarks in Section 4 will close this paper.

2. Minimum Vector Variance (MVV)

The use of determinant in the computation has its drawback. In higher dimensions, computation of determinant is more complicated. As an alternative measure to the long and tedious computation of covariance determinant in data concentration, MVV uses vector variance (VV) for a faster algorithm without changing the structures of Fast MCD algorithm. MVV estimators possess three major properties of a good robust estimator i.e. high breakdown point, affine equivariance and computational efficiency (Herwindiati et al. [13]). The main method used in the estimation of MVV is the Mahalanobis squared distances (MSD) which is defined as

$$d_i^2 = (x_i - \mu)^t \Sigma^{-1} (x_i - \mu), \quad i = 1, 2, ..., n, \tag{1}$$

where n represents the number of observations. Consider a dataset $x = \{x_1, x_2, ..., x_n\}$ of p-variate observations and let $H \subseteq X$. The optimal value of the number of data involved in the computation of MVV estimators namely M_{MVV} and S_{MVV} is $h = \left\lfloor \frac{n+p+1}{2} \right\rfloor$ data points which generate a covariance matrix S_{MVV} having minimum $Tr(S_{MVV}^2)$ among all possible sets of h data. However, it is no guarantee that the iteration process for sets of h data from only one initial h-set can generate the final value of $Tr(S_{MVV}^2)$ as a global minimum of the MVV objective function. The approximation of MVV estimators can be obtained by taking many initial choices of h-subsets. To compute the MVV estimators, we used the MVV algorithm proposed in Yahaya et al. [25] by applying concentration step (C-step) for each initial subset, and later choose a specific number of subsets that produce the lowest vector variance. From the algorithm, the location estimator is defined as

$$M_{MVV} = \frac{1}{h} \sum_{i=1}^{h} x_i \tag{2}$$

and the scatter estimator by

$$S_{MVV} = \frac{1}{h} \sum_{i=1}^{h} (x_i - M_{MVV})(x_i - M_{MVV})^t. \tag{3}$$

Scatter estimators are typically calibrated to be consistent for the normal distribution, thus the consistency and correction factors are needed to guarantee Fisher consistency for the scatter estimator and improve its biasness for small sample behaviour. Fisher consistency is a standard concept in robust statistics and it means that the functionals evaluated at the model distribution F return the true parameter values, μ and $\sum$ (Croux and Rousseeuw [6]). We take $c(h)$ as the approximation of consistency factor, where it can be obtained from elliptical truncation in the multivariate normal distribution based on squared distance. If $x_i \sim N(\mu, \Sigma)$, then $c(h)$ is defined as

$$c(h) = \frac{h/n}{P(\chi^2_{p+2} < \chi^2_{p,h/n})}, \tag{4}$$

where $\chi^2_{p,h/n}$ is the h/n-quantile of χ^2_p distribution. This formula is derived by Butler et al. [4] and further discussed in Croux and Haesbroeck [5] based on the functional form of the MCD estimator. Albeit this process guarantees consistency under normal distribution, the consistency factor alone is not sufficient to make the MVV estimator unbiased for small sample sizes. To overcome the insufficiency problem we also include correction factor $\vartheta^{\alpha}_{h,n,p}$. The computation of the factor was based on Pison et al. [18]. Next, we determine the MVV scatter as follows:

$$\vartheta^{\alpha}_{p,n} c(h) S_{MVV} = \frac{\vartheta^{\alpha}_{p,n} c(h)}{h} \sum_{i=1}^{h} (x_i - M_{MVV})(x_i - M_{MVV})^t. \tag{5}$$

3. Simulation Experiments and Results

Before we present the analysis and results of the performance of MVV, we will discuss briefly on the distributional of the robust MSD. The application of robust estimators in place of the mean and covariance structure in MSD will cause the distributional properties of the classical MSD to change (Rousseeuw and van Zomeren [23] and it is difficult to identify the corresponding robust MSD distribution. To demonstrate and compare the performance of MVV estimator with MCD when working on MSD in detecting outliers, we need a better understanding about the distribution of MVV and MCD on robust MSD in order to be able to obtain appropriate cutoff values, but the distribution of the exact MVV estimators is not known in closed form Herwindiati [12]. Even the distribution of the well known robust Mahalanobis distance derived from MCD is still questionable for applications (Riani et al. [19]) and (Fauconnier and Haesbroeck [8]). Hence, in this study, we used quantile in estimating the distribution obtained via Monte Carlo method. This study focuses on datasets of dimensions 2 and 5 $(\text{i.e. } p = 2, 5)$ with reasonable values of sample size *n*. To ensure a smooth computation for MCD, the setting of the starting value for *n* follows the rule of thumb, where $n/p > 5$ (Rousseeuw and van Zomeren [23]). For each combination of *n* and *p*, 5000 samples from standard multivariate normal distribution $MVN_p(0, I_p)$ were generated. The MVV and MCD of mean vector and covariance matrix estimators and their corresponding robust MSD statistics were calculated. Then, the cutoff values for MSD issued from MVV and MCD were based on the 95% quantile of the 5000 statistics.

Performance evaluation

This study compares the performance of MSD procedures issued from the classical estimators, MVV and MCD with regards to robustness (Type I error) and their effectiveness in detecting outliers (probability of detection). To check on the level of robustness, we consider Bradley's liberal criterion of robustness as a reference by comparing the empirical Type I error value with the nominal value, $\alpha = 0.05$. Bradley [3] specified three criteria for robustness namely stringent, moderate, and liberal which are, respectively, defined as $\alpha \pm 0.1\alpha$, $\alpha \pm 0.2\alpha$, and $\alpha \pm 0.5\alpha$. A statistic is considered robust if its empirical Type I error rates lie in one of the ranges. The closer the value to α, the more robust is the statistic or in other words the procedure has better control of Type I error rates. As for the probability of detection, the higher the value, the more effective is the procedure in detecting outliers.

To check on the strengths and weaknesses of the estimators, various scenarios (conditions) were created by simulating the data using different number of observations (*n*), different number of correlated and uncorrelated variables (*p*), and contaminate the data using different proportion of outliers (ε) and several mean shifts values (μ_1). Therefore we have considered a contaminated model by using a mixture of normal,

$$(1-\varepsilon)N_p(0, I_p) + \varepsilon N_p(\mu_1, I_p), \tag{6}$$

where ε is set to be 0, 0.1 or 0.2, while μ_1 is a vector of size *p* with value of 0 (when there is no change), 3 or 5. The mixture model in equation (6) will produce clean and contaminated data from parameters on the left and right side of the plus sign, respectively. To investigate each condition, we randomly generated data using 1000 replications based on the combinations of *n* and *p*. This study also looked into the cases of correlation and uncorrelated variables. For that purpose, the contaminated model for both cases should be modeled differently. To differentiate between the two cases, the uncorrelated case is referred as Case A and the correlated as Case B. Data for Case A were generated using the mixture model in equation (6), while for Case B, the data were generated using the following mixture model

$$(1-\varepsilon)N_p(0, \Sigma_1) + \varepsilon N_p(\mu_1, \Sigma_1), \tag{7}$$

where Σ_1 was set to be a matrix of size *p* with 1 on the main diagonal and 0.9 elsewhere. We used the value of Σ_1 to examine whether correlation between variables affects the probability of detection and Type I error of each MSD procedure. The results of the performance between the different MSD procedure namely the classical MSD (MSD_O), MSD issued from MCD (MSD_{MCD}), and MSD issued from MVV (MSD_{MVV}) are compared and discussed in the following section. As mentioned before, the performance of the each MSD is judged based on its effectiveness (probability) in detecting outliers and the ability to control Type I error rate. The probability of detection was determined by computing the statistic (MSD) using the contaminated data and estimated as the proportion of statistic values that were above the cutoff values. Conversely, to determine the Type I error rates, the statistic was calculated from the clean data.

Results

The results of the investigation are presented in the form of tables and figures for the Type I error and the corresponding probability of detection, respectively. Type I error rates for Case A are presented in Tables 1 and 2 while Case B are presented in Tables 3 and 4. The estimators are considered robust when the empirical Type I error rates are within the Bradley's criteria. The values with asterisks are considered robust. The number of asterisks denotes the level of robustness such that 3 asterisks imply that the value is stringently robust, 2 asterisks are moderately robust, and only 1 asterisk is liberally robust. For the purpose of comparison, the empirical Type I error rates having the smallest difference with the nominal value among the three procedures are highlighted. Thus, across the row, the highlighted cell indicates that the corresponding procedure generates the best result among the others by producing the closest empirical Type I error to the nominal value.

As for Case A, when $p = 2$, the MSD_{MVV} column has the highest number of robust values and the most number of highlighted cells. There are 14 robust values with 6 values fall within the stringent criteria. In contrast, there are only 2 robust values for MSD_{MCD} and only 5 for MSD_{O}. When the dimension increases to $p = 5$, the number of cells with robust values dwindle to 10 for MSD_{MVV} and the number of highlighted cells equals to MSD_{O}. There is improvement in MSD_{O} when almost all of the Type I error rates for this procedure are considered robust.

Table 1. Type I error rate for $p = 2$ in Case A

n	ε	μ_1	MSD_O	MSD_{MVV}	MSD_{MCD}
10	0	0	0.0530***	**0.0520*****	**0.0520*****
	0.1	3	0.0170	**0.0450*****	0.0290*
		5	0.0160	**0.0450*****	0.0250*
	0.2	3	0.0180	**0.0330***	0.0210
		5	0.0180	**0.0330***	0.0110
25	0	0	0.0590**	0.0530***	**0.0480*****
	0.1	3	0.0290*	**0.0390****	0.0280*
		5	0.0230	**0.0390****	0.0290*
	0.2	3	**0.0280***	0.0190	0.0090
		5	**0.0240**	0.0190	0.0050
50	0	0	0.0560**	**0.0540*****	0.0580**
	0.1	3	0.0200	**0.0350***	0.0230
		5	0.0160	**0.0340***	0.0230
	0.2	3	**0.0210**	0.0180	0.0080
		5	0.0160	**0.0170**	0.0060
100	0	0	0.0550	**0.0490*****	0.0460***
	0.1	3	0.0210	**0.0300***	0.0200
		5	0.0160	**0.0290***	0.0200
	0.2	3	**0.0210**	0.0150	0.0050
		5	**0.0160**	0.0150	0.0040
Number of highlighted			5	14	2

Table 2. Type I error rate for $p = 5$ in Case A

n	ε	μ_1	MSD_O	MSD_{MVV}	MSD_{MCD}
30	0	0	0.0460***	**0.0500*****	0.0430**
	0.1	3	0.0280*	**0.0300***	0.0100
		5	0.0260*	**0.0330***	0.0100
	0.2	3	**0.0300***	0.0210	0.0050
		5	**0.0320***	0.0200	0.0000
50	0	0	0.0530***	**0.0490*****	0.0650*
	0.1	3	0.0270*	**0.0350***	0.0130
		5	0.0260*	**0.0370***	0.0130
	0.2	3	**0.0260***	0.0220	0.0040
		5	**0.0250***	0.0230	0.0020
100	0	0	**0.0540*****	0.0380*	0.0320*
	0.1	3	0.0290*	**0.0300***	0.0140
		5	0.0280*	**0.0320***	0.0140
	0.2	3	**0.0300***	0.0170	0.0020
		5	**0.0290***	0.0190	0.0020
200	0	0	**0.0430****	0.0390*	0.0410**
	0.1	3	0.0250*	**0.0350***	0.0200
		5	0.0240	**0.0350***	0.0200
	0.2	3	**0.0270***	0.0220	0.0010
		5	**0.0270***	0.0220	0.0010
Number of highlighted			10	10	0

Table 3. Type I error rate for $p = 2$ in Case B

n	ε	μ_1	MSD_O	MSD_{MVV}	MSD_{MCD}
30	0	0	0.0530***	0.0740*	**0.0520*****
	0.1	3	0.0260*	**0.0540*****	0.0350*
		5	0.0180	**0.0480*****	0.0300*
	0.2	3	0.0270*	**0.0410****	0.0380*
		5	0.0190	**0.0280***	0.0220
50	0	0	0.0590**	0.0900	**0.0480*****
	0.1	3	**0.0390***	0.0640*	0.0380*
		5	0.0290*	**0.0520*****	0.0300*
	0.2	3	0.0370*	**0.0490*****	0.0340*
		5	**0.0280***	0.0270*	0.0160
100	0	0	**0.0560****	0.0980	0.0580**
	0.1	3	0.0310*	0.0660*	**0.0340***
		5	0.0230	**0.0580****	0.0250*
	0.2	3	0.0310*	**0.0380***	0.0280*
		5	0.0230	**0.0350***	0.0090
200	0	0	0.0550***	0.0660*	**0.0460*****
	0.1	3	0.0360*	**0.0530*****	0.0250*
		5	0.0240	**0.0410****	0.0210
	0.2	3	0.0310*	**0.0330***	0.0190
		5	**0.0220**	0.0210	0.0070
Number of highlighted			4	12	4

Table 4. Type I error rate for $p = 5$ in Case B

n	ε	μ_1	MSD_O	MSD_{MVV}	MSD_{MCD}
30	0	0	**0.0460****	0.0980	0.0430**
	0.1	3	**0.0430*****	0.0810	0.0320*
		5	**0.0370***	0.0760	0.0320*
	0.2	3	**0.0420****	0.0680*	0.0360*
		5	0.0350*	**0.0450*****	0.0340*
50	0	0	**0.0530*****	0.1050	0.0650*
	0.1	3	0.0460***	0.0870	**0.0480**
		5	**0.0380***	0.0690*	0.0350*
	0.2	3	**0.0420****	0.0830	0.0390*
		5	0.0360*	**0.0550*****	0.0400**
100	0	0	**0.0540*****	0.0800	0.0320*
	0.1	3	**0.0470*****	0.0630*	0.0340*
		5	0.0380*	**0.0580****	0.0230
	0.2	3	**0.0460*****	0.0660*	0.0340*
		5	0.0370*	**0.0450*****	0.0300*
200	0	0	0.0430**	**0.0500*****	0.0410**
	0.1	3	0.0370*	**0.0440****	0.0400**
		5	0.0340*	**0.0430****	0.0330*
	0.2	3	0.0350*	**0.0470*****	0.0330*
		5	0.0310*	**0.0360***	0.0260*
Number of highlighted			10	9	1

However, the result for MSD_{MCD} worsens, as this procedure can generate robust Type I error rates under ideal condition only, and furthermore it produces no highlighted cells. Across both dimensions (*p*), we notice that MSD_{MVV} in general is in control of Type I error rates even when sample size is small, nevertheless it falls short when $\varepsilon = 0.2$ and $\mu_1 = 5$.

As we move to Case B (refer to Tables 3 and 4), one can observe that when $p = 2$ there are more robust values under MSD_{MVV} column than the other two (refer to Table 3). This procedure also produces the highest number of stringent values and even the highlighted cells outnumbered the rest. However, the nice result for MSD_{MVV} is marred by a few inflated Type I error rates (non-robust) under ideal condition. In contrast, all the non-robust values belong to MSD_{MCD} as well as MSD_O are conservative (small). As one can observe in Table 4, when $p = 5$, there is a remarkable improvement in MSD_O and MSD_{MCD}. All the Type I error rates for MSD_O are robust with a considerable number of stringent values. Under MSD_{MCD} column, we notice that almost all the Type I error rates are robust even though this approach does not produce any stringent values. Despite a reduction in the number of robust values for MSD_{MVV}, this approach produces quite a number of stringent robust Type I error rates. However, the problem of inflated Type I error rates in the case of $p = 2$ continues to occur and worsen when $p = 5$, MSD_{MVV} even losses to MSD_O in the count of highlighted cells. The comparison between the two dimensions for Case B shows that the Type I error rates for most conditions of MSD_{MVV} increase when the dimension increase.

The investigation continues to check on the effectiveness of the procedures in detecting changes. The results for probability of detection when the variables are assumed to be uncorrelated (Case A) are displayed in Figures 1 and 2. Overall, MSD_{MVV} and MSD_{MCD} perform better than the MSD_O when outliers are present. As the number of variables, sample size and the shift in the mean increase, the probability of detection for both procedures approaches 1. However, in situation where there are 20% outliers with mean shift equals to 3, MSD_{MVV} performs better than the MSD_{MCD}. Under most conditions, MSD_{MVV} is on par with MSD_{MCD} in detecting outliers under this case (uncorrelated).

The effectiveness of the procedures in detecting outliers when the variables are assumed to be correlated (Case B) is illustrated in Figures 3 and 4. From the figures, we observe that the MSD_{MVV} outperforms the other two procedures in detecting outliers for almost all conditions. The lines representing MSD_{MVV} are quite a distance above the lines for MSD_{MCD} and the MSD_{O}. For this case, MSD_{MCD} performs as bad as the classical MSD in detecting outliers.

4. Conclusion

Apart from having the same properties with the popular MCD, MVV has the edge over MCD with respect to computational efficiency. Thus, in this paper, we proposed to use this estimator as alternatives to the classical mean vector and covariance matrix in Mahalanobis squared distance (MSD) procedure. The investigation which covered both the uncorrelated and correlated situations revealed that in general, the MSD issued from MVV estimators outperformed the classical MSD and MSD issued from MCD in terms of probability of detection and did considerably good in controlling Type I error rates. However, this approach needs further consideration to identify the problem of inflated Type I error rates that occur even under ideal conditions of correlated case.

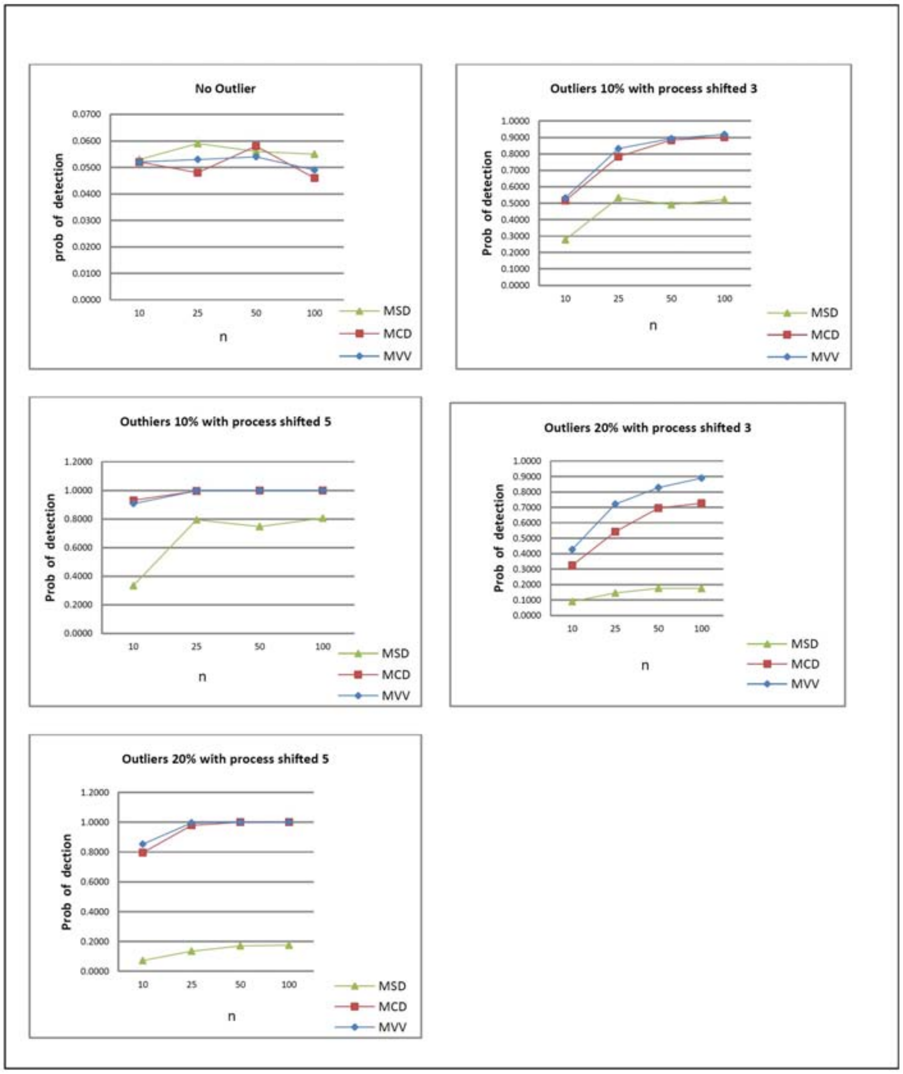
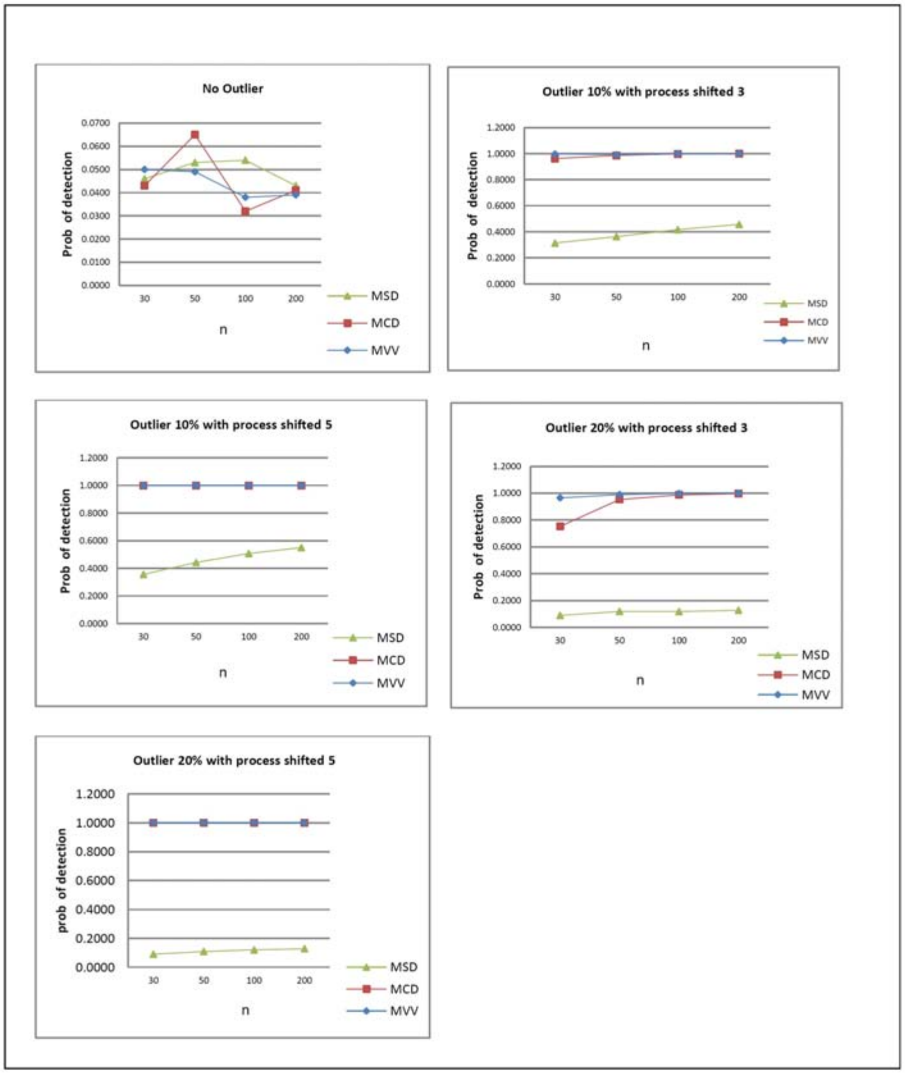


Figure 1. Probability of signal when $p = 2$ for Case A.



No Outlier

Outlier 10% with process shifted 3

Outlier 10% with process shifted 5

Outlier 20% with process shifted 3

Outlier 20% with process shifted 5

Figure 2. Probability of signal when $p = 5$ for Case A.

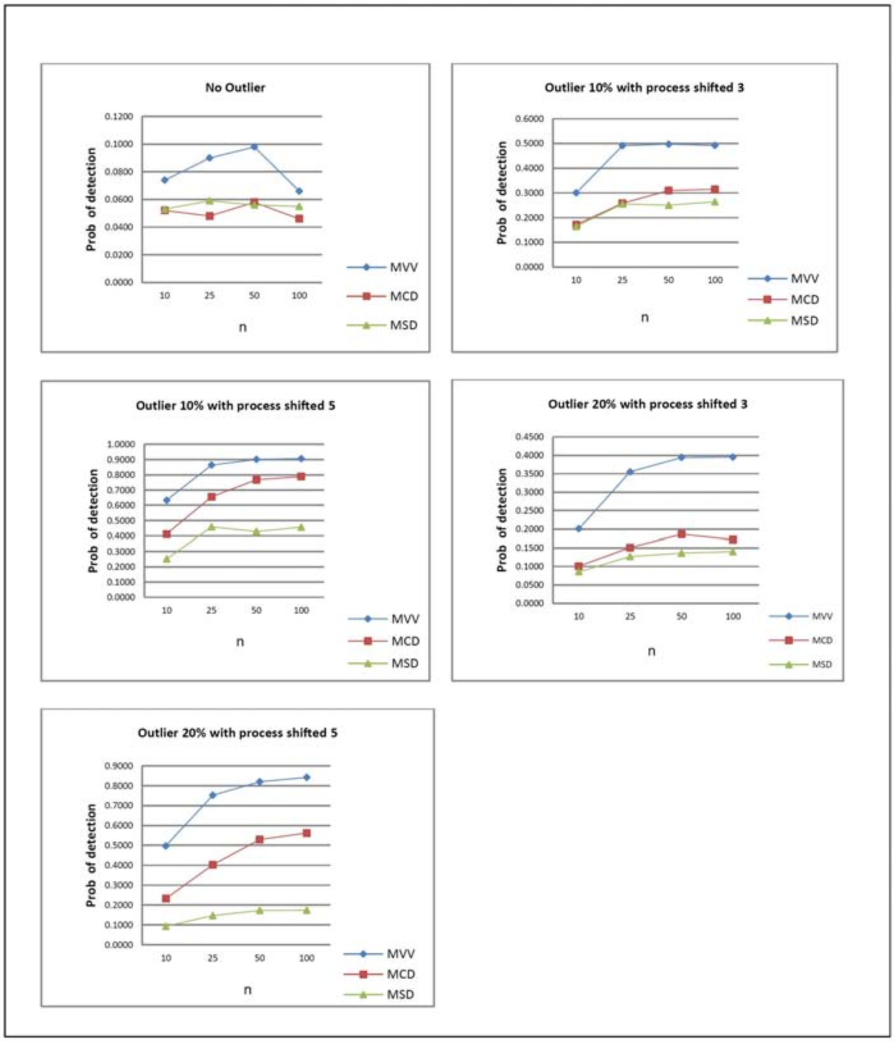


Figure 3. Probability of detection when $p = 2$ for Case B.

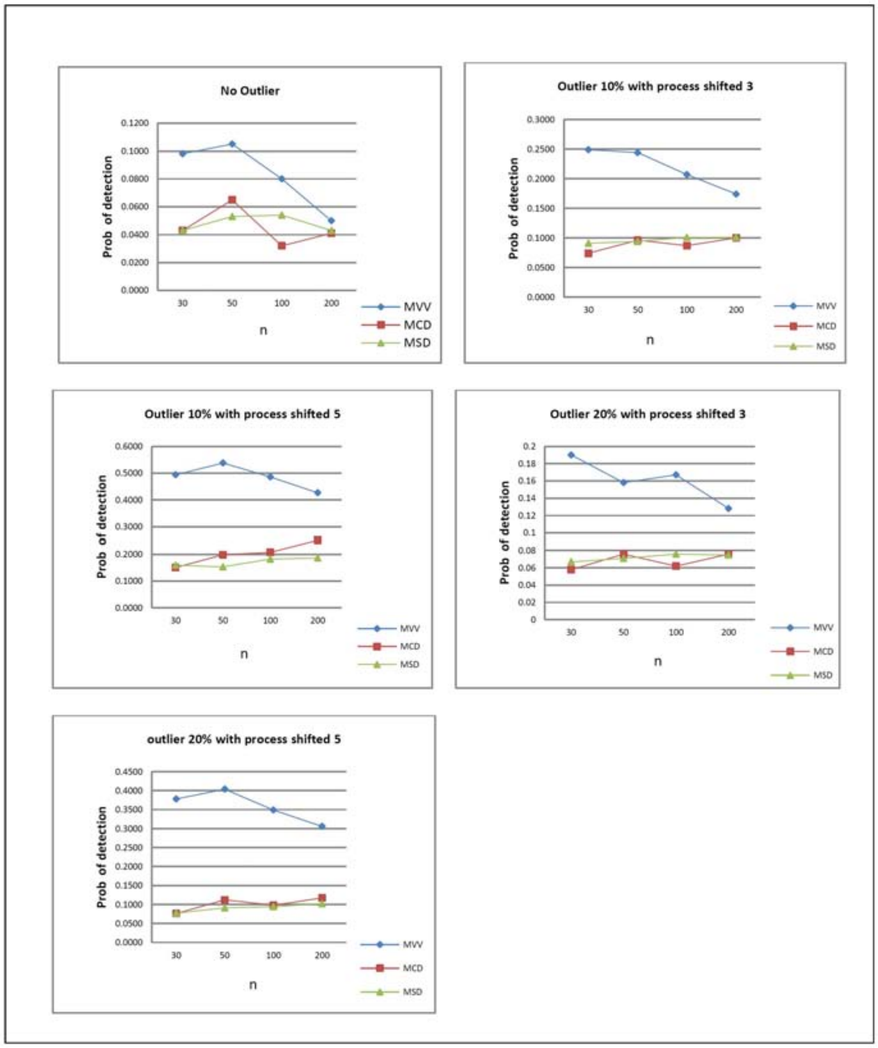


Figure 4. Probability of detection when $p = 5$ for Case B.

References

[1] F. Angiulli and C. Pizzuti, Outlier mining and large high-dimensional data sets, IEEE Trans. Knowledge Data Engineering 17(2) (2005), 203-215.

[2] N. Billor, A. S. Hadi and P. F. Velleman, BACON: blocked adaptive computationally efficient outlier nominators, Comput. Statist. Data Anal. 34 (2000), 279-298.

[3] J. V. Bradley, Robustness? British J. Math. Stat. Psychol. 31 (1978), 144-152.

[4] R. W. Butler, P. L. Davies and M. Jhun, Asymptotics for the minimum covariance determinant estimator, Ann. Statist. 21 (1993), 1385-1400.

[5] C. Croux and G. Haesbroeck, Influence function and efficiency of the minimum covariance determinant scatter matrix estimator, J. Multivariate Anal. 71 (1999), 161-190.

[6] C. Croux and P. J. Rousseeuw, A class of high-breakdown scale estimators based on subranges, Comm. Statist. Theory Meth. 21(7) (1992), 1935-1951.

[7] M. A. Djauhari, A measure of multivariate data concentration, J. Appl. Prob. Stat. 2 (2007), 139-155.

[8] C. Fauconnier and G. Haesbroeck, Outliers detection with the minimum covariance determinant estimator in practice, Stat. Methodol. 6(4) (2009), 363-379.

[9] A. S. Hadi, Identifying multivariate outlier in multivariate data, J. R. Stat. Soc. B 53 (1992), 761-771.

[10] D. M. Hawkins, The feasible solution algorithm for the minimum covariance determinant estimator in multivariate data, Comput. Statist. Data Anal. 17 (1994), 197-210.

[11] D. M. Hawkins and D. J. Olive, Improved feasible solution algorithm for high breakdown estimation, Comput. Statist. Data Anal. 30 (1999), 1-11.

[12] D. E. Herwindiati, A new criterion in robust estimation for location and covariance matrix, and its application for outlier labeling, Ph.D. Thesis, Institut Teknologi Bandung, 2006.

[13] D. E. Herwindiati, M. A. Djauhari and M. Mashuri, Robust multivariate outlier labeling, Comm. Statist. Comput. Simul. 36 (2007), 1287-1294.

[14] M. Hubert, P. J. Rousseeuw and S. Van Aelst, Multivariate outlier detection and robustness, Handbook of Statistics, Volume 23: Data Mining and Computation in

Statistics, C. R. Rao, E. J. Wegman and J. L. Solka, eds., Elsevier, North-Holland, 2005, pp. 263-302.

[15] K. V. Mardia, J. T. Kent and J. M. Bibby, Multivariate Analysis, 7th Printing, Academic Press, London, 2000.

[16] R. A. Maronna, R. D. Martin and V. J. Yohai, Robust Statistics: Theory and Methods, John Wiley and Sons, New York, NY, USA, 2006.

[17] R. A. Maronna and R. Zamar, Robust estimation of location and dispersion for high-dimensional datasets, Technometrics 44 (2002), 307-317.

[18] G. Pison, S. Van Aelst and G. Willems, Small sample corrections for LTS and MCD, Metrika 55(1-2) (2002), 111-123.

[19] M. Riani, A. C. Atkinson and A. Cerioli, Finding an unknown number of multivariate outliers, J. R. Stat. Soc. Ser. B Stat. Methodol. 71(2) (2009), 447-466.

[20] P. J. Rousseeuw, Multivariate estimation with high breakdown point, Paper Appeared in Mathematical Statistics and Applications B, W. Grossman, G. Pflug, I. Vincze and W. Wertz, eds., D. Reidel Publishing Company, 1985, pp. 283-297.

[21] P. J. Rousseeuw and A. M. Leroy, Robust Regression and Outlier Detection, John Wiley, New York, 1987.

[22] P. J. Rousseeuw and K. van Driessen, A fast algorithm for the minimum covariance determinant estimator, Technometrics 41 (1999), 212-223.

[23] P. J. Rousseeuw and B. C. van Zomeren, Unmasking multivariate outliers and leverage points, J. Amer. Statist. Assoc. 85(411) (1990), 633-639.

[24] R. J. Serfling, Approximation Theorems of Mathematical Statistics, John Wiley, New York, 1980.

[25] S. S. S. Yahaya, H. Ali and Z. Omar, An alternative Hotelling T^2 control chart based on minimum vector variance (MVV), Modern Applied Science 5(4) (2011), 132-151.

[26] D. L. Woodruff and D. M. Rocke, Computable robust estimation of multivariate location and shape in high dimension using compound estimators, J. Amer. Statist. Assoc. 89 (1994), 888-896.