Far East Journal of Mathematical Sciences (FJMS)

Published Online: November 2015
http://dx.doi.org/10.17654/FJMSDec2015_819_830
Volume 98, Number 7, 2015, Pages 819-830 **ISSN: 0972-0871**

ENHANCING MINIMUM VECTOR VARIANCE ESTIMATORS USING REWEIGHTED SCHEME

Hazlina Ali, Sharipah Soaad Syed Yahaya and Zurni Omar

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
zurni@uum.edu.my

Abstract

Minimum vector variance (MVV) is one of the latest contributions in the study of multivariate robust estimators. MVV estimators possess three important properties of a good robust estimator, namely, high breakdown point, affine equivariance and computational efficiency. However, highly robust affine equivariant estimators with the best breakdown point commonly have to compensate with low statistical efficiency. In order to cater this drawback, a reweighted minimum vector variance (RMVV) which is capable of increasing the efficiency while retaining the highest breakdown point is proposed in this paper. A simulation study was conducted to investigate the asymptotic

Received: April 2, 2015; Revised: June 17, 2015; Accepted: July 2, 2015

2010 Mathematics Subject Classification: 65C60.

Keywords and phrases: robust estimation, minimum vector variance, reweighted minimum vector variance, outliers' detection.

The research is financed by Ministry of Education Malaysia under Fundamental Research Grant Projects (FRGS) No. 12801.

Communicated by K. K. Azad

relative efficiency and finite-sample behavior of the estimators for several types of distributions. The numerical results revealed that the reweighed scheme is able to attain higher efficiency compared to MVV estimators.

1. Introduction

Several robust estimators of multivariate location and scatter have been proposed since Maronna's pioneering paper on multivariate *M*-estimation (Maronna [1]). One of the latest contributions in the study of multivariate robust estimators is the minimum vector variance (MVV) proposed by Herwindiati [2]. She had proven that MVV estimators possess three major properties of a good robust estimators, i.e., high breakdown point (BP = 0.5), affine equivariance and computational efficiency. Apart from being on par with the popular minimum covariance determinant (MCD) for its robustness, this estimator has the edge over MCD in terms of computational efficiency (Herwindiati et al. [3] and Djauhari et al. [4]). In addition, MVV estimators are more effective in detecting outliers and in controlling Type I error compared with MCD (Ali and Yahaya [5]). Nevertheless, some drawbacks such as inconsistency under normal distribution and biased for small sample size were discovered by Ali et al. [6]. They improved the MVV estimators in terms of consistency under normal distribution and unbiasedness by multiplying the MVV scatter estimator with the consistency and correction factors, respectively. The improved MVV estimators were then applied in the Hotelling T^2 control chart and the numerical result showed that there was a great improvement in the control limit values while maintaining its good performance in terms of false alarm and probability of detection.

However, highly robust affine equivariant estimators with the best breakdown point commonly suffer from low statistical efficiency. Therefore, to make MVV estimators more practical in statistical inference, MVV estimators should offer a reasonable efficiency under normal model and manageable asymptotic distribution. Thus, Ali et al. [7] had investigated the efficiency of MVV estimators based on the asymptotic relative efficiency (ARE). The result indicated that there was a conflict in the statistical efficiency of MVV estimators for different breakdown points. To overcome this issue in MVV, in this paper, we propose to adopt the reweighted version as suggested by Rousseeuw and van Zomeren [8].

The outline of this paper is as follows: Section 2 proposes reweighted version of MVV estimators followed by the investigation on their asymptotic efficiency in Section 3. Section 4 presents the numerical results of the investigation on the finite-sample behavior of the estimator using simulation technique and finally, the conclusion is given in the last section.

2. Improvement of Minimum Vector Variance

As discussed in Ali et al. [7], the distribution of robust MSD based on MVV and MCD estimators is asymptotically equal to χ_p^2 distribution. They defined the reweighted minimum vector variance (RMVV) estimators of location and scatter as follows:

$$\boldsymbol{m}_{\boldsymbol{RMVV}}^{\boldsymbol{raw}} = \frac{\sum_{i=1}^{n} w_i \boldsymbol{x_i}}{m}, \tag{1}$$

$$\boldsymbol{S}_{\boldsymbol{RMVV}}^{\boldsymbol{raw}} = \frac{\sum_{i=1}^{n} w_i (\boldsymbol{x_i} - \boldsymbol{m}_{\boldsymbol{RMVV}}^{\boldsymbol{raw}})(\boldsymbol{x_i} - \boldsymbol{m}_{\boldsymbol{RMVV}}^{\boldsymbol{raw}})^t}{m}, \tag{2}$$

where m represents the number of observations with $d_{(MVV)i}^2 \leq \chi_{p,\,0.025}^2$ and

$$w_i = \begin{cases} 0, & \text{if } d_{i_{MVV}}^2(\boldsymbol{x_i}, \boldsymbol{m_{MVV}}) > \chi_{p,\,0.025}^2, \\ 1, & \text{otherwise.} \end{cases}$$

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 reweighted estimator in order to improve its biasness for small sample behavior. The consistency factor $c^*(m)$ is defined as

$$c^*(m) = \frac{m/n}{P(\chi_{p+2}^2 < \chi_{p,\,m/n}^2)}. \tag{3}$$

However, the consistency factor is still not sufficient to make the RMVV estimator unbiased for small sample sizes. To overcome this drawback, the correction factor, $\vartheta^{*\alpha}_{m,n,p}$ was computed for several sample sizes n and dimension p via simulation approach. Then data sets $X^{(j)} \in \mathbb{R}^{n \times p}$ from standard normal distribution $N_p(0, I)$ were generated. For each data set $X^{(j)}$, $j = 1, \ldots, r$, we then calculated the RMVV estimators of location and scatter as given in equations (1) and (2) followed by $\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}(\boldsymbol{j})}_{\boldsymbol{RMVV}}$. If the estimator is unbiased, then $E[\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}] = \boldsymbol{I_p}$. Thus, the pth root of the determinant of $\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}$ is expected to be equal to 1. The mean of the pth root of the determinant is given by

$$r(|\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}|) = \frac{1}{r}\sum_{j=1}^{r}(|\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}(\boldsymbol{j})}_{\boldsymbol{RMVV}}|)^{1/p},$$

where $|\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}|$ denotes the determinant of a square matrix $\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}$. We then performed $r = 1000$ simulations for different sample sizes n and dimensions p, with value of $\alpha = 0.05$. The correction factor for $\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}$ is given as:

$$\vartheta^{*\alpha}_{m,n,p} := \frac{1}{r(|\boldsymbol{c}^*(\boldsymbol{m})\boldsymbol{S}^{\boldsymbol{raw}}_{\boldsymbol{RMVV}}|)}. \tag{4}$$

Next, we calculate the RMVV location and scatter as follows:

$$\boldsymbol{m}_{\boldsymbol{RMVV}} = \frac{\sum_{i=1}^{n} w_i \boldsymbol{x_i}}{m}, \tag{5}$$

$$\boldsymbol{S}_{\boldsymbol{RMVV}} = \vartheta^{*\alpha}_{m,n,p} c^*(m) \frac{\sum_{i=1}^{n} w_i(\boldsymbol{x_i} - \boldsymbol{m}_{\boldsymbol{RMVV}})(\boldsymbol{x_i} - \boldsymbol{m}_{\boldsymbol{RMVV}})^t}{m}. \tag{6}$$

Finally, the square robust Mahalanobis distances were obtained by using the following formula:

$$d^2_{i_{RMVV}}(\boldsymbol{x_i}, \boldsymbol{m_{RMVV}}, \boldsymbol{S_{RMVV}}) = (\boldsymbol{x_i} - \boldsymbol{m_{RMVV}})^t \boldsymbol{S}^{-1}_{\boldsymbol{RMVV}}(\boldsymbol{x_i} - \boldsymbol{m_{RMVV}}). \tag{7}$$

The approximation of RMVV solution was based on the generalization of MVV algorithm as described in Ali and Yahaya [5] and Yahaya et al. [9]. The complete algorithm of RMVV had been discussed in Ali et al. [7]. Two typical choices of breakdown point (BP) were considered, i.e., 0.5 with $h = [(n + p + 1)/2]$ and 0.25 with $h = (0.75)n.$

3. Efficiency of Reweighted Minimum Vector Variance

To gain more insight into the RMVV estimators and observe how reweighting affects their performance, we computed the asymptotic relative efficiency. The computation of asymptotic relative efficiency (ARE) was based on the definition given by Serfling [10]. For any parameter $\boldsymbol{\theta} \in \mathbb{R}^p$, and two estimators $\hat{\boldsymbol{\theta}}^{(j)}$ which are p-variate normal with mean $\boldsymbol{\theta}$ and non-singular covariance matrices $\Sigma_j(F)/n$, where $j = 1, 2$ and F is the corresponding distribution, the ARE of $\hat{\boldsymbol{\theta}}^{(\mathbf{2})}$ to $\hat{\boldsymbol{\theta}}^{(\mathbf{1})}$ is

$$\text{ARE}(\hat{\boldsymbol{\theta}}^{(\mathbf{2})}, \hat{\boldsymbol{\theta}}^{(\mathbf{1})}, F) = \left(\frac{|\Sigma_1(F)|}{|\Sigma_2(F)|}\right)^{1/p}. \tag{8}$$

The higher value in (8) indicates that the estimator is more efficient. Table 1 shows the asymptotic relative efficiency (ARE) of the RMVV scatter estimators with different breakdown points of 0.25 and 0.5 with relative to the MVV estimator with breakdown point of 0.5 ($\text{MVV}_{0.5}$) at normal model. They are denoted as $\text{RMVV}_{0.25}$ and $\text{RMVV}_{0.5}$, respectively, and computed using the following equation:

$$\text{ARE}(S_{RMVV}, S_{MVV0.5}) = (S_{MVV0.5}/S_{RMVV})^{1/p}. \tag{9}$$

Note that the ARE for $\text{RMVV}_{0.25}$ is lower than $\text{RMVV}_{0.5}$ which indicates that the latter is more efficient than the former for all p's. Furthermore, the values for $\text{RMVV}_{0.5}$ are consistent across all p's, but not in the case of $\text{RMVV}_{0.25}$.

We could observe that the ARE values for $RMVV_{0.25}$ ascend with the increase in p's, but the values are below 0.85. From Table 1, we can deduce that by reweighting MVV, we can achieve high efficiency while simultaneously maintaining highest breakdown point. To show the effect of BP on efficiency, let us refer to Table 2. The first row records the efficiency of MVV estimator while the second row records the efficiency of the reweighted version of the estimator, RMVV. Apparently, the MVV estimator is more efficient when $BP = 0.25$, however, after reweighting the estimator, the efficiency at $BP = 0.5$ improves and outdoes $BP = 0.25$. The result suggests that the $RMVV_{0.5}$ estimators possess both high efficiency and high breakdown point, hence, making these estimators more appealing. Nevertheless, we should be aware that the gains in efficiency come at the price of a larger bias under contamination. The reason is that higher efficiency can only be obtained by increasing tuning parameters, which in turn affects the bias under contamination (Rousseeuw [11]). Our study then continued with the investigation on finite-sample robustness of RMVV estimators to support the above ARE results. For that purpose, a simulation study was conducted and discussed in the following section.

Table 1. Asymptotic relative efficiency of the scatter matrix for the RMVV estimator with different breakdown points ($BP = 0.25$ and 0.5) w.r.t the MVV estimator with ($BP = 0.5$) at normal model

	$p = 2$	$p = 5$	$p = 10$	$p = 15$	$p = 20$
$RMVV_{0.25}$ w.r.t $MVV_{0.5}$	0.6984	0.7490	0.8012	0.8293	0.8489
$RMVV_{0.5}$ w.r.t $MVV_{0.5}$	1.0217	0.9984	1.0015	0.9975	0.9987

Table 2. Asymptotic relative efficiency of the scatter matrix for the MVV and RMVV estimators with $BP = 0.25$ with relative to MVV and RMVV estimators with $BP = 0.5$, respectively

	$p = 2$	$p = 5$	$p = 10$	$p = 15$	$p = 20$
$MVV_{0.25}$ w.r.t $MVV_{0.5}$	1.4073	1.3225	1.2439	1.2024	1.1760
$RMVV_{0.25}$ w.r.t $RMVV_{0.5}$	0.9620	0.9922	0.9951	0.9997	0.9996

4. Finite-sample Robustness

To study on the finite-sample robustness of the RMVV location estimator, we performed simulations on contaminated data sets based on shifted outliers. We consider this type of contamination due to the fact that, for distance based method, the shift outliers are the most difficult to detect (Rocke and Woodruff [12]). In each simulation, we generated $L = 1000$ data sets of $N(0, I_p)$ with $p = 2,\ 5$ and 10 representing small, medium and slightly high numbers of dimensions with reasonable values of sample sizes of $n = 50,\ 100$, 200 and 500. We have considered a contaminated model by using a mixture of normal

$$(1-\varepsilon)N_p(0,\ I_p) + \varepsilon N_p(\mu_1,\ I_p),$$

where ε is the proportion of outliers. We consider ε to be 0.1 or 0.2, while μ_1 is the shift in the mean with value of 3 or 5. Manipulation on the mean shifts and percentage of outliers generate 4 different types of contaminated distributions which are categorized as mildly, moderately and extremely contaminated as follows:

(1) $(0.9)N_p(0,\ I_p) + (0.1)N_p(3,\ I_p)$ - mild contamination,

(2) $(0.8)N_p(0,\ I_p) + (0.2)N_p(3,\ I_p)$ - moderate contamination,

(3) $(0.9)N_p(0,\ I_p) + (0.1)N_p(5,\ I_p)$ - moderate contamination,

(4) $(0.8)N_p(0,\ I_p) + (0.2)N_p(5,\ I_p)$ - extreme contamination.

To measure robustness, we used the mean squared error (MSE) and bias (Rousseeuw et al. [13]). Then MSE and bias of the mean (location) vectors are computed for each simulation, as in Roelant et al. [14],

$$MSE(\hat{\mu}_{RMVV}) = n\left[\frac{\sum_{j=1}^{p}\sum_{l=1}^{L}\{(\hat{\mu}_{RMVV})_j^{(l)}\}^2}{pL}\right], \tag{10}$$

$$bias(\hat{\mu}_{RMVV}) = \left[\frac{1}{p} \sum_{i=1}^{p} \left\{ \frac{\sum_{l=1}^{L} (\hat{\mu}_{RMVV})_j^{(l)}}{L} \right\}^2 \right]^{1/2}, \tag{11}$$

where $l = 1, \ldots, L$ and $j = 1, \ldots, p$.

Tables 3-6 show the MSE and bias from mild, moderate and extreme contaminations for $\mathrm{RMVV}_{0.5}$, $\mathrm{RMVV}_{0.25}$ and $\mathrm{MVV}_{0.5}$ when $p = 2$, 5 and 10, respectively. In general, across the types of contaminations, there is a diminution in the value of MSE when *p* increases. For most conditions, the $\mathrm{RMVV}_{0.25}$ location estimator yields the lowest value of MSE, followed by $\mathrm{RMVV}_{0.5}$ and then $\mathrm{MVV}_{0.5}$. For larger sample sizes, the bias values for all estimators reduce closer to zero.

As shown in Table 3, under mild contamination, $\mathrm{RMVV}_{0.25}$ produces the smallest bias value when $p = 2$ and $n = 50$, but when *n* increases, $\mathrm{RMVV}_{0.5}$ estimator outperforms $\mathrm{RMVV}_{0.25}$. Nonetheless, when *p* increases to 5 and 10, $\mathrm{RMVV}_{0.25}$ reverts back to be the better performer. This situation repeats for moderate contamination with mean shift 5 (see Table 4). Except for $n = 50$, all the other combinations of *p* and *n* for $\mathrm{RMVV}_{0.25}$ generate the smallest bias values. For the other moderate contamination in Table 5 where there are 20% outliers with mean shift 3, $\mathrm{RMVV}_{0.5}$ is more dominant in generating the smallest bias value when $p = 2$. However, the situation changes when *p* increases to 5, whereby $\mathrm{RMVV}_{0.25}$ seems to outperform $\mathrm{RMVV}_{0.5}$. As *p* increases to 10, $\mathrm{RMVV}_{0.25}$ maintains to be better than $\mathrm{RMVV}_{0.5}$, but as *n* increases to 200 and above, $\mathrm{RMVV}_{0.5}$ becomes less bias. Under the condition of extreme contamination as presented in Table 6, $\mathrm{RMVV}_{0.25}$ outperforms $\mathrm{RMVV}_{0.5}$ when $p = 2$ and 5, but when *p* increases to 10, $\mathrm{RMVV}_{0.5}$ is better in terms of bias.

Table 3. Location estimator: 10% outliers with mean shift 3 (mild contamination)

n	50		100		200		500	
$p = 2$	MSE	bias	MSE	bias	MSE	bias	MSE	bias
$\boldsymbol{RMVV_{0.5}}$	3.0032	0.0114	3.1907	0.0032	3.4245	0.0034	3.6031	0.0017
$\boldsymbol{RMVV_{0.25}}$	1.7446	0.0066	1.6572	0.0067	1.6788	0.0052	1.6964	0.0029
$\boldsymbol{MVV_{0.5}}$	3.4706	0.0080	4.0558	0.0047	4.5698	0.0043	5.0813	0.0018
$p = 5$								
$\boldsymbol{RMVV_{0.5}}$	2.0571	0.0056	2.0776	0.0038	2.0825	0.0098	2.0380	0.0044
$\boldsymbol{RMVV_{0.25}}$	1.4971	0.0066	1.4342	0.0037	1.4038	0.0078	1.3780	0.0031
$\boldsymbol{MVV_{0.5}}$	2.1349	0.0127	2.4213	0.0092	2.7794	0.0114	3.1466	0.0048
$p = 10$								
$\boldsymbol{RMVV_{0.5}}$	1.8022	0.0126	1.8966	0.0151	1.8359	0.0136	1.7480	0.0089
$\boldsymbol{RMVV_{0.25}}$	1.4394	0.0092	1.3617	0.0120	1.3560	0.0117	1.3441	0.0079
$\boldsymbol{MVV_{0.5}}$	1.8042	0.0107	1.9474	0.0148	2.0003	0.0154	2.4890	0.0105

Table 4. Location estimator: 10% outliers with mean shift 5 (moderate contamination)

n	50		100		200		500	
$p = 2$	MSE	bias	MSE	bias	MSE	bias	MSE	bias
$\boldsymbol{RMVV_{0.5}}$	2.9739	0.0102	3.1905	0.0031	3.4192	0.0037	3.5983	0.0018
$\boldsymbol{RMVV_{0.25}}$	1.7294	0.0024	1.6315	0.0033	1.6520	0.0017	1.6837	0.0009
$\boldsymbol{MVV_{0.5}}$	3.4553	0.0072	4.0614	0.0044	4.5754	0.0043	5.0873	0.0019
$p = 5$								
$\boldsymbol{RMVV_{0.5}}$	2.0643	0.0057	2.0525	0.0043	2.0807	0.0100	2.0394	0.0045
$\boldsymbol{RMVV_{0.25}}$	1.4904	0.0065	1.4372	0.0042	1.4036	0.0077	1.3773	0.0031
$\boldsymbol{MVV_{0.5}}$	2.1370	0.0133	2.4064	0.0073	2.7612	0.0114	2.9932	0.0047
$p = 10$								
$\boldsymbol{RMVV_{0.5}}$	1.7936	0.0117	1.8934	0.0157	1.8708	0.0121	1.7473	0.0090
$\boldsymbol{RMVV_{0.25}}$	1.4463	0.0120	1.3674	0.0130	1.3545	0.0112	1.3464	0.0080
$\boldsymbol{MVV_{0.5}}$	1.7902	0.0107	1.9474	0.0148	2.1854	0.0129	2.4737	0.0117

Table 5. Location estimator: 20% outliers with mean shift 3 (moderate contamination)

n	50		100		200		500	
$p = 2$	MSE	bias	MSE	bias	MSE	bias	MSE	bias
$\boldsymbol{RMVV_{0.5}}$	2.7968	0.0109	2.9460	0.0034	3.1887	0.0043	3.1918	0.0037
$\boldsymbol{RMVV_{0.25}}$	1.6785	0.0168	1.6481	0.0188	1.7205	0.0177	1.7133	0.0145
$\boldsymbol{MVV_{0.5}}$	3.1571	0.0142	3.6666	0.0011	4.2234	0.0043	4.4614	0.0044
$p = 5$								
$\boldsymbol{RMVV_{0.5}}$	2.0241	0.0111	2.0430	0.0033	2.0979	0.0110	2.0288	0.0053
$\boldsymbol{RMVV_{0.25}}$	1.4396	0.0058	1.3793	0.0030	1.3890	0.0071	1.4244	0.0036
$\boldsymbol{MVV_{0.5}}$	2.0840	0.0109	2.3109	0.0060	2.6152	0.0143	2.8112	0.0147
$p = 10$								
$\boldsymbol{RMVV_{0.5}}$	1.7607	0.0128	1.9035	0.0154	1.9249	0.0103	1.8271	0.0093
$\boldsymbol{RMVV_{0.25}}$	1.4070	0.0114	1.3522	0.0132	1.3428	0.0111	1.4040	0.0095
$\boldsymbol{MVV_{0.5}}$	1.7682	0.0140	1.9238	0.0166	2.1292	0.0120	2.4336	0.0129

Table 6. Location estimator: 20% outliers with mean shift 5 (extreme contamination)

n	50		100		200		500	
$p = 2$	MSE	bias	MSE	bias	MSE	bias	MSE	bias
$\boldsymbol{RMVV_{0.5}}$	2.7651	0.0056	2.9251	0.0017	3.1455	0.0046	3.1678	0.0037
$\boldsymbol{RMVV_{0.25}}$	1.5730	0.0021	1.5113	0.0023	1.5761	0.0017	1.5600	0.0022
$\boldsymbol{MVV_{0.5}}$	3.1161	0.0110	3.6765	0.0019	3.8672	0.0049	4.4345	0.0038
$p = 5$								
$\boldsymbol{RMVV_{0.5}}$	2.0273	0.0129	2.0449	0.0026	2.0951	0.0108	2.1688	0.0037
$\boldsymbol{RMVV_{0.25}}$	1.4407	0.0060	1.3797	0.0032	1.3915	0.0071	1.4221	0.0037
$\boldsymbol{MVV_{0.5}}$	2.0922	0.0117	2.3082	0.0048	2.5998	0.0143	3.2375	0.0039
$p = 10$								
$\boldsymbol{RMVV_{0.5}}$	1.7522	0.0104	1.9073	0.0151	1.9257	0.0108	1.8272	0.0093
$\boldsymbol{RMVV_{0.25}}$	1.4154	0.0108	1.3397	0.0153	1.3439	0.0111	1.2739	0.0093
$\boldsymbol{MVV_{0.5}}$	1.7605	0.0116	1.9352	0.0149	2.2967	0.0137	2.4251	0.0132

5. Conclusion

The result of the investigation showed that the conflict between efficiency and breakdown point also occurred in MVV estimators. Hence, to maintain the highest breakdown value while simultaneously achieving high efficiency, this study developed a one-step reweighted version of minimum vector variance (MVV) estimator. The finding proved that reweighting leads to improvement in efficiency and at the same time maintaining the highest breakdown value as shown in the case of $RMVV_{0.5}$ estimators, thus making these estimators more appealing. Even though $RMVV_{0.25}$ has lower efficiency and breakdown point than $RMVV_{0.5}$, the estimators perform better in terms of MSE and bias especially under mild and extreme contaminations.

References

[1] R. A. Maronna, Robust *M*-estimators of multivariate location and scatter, Ann. Statist. 1 (1976), 51-67.

[2] 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.

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

[4] M. A. Djauhari, M. Mashuri and D. E. Herwindiati, Multivariate process variability monitoring, Comm. Statist. Theory Methods 37(11) (2008), 1742-1754.

[5] H. Ali and S. S. Syed Yahaya, On robust Mahalanobis distance issued from minimum vector variance, Far East J. Math. Sci. (FJMS) 74(2) (2013), 249-268.

[6] H. Ali, S. S. Syed Yahaya and Z. Omar, Robust Hotelling T^2 control chart with consistence minimum vector variance, Math. Probl. Eng. 2013 (2013), Article ID 401350, 8 pp. http://dx.doi.org/10.11555/2013/401350.

[7] H. Ali, S. S. Syed Yahaya and Z. Omar, The efficiency of reweighted minimum vector variance, Publish in Proceedings of the 3rd International Conference on Mathematical Sciences, AIP Conf. Proc. 1602, Vol. 1151, 2014. http://dx.doi.org/ 10.1063/1.4882629.

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

[9] S. S. Syed Yahaya, H. Ali and Z. Omar, An alternative Hotelling T^2 control chart based on minimum vector variance (MVV), Modern Appl. Sci. 5(4) (2011). DOI: 10.5539/mas.v5n4p132.

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

[11] P. J. Rousseeuw, Unconventional features of positive-breakdown estimators, Statist. Probab. Lett. 19 (1994), 417-431.

[12] D. Rocke and D. Woodruff, Identification of outliers in multivariate data, J. Amer. Statist. Assoc. 91 (1996), 1047-1061.

[13] P. J. Rousseeuw, K. Van Driessen, S. Van Aelst and J. Agullo, Robust multivariate regression, Technometrics 46(3) (2004), 293-305. DOI: 10.1198/004017004000000329

[14] E. Roelant, S. Van Aelst and G. Willems, The minimum weighted covariance determinant estimator, Metrika 70(2) (2009), 177-204. DOI: 10.1007/500184-008-0186-3.