INSIST Vol. 2 No. 1, April 2017 (1 - 5)
http://insist.unila.ac.id/
DOI: 10.23960/ins.v2i1.23

eISSN: 2502-8588

Received: 20/08/2016
Accepted: 19/01/2017
Published online: 01/04/2017

Robust Estimation of Generalized Estimating Equation when Data Contain Outliers

Khoirin Nisa[1,*], Netti Herawati[2]

***Abstract*—**In this paper, a robust procedure for estimating parameters of regression model when generalized estimating equation (GEE) applied to longitudinal data that contains outliers is proposed. The method is called 'iteratively reweighted least trimmed square' (IRLTS) which is a combination of the iteratively reweighted least square (IRLS) and least trimmed square (LTS) methods. To assess the proposed method a simulation study was conducted and the result shows that the method is robust against outliers.

***Keywords*—**GEE, IRLS, LTS, longitudinal data, regression model.

I. INTRODUCTION

LONGITUDINAL studies are increasingly common in many scientific research areas, for example in the social, biomedical, and economical fields. In longitudinal studies, multiple measurements are taken on the same subject at different points in time. Thus, observations for the same subject are correlated. The analysis of data resulting from such studies often becomes complicated due to the within-subject correlation. This correlation must be considered for any appropriate analysis method.

Generalized linear models (GLM) as described by McCullagh and Nelder [1] is a standard method used to fit regression models for univariate data that are presumed to follow an exponential family distribution. The association between the response variable and the covariates is given by the link function. GLM assume that the observations are independent and do not consider any correlation between the outcome of the *n* observations. Liang and Zeger [2] introduced an approach to this correlation problem using GEE to extend GLM into a regression setting with correlated observations within subjects.

The GEE method of Liang and Zeger gives consistent estimators of the regression parameter. The parameter estimates are consistent even when the variance structure is miss-specified under mild regularity conditions. However, problems can occurs when data contain outliers. The method is not robust against outliers since it is based on score equations from the quasi likelihood method of estimation. The working correlation matrix would be affected by the outliers and also the parameter estimates. In this situation, we need a robust method that can minimize the effect of outliers.

In recent years, a few authors have considered robust methods for longitudinal data analysis. For example, Qaqish and Preisser [3] proposed a resistant version of the GEE using M-type estimation by involving down-weighting influential data points. Gill [4] proposed a robustified likelihood based on multivariate normal distribution. Jung and Ying [5] proposed an adaptation of the Wilcoxon-Mann-Whitney method of estimating linear regression parameters for use in longitudinal data analysis under the working independence model. And recently, Abebe *et al*. [6] proposed a robust GEE using iterated reweighted rank-based estimation.

In this paper, we adopt the LTS [7] method for robust linear regression in the sense of trimming the data for estimating the regression coefficients so that the observations with high residuals are not included in the parameter estimation. In Section 2 we present a brief review of GEE. In Section 3 we describe our proposed method IRLTS. In Section 4 we discuss some results from our simulation study.

II. GENERALIZED ESTIMATING EQUATION AND IRLS METHOD

Let Y_{ij}, $j = 1, ..., m_i$, $i = 1, ..., n$ represent the *j*th measurement on the *i*th subject. There are m_i measurements on subject *i* and $N=\sum_{i=1}^{n} m_i$ total measurements. Assume that the marginal distribution of y_{ij} is of the exponential class of distributions and is given by:

$$f(y,\theta,\phi) = \exp\{y\theta - b(\theta)/a(\phi) + c(y,\phi)\}$$

where $a(.)$, $b(.)$, and $c(.)$ are given, θ is the canonical parameter, and ϕ is the dispersion parameter.

Let the vector of measurements on the *i*th subject be $\mathbf{Y}_i = [Y_{i1}, ..., Y_{imi}]^T$ with corresponding vector of means $\boldsymbol{\mu}_i = [\mu_{i1}, ..., \mu_{i\ mi}]^T$ and $\mathbf{X}_i = [X_{i1}, ..., X_{i\ mi}]^T$ be the m_ixp matrix of covariates. In general, the components of $\mathbf{Y}_i$ are correlated but $\mathbf{Y}_i$ and $\mathbf{Y}_k$ are independent for any $i \neq k$. To model the relation between the response and covariates, we can use a regression model similar to the generalized linear models:

$$g(\boldsymbol{\mu}_i) = \boldsymbol{\eta}_i = \mathbf{X}_i\ \boldsymbol{\beta}$$

where $\boldsymbol{\mu}_i = E(\mathbf{Y}_i|\mathbf{X}_i)$, g is a specified link function, and $\boldsymbol{\beta} = [\beta_1, ..., \beta_p]^T$ is a vector of unknown regression coefficients to be estimated. The GEE for estimating the $p\times 1$ vector of regression parameters $\boldsymbol{\beta}$ is is given by :

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \frac{\partial \boldsymbol{\mu}_i^T}{\partial \boldsymbol{\beta}} \mathbf{V}_i^{-1}[\mathbf{Y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta})] = 0 \qquad (1)$$

where $\mathbf{V}_i$ be the covariance matrix of $\mathbf{Y}_i$ modeled as $\boldsymbol{V}_i = \lambda \mathbf{A}_i^{1/2}\mathbf{R}(\boldsymbol{\alpha})\mathbf{A}_i^{1/2}$, $\mathbf{A}i$ is a diagonal matrix of variance function $V(\mu_{ij})$, and $\mathbf{R}(\alpha)$ is the working correlation matrix of $\mathbf{Y}_i$ indexed by a vector of parameters $\boldsymbol{\alpha}$. Solutions to (2) are obtained by alternating between estimation of λ, $\boldsymbol{\alpha}$ and θ.

There are several specific choices of the form of working correlation matrix $\mathbf{R}_i(\alpha)$ commonly used to model

[1]Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Lampung, Jalan Prof. Soemantri Brojonegoro No. 1, Bandar Lampung, Indonesia.

[*]Correspondence to Khoirin Nisa, email: khoirin.nisa@fmipa.unila.ac.id. Tel.:+62 721 701609; fax: +62 721 702767.

the correlation matrix of $\mathbf{Y}_i$. A few of the choices are shown below, one can refer to [1] for additional choices. The dimension of the vector $\boldsymbol{\alpha}$, which is treated as a nuisance parameter, and the form of the estimator of $\boldsymbol{\alpha}$ are different for each choice. Some typical choices are:

1. $\mathbf{R}_i(\boldsymbol{\alpha}) = \mathbf{R_0}$, a fixed correlation matrix. For $\mathbf{R}_0 = \mathbf{I}$, the identity matrix, the GEE reduces to the independence estimating equation.
2. Exchangeable: $Cor(Y_{ij}, Y_{ik}) = \alpha, \quad j \neq k$.
3. Autoregressive-1: $Cor(Y_{ij}, Y_{ik}) = \alpha^{|j-k|}$.
4. Unstructured: $Cor(Y_{ij}, Y_{ik}) = \alpha_{jk}$.

Solving for $\boldsymbol{\beta}$ is done with iteratively reweighted least squares (IRLS). The following is the algorithm for fitting the specified model using GEEs [3] :

1. Compute an initial estimate of $\hat{\boldsymbol{\beta}}_{GEE}$, for example with an ordinary generalized linear model assuming independence.
2. A current estimate $\hat{\boldsymbol{\beta}}_{GEE}$ is updated by regressing the working response vector
$$\mathbf{Z}^* = \mathbf{X}\hat{\boldsymbol{\beta}} + \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\beta}}(\mathbf{y} - \hat{\boldsymbol{\mu}})$$
on $\mathbf{X}$. A new estimate $\hat{\boldsymbol{\beta}}_{new}$ is obtained by :
$$\hat{\boldsymbol{\beta}}_{new} = (\mathbf{X^T W^* X})^{-1}\mathbf{X^T W^* Z^*} \quad (2)$$
where $\mathbf{W}^*$ is a block diagonal weight matrix whose ith block is the m_ixm_i matrix
$$\mathbf{W}_i^* = \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right)^{-1} \mathbf{A}_i^{-1}\mathbf{R}_i(\hat{\boldsymbol{\alpha}})\mathbf{A}_i^{-1}\left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right)^{-1} .$$
3. Use $\hat{\boldsymbol{\beta}}_{new}$ to update $\hat{\boldsymbol{\eta}} = \mathbf{X}\hat{\boldsymbol{\beta}}_{new} = \mathbf{HZ}^*$, where $\mathbf{H} = \mathbf{X}(\mathbf{X^T W^* X})^{-1}\mathbf{X^T W}^*$.
4. Iterate until convergence.

III. Iteratively Reweighted Least Trimmed Square Algorithm

First let us briefly recall that the robust estimation of regression parameters using LTS method is given by:
$$\hat{\boldsymbol{\beta}}_{LTS} = \arg\min \sum_{i=1}^{h} e_i^2 \quad (3)$$
which is based on the ordered absolute residuals $|e_1| \leq |e_2| \leq \ldots \leq |e_n|$. LTS estimation is calculated by minimizing the h ordered squares residuals, where h can be chosen within $\frac{n}{2} + 1 \leq h \leq \frac{3n}{4} + \frac{p+1}{4}$, with n and p being sample size and number of parameters respectively. When h = [n/2], , LTS locates that half of the observations which has the smallest estimated variance. In that case, the breakdown point is 50%. When h is set to the sample size, LTS and ordinary least square (OLS) coincide.

In [7] Rousseeuw and Leroy shows $n^{1/2}$ consistency and asymptotic normality of LTS in the location–scale model. Víšek [8] extends this to the regression model with random regressors, the proof for fixed regressors is in later series of his papers: [9][10].

When n is very small, it is possible to generate all subsets of size h to determine which one minimizes the LTS criterion. Rousseeuw and Leroy computation of LTS based on subsets of size k requires $q = \binom{n}{k}$ subsets which is usually still too large for realistic applications. When n is small enough:

1. Select h.
2. Generate all possible subsets with k observations, and compute the regression coefficients, say $\widehat{\boldsymbol{\beta}}(1), \ldots, \widehat{\boldsymbol{\beta}}(1)$.
3. Compute the residuals using all n observations, and from this the LTS criterion.
4. The LTS estimate corresponds to the $\widehat{\boldsymbol{\beta}}(l)$ that minimizes the objective function (3).

Rousseeuw and van Driessen [11] propose a fast algorithm for computing LTS. The trick is to iterate a few steps on a large number of starting values, and keep the 10 (say) most promising ones. These are then used for full iteration, yielding the final estimate. The resulting algorithm makes LTS estimation faster.

Our proposed procedure is a combination of IRLS and LTS methods. IRLTS estimator involves computing the hyperplane that minimizes the sum of the smallest h squared residuals and use the weighted least square estimation for $\boldsymbol{\beta}$ in each iteration. To motivate our estimator and following the fast LTS algorithm [11], it is convenient to write IRLTS algorithm with involving the residuals as follow.

Concentration-step:

1. Choose h observations.
2. Compute $\hat{\boldsymbol{\beta}}$ based on h observations using IRLS method .
3. Use the estimate $\hat{\boldsymbol{\beta}}$ to calculate residuals: $e_{ij} = Y_{ij} - \hat{\mu}_{ij}$ based on equation $\hat{\boldsymbol{\mu}}_i = g^{-1}(\mathbf{X_i}\hat{\boldsymbol{\beta}})$ of n observations.
4. Sort $|e_{ij}|$ for j = 1, ..., m_i, i = 1, ... , n in ascending order: $|e_{11}| \leq |e_{12}| \leq \ldots \leq |e_{ij}|$.
5. Choose h observations which have the lowest h residuals, we denote the h observations as subset H.

The repetitions of concentration-step will produce an iteration process.

IRLTS algorithm:

1. Choose h observations.
2. Compute $\hat{\boldsymbol{\beta}}$ based on h observations using IRLS by (2), we obtain $\hat{\boldsymbol{\mu}}_i = g^{-1}(\mathbf{X_i}\hat{\boldsymbol{\beta}})$.
3. Calculate residuals: $e_{ij} = Y_{ij} - \hat{\mu}_{ij}$ of n observations.
4. Sort $|e_{ij}|$ in ascending order: $|e_{11}| \leq |e_{12}| \leq \ldots \leq |e_{ij}|$
5. Choose h_1 observations which have the lowest h_1 residuals, we denote as subset H_1.
6. Run concentration-step on H_1 twice, and we obtain H_1^*.
7. Repeat step 1- step 6 for $\binom{n}{h}$ times

8. From the $\binom{n}{h}$ results, choose the best 10 subsets H_q, q=1,…,10.
9. Run concentration-step on the best 10 subsets H_q until convergence.
10. Choose the best subset H.

IV. Simulation Study

To look at the performance of the proposed method, we have done a simulation study by generating N=1000 observations from 200 subjects with 5 repeated measures. The model for data generation is as follows:

$$u_{ij} = \beta_0+\beta_1 x_{ij}$$

where $\beta_0=\beta_1=1$, i=1,2,… 200 and j=1,2,..,5. The covariates x_{ij} are i.i.d. from a uniform distribution Unif(1,5). For this longitudinal data the normal distributed model is used. We generated data based on the underlying true correlation structures as exchangeable (EXCH) and autoregressive-1 (AR1) with α=0.3 and 0.7. We considered data without outliers (ε = 0%) as well as contaminated data (ε = 10%, 20% and 30%). The contamination is generated from normal distribution N(100,1), we set two cases for the contamination, i.e. randomly spread over the sample (case A) and randomly spread over the half upper x_{ij} values of the sample (case B). For each scenario 1000 Monte Carlo data sets were generated. We evaluated the results using relative efficiency (RE) of IRLTS to IRLS and the mean square error (MSE) of $\hat{\beta}$ which we define as

$$RE_{IRLTS/IRLS} = Var(\hat{\beta}_i^{IRLTS})\{Var(\hat{\beta}_i^{IRLS})\}^{-1}$$

and

$$MSE = \frac{1}{1000}\sum_{s=1}^{1000}\left(\hat{\beta}_i^{(s)} - \beta_i\right)^2, \text{ with } i = 0,1,$$

where Var(.) is the variance. We provide the expected values (E), and the relative efficiency resulted from our simulation in Table I - Table IV and the MSEs in Table V- Table VI.

The efficiency of IRLTS and IRLS for clean data (i.e. when ε = 0%) is almost equal since RE ~ 1 for each case, but IRLTS is more efficient than IRLS when data contain outliers. The parameter estimates of IRLS are much more influenced by the outliers than the parameter estimates of IRLTS. From the expected values we can see that the more outliers contained in the data the larger the deviation of IRLS estimates from the parameter (i.e. $\beta_0 = \beta_1 = 1$), while the parameter estimates of IRLTS are almost stable and close to the parameter.

Table 1. Simulation Result for Longitudinal Data with Exchangeable Correlation Matrix with $\alpha = 0.3$

Case	Coeff.	ε	$E(\hat{\beta}_{IRLS})$	$E(\hat{\beta}_{IRLTS})$	$RE_{(IRLTS/IRLS)}$
Case A	$\hat{\beta}_0$	0%	1.00500	1.00499	1.00178
		10%	10.08781	1.02474	0.00294
		20%	17.55050	1.08504	0.00401
		30%	23.73423	1.15940	0.00657
	$\hat{\beta}_1$	0%	0.99846	0.99848	1.00278
		10%	0.99604	0.99948	0.00267
		20%	0.98073	0.99775	0.00393
		30%	0.94544	0.99577	0.00649
Case B	$\hat{\beta}_0$	0%	1.02179	1.01050	0.99466
		10%	-5.20456	1.00195	0.00577
		20%	-10.30572	0.95625	0.01606
		30%	-14.27400	0.84177	0.03562
	$\hat{\beta}_1$	0%	0.99318	0.99693	0.99369
		10%	5.95079	0.98675	0.00483
		20%	9.99122	0.99425	0.05797
		30%	13.21288	1.04650	0.11281

Table 2. Simulation Result for Longitudinal Data with Exchangeable Correlation Matrix with $\alpha = 0.7$

Case	Coeff.	ε	$E(\hat{\beta}_{IRLS})$	$E(\hat{\beta}_{IRLTS})$	$RE_{(IRLTS/IRLS)}$
Case A	$\hat{\beta}_0$	0%	1.01266	1.01187	1.01763
		10%	9.93676	1.01672	0.00433
		20%	17.31940	1.05446	0.00505
		30%	23.68107	1.16184	0.00769
	$\hat{\beta}_1$	0%	0.99600	0.99613	1.01882
		10%	1.04154	1.00103	0.00403
		20%	1.05223	1.00661	0.00478
		30%	0.95609	0.99443	0.00770
Case B	$\hat{\beta}_0$	0%	1.05174	1.01922	0.99052
		10%	-5.20125	1.00911	0.00837
		20%	-10.28303	0.98818	0.00528
		30%	-14.25209	0.83684	0.04304
	$\hat{\beta}_1$	0%	0.98159	0.99245	0.99948
		10%	5.93523	0.98438	0.00722
		20%	9.96316	0.97828	0.00489
		30%	13.21954	1.05368	0.12544

Table 3. Simulation Result for Longitudinal Data with Autoregressive-1 Correlation Matrix with $\alpha = 0.3$

Case	Coeff.	ε	$E(\hat{\beta}_{IRLS})$	$E(\hat{\beta}_{IRLTS})$	RE(IRLTS/IRLS)
Case A	$\hat{\beta}_0$	0%	0.99848	0.99744	1.03077
		10%	10.04370	1.02103	0.00247
		20%	17.60751	1.07994	0.00359
		30%	23.63332	1.15166	0.00621
	$\hat{\beta}_1$	0%	1.00043	1.00079	1.04004
		10%	1.00954	1.00069	0.00217
		20%	0.96174	0.99879	0.00356
		30%	0.97699	0.99897	0.00625
Case B	$\hat{\beta}_0$	0%	1.01800	1.00963	1.00481
		10%	-5.40516	0.99210	0.00113
		20%	-10.07597	0.94809	0.00393
		30%	-14.18714	0.85628	0.05029
	$\hat{\beta}_1$	0%	0.99341	0.99621	1.00108
		10%	5.98712	0.99180	0.00541
		20%	9.89834	0.99916	0.01079
		30%	13.19889	1.04016	0.16313

Table 4. Simulation Result for Longitudinal Data with Autoregressive-1 Correlation Matrix with $\alpha = 0.7$

Case	Coeff.	ε	$E(\hat{\beta}_{IRLS})$	$E(\hat{\beta}_{IRLTS})$	RE(IRLTS/IRLS)
Case A	$\hat{\beta}_0$	0%	1.00024	1.00079	1.07947
		10%	9.93814	1.00872	0.00412
		20%	17.52048	1.07392	0.00486
		30%	23.77262	1.16751	0.00790
	$\hat{\beta}_1$	0%	0.99960	0.99939	1.06915
		10%	1.04489	1.00404	0.00379
		20%	0.98595	1.00065	0.00481
		30%	0.94520	0.99331	0.00785
Case B	$\hat{\beta}_0$	0%	1.03666	1.01330	0.99589
		10%	-5.18276	1.01138	0.00618
		20%	-10.37403	0.97097	0.00484
		30%	-14.34357	0.80114	0.04723
	$\hat{\beta}_1$	0%	0.98765	0.99545	0.99343
		10%	5.92778	0.98511	0.00464
		20%	10.00611	0.98523	0.00461
		30%	13.25732	1.07041	0.14581

The consistency of the estimators is assessed through their MSEs (see Table V and Table VI). When data contain outliers, the MSEs of IRLTS are relatively small compared to the MSEs of the classical GEE (IRLS). From the result we conclude that IRLTS is robust against outliers.

Table 5. Mean Square Error of Parameter Estimates for Data with Exchangeable Correlation Matrix

Case	Coeff.	ε	$\alpha = 0.3$		$\alpha = 0.7$	
			IRLS	IRLTS	IRLS	IRLTS
Case A	$\hat{\beta}_0$	0%	0.01664	0.01666	0.03141	0.03194
		10%	90.44780	0.02372	88.48682	0.03765
		20%	291.41893	0.07749	283.90429	0.09182
		30%	544.84787	0.20944	540.15839	0.22414
	$\hat{\beta}_1$	0%	0.00165	0.00165	0.00299	0.00305
		10%	0.87157	0.00233	0.94041	0.00378
		20%	1.91855	0.00754	1.92692	0.00924
		30%	3.08263	0.02001	2.82052	0.02174
Case B	$\hat{\beta}_0$	0%	0.01726	0.01680	0.03467	0.03206
		10%	41.61809	0.01801	41.89402	0.02886
		20%	136.26362	0.13754	133.06421	0.03051
		30%	240.63699	0.28655	240.56976	0.36852
	$\hat{\beta}_1$	0%	0.00171	0.00167	0.00346	0.00318
		10%	24.88252	0.00197	24.75140	0.00309
		20%	82.04616	0.06984	80.97956	0.00361
		30%	149.95039	0.09196	150.19042	0.11242

Table 6. Mean Square Error of Parameter Estimates for Data with Autoregressive-1 Correlation Matrix

Case	Coeff.	ε	$\alpha = 0.3$		$\alpha = 0.7$	
			IRLS	IRLTS	IRLS	IRLTS
Case A	$\hat{\beta}_0$	0%	0.01309	0.01349	0.02329	0.02515
		10%	89.85427	0.02033	87.75463	0.03249
		20%	292.90715	0.06774	289.43715	0.08572
		30%	539.61490	0.19286	543.38625	0.22402
	$\hat{\beta}_1$	0%	0.00131	0.00136	0.00231	0.00247
		10%	0.88483	0.00192	0.86341	0.00328
		20%	1.85020	0.00659	1.77790	0.00856
		30%	2.98370	0.01864	2.71468	0.02134
Case B	$\hat{\beta}_0$	0%	0.01354	0.01337	0.02704	0.02577
		10%	87.20910	0.05218	42.39676	0.02589
		20%	164.20519	0.16609	135.05622	0.02836
		30%	237.06421	0.34325	243.38623	0.41559
	$\hat{\beta}_1$	0%	0.00135	0.00133	0.00273	0.00258
		10%	27.51679	0.01437	24.80085	0.00262
		20%	86.50801	0.07905	81.76371	0.00323
		30%	149.53554	0.11950	151.10469	0.13076

V. Conclusion

Our proposed method have two different iterations in its procedure, one is the iteration for the estimation of regression parameter using IRLS method, and the other iteration is for selecting the best subset H for calculating the parameter estimate. We have shown that this procedure can minimize the effect of outliers on parameter estimation; IRLTS can produce a relatively efficient and consistent estimator compared to the classical GEE (IRLS). Base on the MSE, IRLTS performs much better than the classical GEE. Hence, robust GEE using IRLTS is a good choice for longitudinal data analysis when data contains outliers.

References

[1] P. McCullagh and J.A. Nelder, *Generalized Linear Models*. London : Chapmann and Hall, 1989.

[2] K.Y. Liang and S.L. Zeger, “Longitudinal data analysis using generalized linear models”, *Biometrika*, vol 73, pp. 13-22, 1986.

[3] B.F. Qaqish and J.S. Preisser, “Resistant fits for regression with correlated outcomes an estimating equations approach”, *Journal of Statistical Planning and Inference*, vol. 75, pp. 15-431, 1999.

[4] P.S. Gill, “A Robust Mixed Linear Model Analysis for Longitudinal Data”, *Statistics in Medicine*, vol. 19, pp. 975-987, 2000.

[5] S.H. Jung and Z. Ying, “Rank-Based Regression With Repeated Measurements Data”, *Biometrika*, vol. 90, pp. 732-740, 2003.

[6] Abebe, A., McKean, J. W. & Kloke, J. D., Bilgic, Y. “Iterated Reweighted Rank-Based Estimates for GEE Models”, submitted.

[7] R.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.

[8] J.A. Víšek, “The Least Trimmed Squares – random carriers”. *Bulletin of the Czech Econometric Society*, vol. 6, pp. 1-30, 1999.

[9] J.A. Víšek, The Least Trimmed Squares. Part II: $\sqrt{n}$-consistency. *Kybernetika*, vol. 42, pp. 181-202, 2006a.

[10] J.A. Víšek, “The Least Trimmed Squares. Part III: Asymptotic normality”. *Kybernetika*, vol. 42, pp. 203-224, (2006b).

[11] P.J. Rousseeuw, and K. van Driessen, “ Computing LTS Regression for Large Data Sets”. *Data mining and Knowledge Discovery*, vol. 12, pp. 29-45, 2006.