INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

MARS MANAGEMENT ACADEMIC RESEARCH SOCIETY
EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2222-6990

Robust Weiszfeld's Structural Equation Modelling in Covariance Matrices

Hafeez Ahmed, Maznah Mat Kasim, Malina Zulkifli

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v8-i12/5322 DOI: 10.6007/IJARBSS/v8-i12/5322

Received: *28 Nov 2018,* ***Revised:*** *17 Dec 2018,* ***Accepted:*** *26 Dec 2018*

Published Online: 30 Dec 2018

In-Text Citation: (Ahmed, Kasim, & Zulkifli, 2018)
To Cite this Article: Ahmed, H., Kasim, M. M., & Zulkifli, M. (2018). Robust Weiszfeld's Structural Equation Modelling in Covariance Matrices. *International Journal of Academic Research in Business and Social Sciences*, *8*(12), 1773–1784.

Vol. 8, No. 12, 2018, Pg. 1773 - 1784

http://hrmars.com/index.php/pages/detail/IJARBSS | **JOURNAL HOMEPAGE**

INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
www.hrmars.com
ISSN: 2222-6990

Robust Weiszfeld's Structural Equation Modelling in Covariance Matrices

Hafeez Ahmed, Maznah Mat Kasim, Malina Zulkifli
School of Quantitative Science, Universiti Utara Malaysia

Abstract
The existing procedures for structural equation modelling involve in goodness of fit test for the sample covariance matrix by the structural model can no longer work in high dimensional datasets. The sample covariance can easily be influenced by outlier's presence in the datasets. This affect the estimation of the sample mean and sample covariance not being accurate and as well cause inefficiency in the computation. Therefore, there is need to suggest a robust covariance estimator the L_1-median with Weiszfeld algorithm that will resolve the outliers problem in high dimensional dataset. This test is subjected to conditions of sample size (n), variables (p) percentages of outliers (ε) with $\alpha = 0.05$. Simulation study carried out and the results show that when variable is minor both test performed but the new robust covariance test is better. At both middle and greater variables the existing test cannot compute the rate when $p > n$ cases. Generally, the results shows that the newly incorporated robust covariance test performed better compare to the existing test.

Introduction
The original methods for structural equation model include fitting the regular sample covariance matrix by a suggested maximum likelihood structural model. Since the assumption of normality required in the original estimation methods is frequently not satisfied in high dimensional datasets, Huber (1964). When number of observation, (p) is larger than the sample size, (n). The sample covariance matrix is easily influenced by a few outliers present, the usual exercise of modeling the sample covariance matrix can lead to inaccurate estimates as well as overestimated fit in high dimensional cases. Hence, causing the inability of the model to provide adequate fit or statistical explanation. However, when outliers exist in the data, the use of sample mean vector will result in poor estimation.

Thus, we need estimators which are robust to the existence of outliers to resolve the problem. In this research, in order to overcome these problem, several literatures are presented in statistics that highlight the importance of robustness of the sample covariance matrices. A wide range of robust estimators of multivariate location and scatter are available.

Some of them are based on the minimization of a robust scale of Mahalanobis distance such as M-estimator (Campbell, 1980), minimum volume ellipsoid (MVE), minimum covariance determinant (MCD) estimates {Rousseeuw & Driessen, 1999, Pison, Van Aelst & Willems, 2003), S-estimates (Davies, 1987), and τ-estimates (Bianco & Boente, 2002). Others are based on projections, for example, the Stahel-Donoho estimate (SDE), *P*-estimates (Maronna, Stahel, & Yohai, 1992) and Kurtosis1 (Pena & Prieto, 2001). But, the minimum volume ellipsoid (MVE) and minimum covariance determinant (MCD) estimator introduced by Rousseeuw (1984; 1985) has received a considerable attention by scientific community and widely used in practice.
For high dimensional data computing, the MVE takes too much time and find it challenging to resolve the outlier's problem. Therefore, we proposed a modification of ML-test by incorporating the Weiszfeld's algorithm covariance matrix (Vardi & Zhang, 2000 Cardot & Godichon-Baggioni, 2017), for developing a new robust covariance matrix in the presence of outliers. To resolve the outlier's problem in high dimensional data sets. The most useful model in data analysis and outliers in high dimension is possibly the modelling of data by high dimensional data set with the median Weiszfeld's algorithm covariance matrix they are provably immune to outliers. Given data with a large fraction of extreme outliers, a robust estimator guarantees the returned value is still within the non-outlier part of the data (Tang & Phillips, 2016 and Arrigoni, Rossi, Fragneto & Fusiello, 2018).

Methodology

Mathematically, the original ML-test is given as:

$$F_{ML}(\hat{\theta}) = log|\boldsymbol{\Sigma}(\hat{\theta})| + tr\left(S\theta^{-1}(\hat{\theta})\right) - log|S| - (p+q) \quad (1)$$

where $\boldsymbol{\Sigma}(\hat{\theta})$ is the variance structure, $\hat{\theta}$ estimated parameters, tr is the trace of a matrix, S sample variance matrix, θ^{-1} inverse of a matrix, p endogenous latent factors observation, and q the exogenous latent factors observation (Bollen, Kirby, Curran, Paxton & Chen, 2007).
The sample covariance matrix is most often assumed as the estimate of the population covariance matrix is most often assumed for this purposes is as:

$$S = 1/n \sum_{i=1}^{n} (\mathrm{x}^{(\mathrm{i})} - \bar{\mathrm{x}})(\mathrm{x}^{(\mathrm{i})} - \bar{\mathrm{x}})', \quad (2)$$

where $\bar{\mathrm{x}} = \sum \mathrm{x}^{(\mathrm{i})}/\mathrm{n}$ is the mean of the sample covariance matrix.
In order to develop robust ML-test with SEM denoted by $\mathrm{ML}_{\mathrm{SEM}}$, the covariance of L_1-Weiszfeld, $S_{L_1-\mathrm{Weiszfeld}}$ where $i = 1, 2, \ldots, m$ is replaced into Equation 1.
Therefore, the test now become the following,

$$F_{ML}(\hat{\theta}) = log|\boldsymbol{\Sigma}(\hat{\theta})| + tr\left(S_{L_1-\mathrm{Weiszfeld}}\theta^{-1}(\hat{\theta})\right) - log|S_{L_1-\mathrm{Weiszfeld}(i)}| - (p+q) \quad (3)$$

Where $S_{L_1-\mathrm{Weiszfeld}} = 1/N \sum_{i=1}^{m} S_{L_1-\mathrm{Weiszfeld}(i)}$ is the pooled sample covariance matrix of L_1-Weiszfeld estimator.
$L_1 - \mathrm{Weiszfeld}(i)$ is the number of subgroup where the stability of matrices is hypothesized.
$N = n_1 + n_2 + \cdots + n_{L_1-\mathrm{Weiszfeld}}$; $n_i = i$-th sample size. The measure of estimation that is performed in ML-test is the Type I error test rate on equation 1 and 3. The original test now become SEM_O-test and modified $\mathrm{SEM}_{Weiszfeld}$-test, we be compared in terms performance.

Results and Discussions

The main result is on robust covariance goodness of fit test in the presence of outliers. We compare original ML_{SEM}-test and modified $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$- test in terms of Type I error rate. For each of the test 3 types of data contaminations are used to examine the strength and weakness of the tests. Also, all these tests have been open to various conditions which are number of variables (p), sample size (n), percentage of outliers (ε) and Mean Shift (μ). The summary of test comparison are in form of table. Start with sample size (n), in the first column in each of the table, followed by percentage of outliers (ε) and Mean Shift (μ). The following two columns detailed the Type I error rate of ML_{SEM}-test and the modified $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$- test examined at different sample sizes in the study. The values that is closest to the significance level and within [0.025-0.075] are shaded in the tables with green and pink colours, whereas red colours denotes results cannot be computed.
In addition, Table 3.1 to 3.6 recorded the Type I error rate for each condition are arranged based on the ascending number of variables, specifically, minor, middle and greater variables ($p = 15$ and 18, $p = 20$ and 25, $p = 30$ and 50, respectively, $n = 10, 20, 30, 40, 50, 60$ and 70 with $\alpha = 0.05$.

Type I Error for Minor Number of Variables ($p = 15$ and 18)

In Table 3.1 the Type I error rate of $\mathrm{ML}_{\mathrm{SEM}}$ and $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$-tests are recorded. The overall results show that $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$-test is more robust compared to $\mathrm{ML}_{\mathrm{SEM}}$-test. All the values of Type I error rate of $\mathrm{ML}_{\mathrm{SEM}}$-test fall within the interval $[0.025,\ 0.075]$ in blue when $p = 15$ and $n = 10, 20, 30, 40, 60$ and 70, except for conditions when $\varepsilon = 0$ and $\mu = 0$, $n = 40$ and 70, at $\varepsilon = 10$ and $\mu = 3$ when $n = 30, 40$ and 70, at $\varepsilon = 10$ and $\mu = 5$ when $n = 30, 40$ and 70, at $\varepsilon = 20$ and $\mu = 3$ when $n = 30, 40, 60$ and 70 and $\varepsilon = 20$ and $\mu = 5$ when $n = 30, 40$ and 70. Similarly, for $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$-test shaded in green, when $\varepsilon = 0$ and 10 and $\mu = 0, 3$ and 5 only when $n = 60$ and 70, also, when $\varepsilon = 20$ and $\mu = 3$ and 5 only when $n = 40$, 60 and 70. From the results, it shows that $\mathrm{ML}_{\mathrm{SEM}}$-test is still well perform where 16 out of 30 conditions are non-robust, but $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$-test is more robust compare to $\mathrm{ML}_{\mathrm{SEM}}$-test and performed very well 24 out of 30 fall within the robust interval.
In Table 3.2. The rate contamination cannot be computed when $p > n$ cases. All the values of Type I error rate of $\mathrm{ML}_{\mathrm{SEM}}$-test fall within the interval $[0.025,\ 0.075]$ in blue when $p = 18$ and $n = 10, 20, 30, 40, 60$ and 70, except for conditions when $\varepsilon = 0$ and $\mu = 0$, $n = 40$ and 70, at $\varepsilon = 10$ and $\mu = 3$ when $n = 40, 60$ and 70, at $\varepsilon = 10$ and $\mu = 5$ when $n = 30, 40$ and 60, at $\varepsilon = 20$ and $\mu = 3$ when $n = 30, 40, 60$ and 70 and $\varepsilon = 20$ and $\mu = 5$ when $n = 20, 30\ 40, 60$ and 70. Equally, for $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$-test shaded in green, when $\varepsilon = 0$ and 10 and $\mu = 0, 3$ and 5 only when $n = 40$ and 70, also, when $\varepsilon = 20$ and $\mu = 3$ and 5 only when $n = 60$. From the results, it shows that $\mathrm{ML}_{\mathrm{SEM}}$-test is does not performed well, where 22 out of 30 conditions are non-robust. Besides, $\mathrm{ML}_{\mathrm{SEM}_{Weiszfeld}}$-test is more robust compare to $\mathrm{ML}_{\mathrm{SEM}}$-test and performed very well 23 out of 30 fall within the robust interval.

Table 3.1: Type I error rates for the corresponding $p = 15$

Sample Size (n)	% Outlier (ε)	Mean Shift (μ)	Robust Tests	
			$\mathrm{ML_{SEM}}$	$\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$
10	0	0	0.0502	0.05
20			0.0542	0.0404
30			0.0738	0.0454
40			0.113	0.0578
60			0.0578	0.0404
70			0.0982	0.0494
10	10	3	0.0502	0.0502
20			0.058	0.0336
30			0.1654	0.0276
40			0.4024	0.0396
60			0.0726	0.023
70			0.3354	0.0232
10	10	5	0.05	0.0502
20			0.0608	0.0472
30			0.1466	0.0532
40			0.3248	0.0636
60			0.0696	0.0492
70			0.276	0.0636
10	20	3	0.05	0.05
20			0.0644	0.0476
30			0.2134	0.0648
40			0.4932	0.0868
60			0.085	0.0464
70			0.4346	0.082
10	20	5	0.05	0.05
20			0.0664	0.0476
30			0.1948	0.0648
40			0.4116	0.0868
60			0.0766	0.0464
70			0.353	0.082
			14	24

Shaded region indicate Type I error within [0.025- 0.075]

Table 3.2: Type I error rates for the corresponding $p = 18$

Sample Size (n)	% Outlier (ε)	Mean Shift (μ)	Robust Tests $\mathrm{ML_{SEM}}$	$\mathrm{ML_{SEM_{Weiszfeld}}}$
10	0	0		0.0502
20			0.0502	0.0498
30			0.0578	0.0658
40			0.1286	0.0828
60			0.263	0.0474
70			0.0648	0.0838
10	10	3		0.05
20			0.0502	0.0434
30			0.0702	0.061
40			0.254	0.0838
60			0.5556	0.0406
70			0.0954	0.0828
10	10	5		0.05
20			0.0502	0.0448
30			0.0642	0.0604
40			0.1596	0.0826
60			0.4008	0.0376
70			0.0694	0.0768
10	20	3		0.05
20			0.05	0.0256
30			0.0844	0.0322
40			0.5322	0.0528
60			0.8918	0.0122
70			0.1616	0.0358
10	20	5	0.05	0.0502
20			0.0666	0.036
30			0.294	0.0488
40			0.6166	0.0682
60			0.0984	0.0252
70			0.5894	0.0612
			8	23

Shaded region indicate Type I error within [0.025- 0.075]

Type I Error for Middle Number of Variables ($p = 20$ and 25)

In Table 3.3 the Type I error rate of $\mathrm{ML_{SEM}}$ and $\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$-tests are recorded. The overall results show that $\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$-test is more robust compared to $\mathrm{ML_{SEM}}$-test. The rate contamination cannot be computed when $p > n$ cases. All the values of Type I error rate of $\mathrm{ML_{SEM}}$-test fall within the interval $[0.025, 0.075]$ in blue when $p = 20$ and $n = 10, 20, 30, 40, 60$ and 70, except for conditions when $\varepsilon = 0$ and $\mu = 0$, $n = 40$ and 60, at $\varepsilon = 10$ and $\mu = 3$ when $n = 30, 60$ and 70, at $\varepsilon = 10$ and $\mu = 5$ when $n = 30, 40$ and 70, at $\varepsilon = 20$ and $\mu = 3$ when $n = 40, 60$ and 70 and $\varepsilon = 20$ and $\mu = 5$ when $n = 30, 40{,}60$ and 70. Likewise, for $\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$-test shaded in green, when $\varepsilon = 0$ and 10 and $\mu = 0, 3$ and 5 only when $n = 40$ and 70, also, when $\varepsilon = 20$ and $\mu = 3$ and 5 only when $n = 40$ and 70. From the results, it shows that $\mathrm{ML_{SEM}}$-test is not well perform where 23 out of 30 conditions are non-robust. Also, $\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$-test is more robust compare to $\mathrm{ML_{SEM}}$-test and performed very well 22 out of 30 fall within the robust interval. In Table 3.4. The Type I error rate of $\mathrm{ML_{SEM}}$-test fall within the interval in blue when $p = 25$, except for conditions when $\varepsilon = 0$ and $\mu = 0$, $n = 60$ and 70, at $\varepsilon = 10$ and $\mu = 3$ and 5 when $n = 40, 60$ and 70, at $\varepsilon = 20$ and $\mu = 3$ and 5, none of the n. Also, for $\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$-test shaded in green colours, when $\varepsilon = 0$ and 10 and $\mu = 0, 3$ and 5 only when $n = 30, 40$ and 70, also, when $\varepsilon = 20$ and $\mu = 3$ and 5 only when $n = 40$ and 70. From the results, it shows that $\mathrm{ML_{SEM}}$-test is not well performed where 4 out of 30 conditions are non-robust. Also, $\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$-test is more robust compare to $\mathrm{ML_{SEM}}$-test and perform very well 24 out of 30 fall within the robust interval.

Table 3.3: Type I error rates for the corresponding $p = 20$

Sample Size (n)	% Outlier (ε)	Mean Shift (μ)	Robust Tests $\mathrm{ML_{SEM}}$	$\mathrm{ML_{SEM_{\mathit{Weiszfeld}}}}$
10	0	0		0.0502
20				0.0504
30			0.0548	0.0712
40			0.1438	0.0948
60			0.2994	0.052
70			0.0636	0.0976
10	10	3		0.0502
20				0.0466
30			0.2994	0.0532
40			0.0636	0.0734
60			0.2544	0.0492
70			0.4966	0.0662
10	10	5		0.0502
20				0.0496
30			0.0622	0.068
40			0.289	0.0928
60			0.6072	0.0534
70			0.0862	0.09
10	20	3		0.05
20				0.0364
30			0.0502	0.0624
40			0.1022	0.082
60			0.6274	0.0308
70			0.9306	0.0888
10	20	5	0.05	0.0502
20			0.0756	0.04
30			0.4948	0.0592
40			0.907	0.078
60			0.1148	0.0328
70			0.8584	0.0782
			7	22

Shaded region indicate Type I error within [0.025- 0.075]

Vol. 8 , No. 12, Dec, 2018, E-ISSN: 2222-6990 © 2018 HRMARS

Table 3.4: Type I error rates for the corresponding $p = 25$

Sample Size (n)	% Outlier (ε)	Mean Shift (μ)	Robust Tests $\mathrm{ML_{SEM}}$	$\mathrm{ML_{SEM_{Weiszfeld}}}$
10	0	0		0.05
20				0.0564
30			0.0502	0.0848
40			0.0716	0.1002
60			0.2668	0.0506
70			0.53	0.1174
10	10	3		0.05
20				0.035
30			0.0684	0.046
40			0.3278	0.0712
60			0.6438	0.0376
70			0.1004	0.0604
10	10	5		0.05
20				0.035
30			0.0684	0.046
40			0.3278	0.0712
60			0.6438	0.0376
70			0.1004	0.0604
10	20	3		0.0502
20				0.0522
30			0.2026	0.0722
40			0.4796	0.1018
60			0.0792	0.054
70			0.417	0.0962
10	20	5		0.0502
20				0.0294
30			0.661	0.0386
40			0.9398	0.06
60			0.22	0.0158
70			0.9658	0.0616
			4	24

Shaded region indicate Type I error within [0.025- 0.075]

Type I Error for Greater Number of Variables ($p = 30$ and 50)

In Table 3.5 the Type I error rate of $\mathrm{ML_{SEM}}$ and $\mathrm{ML_{SEM_{Weiszfeld}}}$-tests are recorded. The overall results show that $\mathrm{ML_{SEM_{Weiszfeld}}}$-test is more robust compared to $\mathrm{ML_{SEM}}$-test. The rate contamination cannot be computed when $p > n$ cases. All the values of Type I error rate of $\mathrm{ML_{SEM}}$-test fall within the interval $[0.025,\ 0.075]$ in blue when $p = 30$ and $n = 10, 20, 30, 40, 60$ and 70, except for conditions when $\varepsilon = 0$ and $\mu = 0$, $n = 40$ and 60, at $\varepsilon = 10$ and $\mu = 3$ when $n = 40, 60$ and 70, at $\varepsilon = 10$ and $\mu = 5$ when $n = 40, 60$ and 70, at $\varepsilon = 20$ and $\mu = 3$ when $n = 40, 60$ and 70 and $\varepsilon = 20$ and $\mu = 5$ when $n = 40, 60$ and 70. Likewise, for $\mathrm{ML_{SEM_{Weiszfeld}}}$-test shaded in green, when $\varepsilon = 0$ and $\mu = 0$, when $n = 40$ and 70, also, when $\varepsilon = 10$ and $\mu = 3$ and 5 only when $n = 30, 40$ and 70, when $\varepsilon = 20$ and $\mu = 3$ and 5, when $n = 30, 40$ and 70, From the results, it shows that $\mathrm{ML_{SEM}}$-test performed worst where 29 out of 30 conditions are non-robust. Also, $\mathrm{ML_{SEM_{Weiszfeld}}}$-test is more robust compare to $\mathrm{ML_{SEM}}$-test and performed very well 19 out of 30 fall within the robust interval.

In Table 3.6. The Type I error rate of $\mathrm{ML_{SEM}}$-test fall within the interval in blue colour when $p = 50$, all the conditions are not robust. Similarly, for $\mathrm{ML_{SEM}}_{Weiszfeld}$-test shaded in green, when $\varepsilon = 0$ and $\mu = 0$, when $n = 30, 40$ and 70, also, when $\varepsilon = 10$ and $\mu = 3$ and 5 only when $n = 40$ and 70, when $\varepsilon = 20$ and $\mu = 3$ and 5 only when $n = 30, 40$ and 70. From the results, it shows that $\mathrm{ML_{SEM}}$-test performed worst 30 out of 30 conditions are non-robust. Also, $\mathrm{ML_{SEM}}_{Weiszfeld}$-test is more robust compare to $\mathrm{ML_{SEM}}$-test and performed very well 19 out of 30 fall within the robust interval.

Table 3.5: Type I error rates for the corresponding $p = 30$

Sample Size (n)	% Outlier (ε)	Mean Shift (μ)	Robust Tests $\mathrm{ML_{SEM}}$	$\mathrm{ML_{SEM}}_{Weiszfeld}$
10	0	0		0.05
20				0.046
30				0.072
40			0.0646	0.0912
60			0.3102	0.0468
70			0.5728	0.1046
10	10	3		0.05
20				0.0452
30				0.0786
40			0.5256	0.091
60			0.0896	0.045
70			0.5256	0.1204
10	10	5		0.05
20				0.033
30				0.0646
40			0.0896	0.075
60			0.5256	0.0278
70			0.8006	0.0964
10	20	3		0.0502
20				0.0444
30				0.0808
40			0.6796	0.0922
60			0.905	0.0392
70			0.2632	0.1138
10	20	5		0.05
20				0.0276
30				0.055
40			0.978	0.0706
60			0.999	0.0164
70			0.7322	0.0844
			1	19

Shaded region indicate Type I error within [0.025- 0.075]

Table 3.6: Type I error rates for the corresponding $p = 50$

Sample Size (n)	% Outlier (ε)	Mean Shift (μ)	Robust Tests ML_{SEM}	$\text{ML}_{\text{SEM}_{Weiszfeld}}$
10	0	0		0.05
20				0.0424
30				0.0804
40				0.0922
60			0.7068	0.0398
70			0.9218	0.126
10	10	3		0.0502
20				0.0304
30				0.0596
40				0,0728
60			0.259	0.0254
70			0.9818	0.1066
10	10	5		0.0502
20				0.0494
30				0.0614
40				0.0868
60			0.5792	0.054
70			0.9238	0.0802
10	20	3		0.05
20				0.0406
30				0.077
40				0.0868
60			0.8706	0.0364
70			0.9824	0.124
10	20	5		0.05
20				0.0396
30				0.0724
40				0.0852
60			0.8706	0.0338
70			0.9824	0.12
			0	19

Shaded region indicate Type I error within [0.025- 0.075]

In conclusion, there are 60 conditions involved in evaluating the robustness of test for minor number of variables ($p = 15$ and 18). There are 22 out 60 conditions of ML_{SEM}-test, and 47 condition of $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test fall within the robust interval. When, $p = 15$, there are 14 out 30 and $p = 18$, also 8 out 30 of ML_{SEM}-test, and 24 out of 30 and 23 out of 30 $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test respectively that fall within the interval. Likewise, when ($p = 20$ and 25). There are 11 out 60 conditions of ML_{SEM}-test, and 46 condition of $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test and for, $p = 20$, there are 7 out 30 and $p = 25$, there are 4 out 30 of ML_{SEM}-test, and , $p = 20$, there are 22 out 30 and $p = 25$, there are 24 out 30 of $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test fall within the robust interval. When ($p = 30$ and 50). There are 1 out 60 conditions of ML_{SEM}-test, and 38 condition of $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test and for, $p = 30$, there are 1 out 30 and $p = 50$, none of ML_{SEM}-test, and , $p = 30$, there are 19 out 30 and $p = 50$, there are 19 out 30 of $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test fall within the robust interval.

In general we derive the new statistical robust estimator using the Weiszfeld's algorithm covariance matrix to resolve the outliers' problem in high dimensional data sets. The robust estimator can computed the results based on when $p > n$ with $\text{ML}_{\text{SEM}_{Weiszfeld}}$-test in structural equation modelling.

References

Arrigoni, F., Rossi, B., Fragneto, P., & Fusiello, A. (2018). Robust synchronization in SO (3) and SE (3) via low-rank and sparse matrix decomposition. *Computer Vision and Image Understanding*.

Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. M., & Chen, F. (2007). Latent variable Models under misspecification: two-stage least squares (2SLS) and maximum likelihood (ML) estimators. *Sociological Methods & Research*, *36*(1), 48-86.

Bianco, A., & Boente, G. (2002). On the asymptotic behavior of one-step estimates in Heteroscedastic regression models. *Statistics & probability letters*, *60*(1), 33-47.

Campbell, N. A. (1980). Robust procedures in multivariate analysis I: Robust covariance Estimation. *Applied statistics*, 231-237.

Cardot, H. and Godichon-Baggioni, A. (2017). Fast Estimation of the Median Covariation Matrix with Application to Online Robust Principal Component7s Analysis. TEST, 26, 461-480.

Davies, P. L. (1987). Asymptotic Behaviour of $ S $-Estimates of Multivariate Location Parameters and Dispersion Matrices. *The Annals of Statistics*, *15*(3), 1269-1292.

Huber, P. J. (1964). Robust estimation of a location parameter. *The annals of mathematical Statistics*, *35*(1), 73-101.

Maronna, R. A., Stahel, W. A., & Yohai, V. J. (1992). Bias-robust estimators of multivariate Scatter based on projections. *Journal of Multivariate Analysis*, *42*(1), 141-161.

Rousseeuw, P. J., & Driessen, K. V. (1999). A fast algorithm for the minimum covariance Determinant estimator. *Technometrics*, *41*(3), 212-223.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American statistical Association, 79(388), 871-880.*

Rousseeuw, P. J. (1985). Multivariate estimation with high breakdown point. *Mathematical Statistics and applications*, *8*(283-297), 37.

Tang, P., & Phillips, J. M. (2016). The robustness of estimator composition. In *Advances in Neural Information Processing Systems* (pp. 929-937).

Pison, G., Van Aelst, S., & Willems, G. (2003). Small sample corrections for LTS and MCD. In *Developments in Robust Statistics* (pp. 330-343). Physica, Heidelberg.

Peña, D., & Prieto, F. J. (2001). Multivariate outlier detection and robust covariance matrix Estimation. *Technometrics*, *43*(3), 286-310.

Vardi, Y. and Zhang, C.-H. (2000). The multivariate L1-median and associated data depth. Proc. Natl. Acad. Sci. USA, 97(4):1423-1426.