METHODOLOGY

TELBS robust linear regression method

MA Tabatabai[1]
WM Eby[1]
H Li[1]
S Bae[2]
KP Singh[2]

[1]Department of Mathematical Sciences, Cameron University, Lawton, OK, [2]Department of Medicine, University of Alabama, Birmingham, AL, USA

Correspondence: WM Eby
Department of Mathematical Sciences, Cameron University, Lawton, OK, USA
Tel +1 580 581 2395
Fax +1 580 581 2616
Email weby@cameron.edu

Abstract: Ordinary least squares estimates can behave badly when outliers are present. An alternative is to use a robust regression technique that can handle outliers and influential observations. We introduce a new robust estimation method called TELBS robust regression method. We also introduce a new measurement called $S_h(i)$ for detecting influential observations. In addition, a new measure for goodness of fit, called R^2_{RFPR}, is introduced. We provide an algorithm to perform the TELBS estimation of regression parameters. Real and simulated data sets are used to assess the performance of this new estimator. In simulated data with outliers, the TELBS estimator of regression parameters performs better in comparison with least squares, M and MM estimators, with respect to both bias and mean squared error. For rat liver weights data, none of the estimators (least squares, M, and MM) are able to estimate the parameters accurately. However, TELBS does give an accurate estimate. Using real data for brain imaging, the TELBS and MM methods were equally accurate. In both of these real data sets, the $S_h(i)$ measure was very effective in identifying influential observations. The robustness and simplicity of computations of TELBS model parameters make this method an appropriate one for analysis of linear regression. Algorithms and programs have been provided for ease in implementation, including all relevant statistics necessary to perform a complete analysis of linear regression.

Keywords: robust linear regression, least squares estimator, M and MM estimators, magnetic resonance imaging, Cook's distance, detection of influential observations, Studentized residual

Background

Linear regression is one of the most popular and widely used models for analyzing the effect of explanatory variables on a response variable, and it has widespread application in every field of study, including biomedical research. The least squares method is a commonly used method for parameter estimation of regression coefficients, but with the presence of outliers in the data, regression parameter estimates, hypothesis testing, and predictions may no longer be reliable. In biomedical research it is an absolute necessity to make reliable conclusions from the data and to be able to identify and analyze data containing outliers and influential observations. The primary purpose of a robust linear regression is to fit a model which gives resilient results in the presence of outliers. Rousseeuw and Leroy[1] define vertical outliers as those datum points that have outlying values in the direction of the response variable, while leverage points are outliers in the direction of explanatory variables. Good leverage points are those datum points which are located near the regression line. The ordinary least squares method

65

minimizes the variance of residuals. The presence of good leverage does not affect ordinary least squares estimates of the regression coefficients, but does affect the estimated standard error of the estimated coefficients. On the other hand, bad leverage points influence both the least squares estimates of parameters as well as their estimated standard errors. In addition to leverage points and outliers, an observation may be influential if its removal would significantly change the estimate of regression coefficients. Edgeworth[2] proposed a robust method called the least absolute deviation or L^1 norm. Huber[3] introduced the M estimation method for regression. The M estimation has unbounded influence and is not robust with respect to leverage points, but it is fairly robust with respect to vertical outliers. It uses a certain cost function of residuals. Rousseeuw[4] introduced the least trimmed squares estimates. Rousseeuw and Yohai[5] introduced the S estimator. Yohai and Zammar[6] introduced the τ estimator of linear regression coefficients, which is a high efficiency estimator and has a high breakdown point. Tabatabai and Argyros[7] extended the τ estimates to nonlinear regression models (see Fasano[8] for additional use of such nonlinear robust methods). The most popular and widely used method of robust regression is MM estimation of linear regression parameters. The MM estimator was originally introduced by Yohai.[9] It has high asymptotic efficiency under the normal distribution assumption and can possibly attain a 50% breakdown point. Linear regression methods are frequently used to assist scientists in biomedical research.

Recently, Yang et al[10] used robust regression to analyze voxelwise correlations between functional and structural neuroimaging modalities. Yeo et al[11] introduced a new robust regression-based exon array protocol for exon array analysis which led to discoveries about the complexity of alternative RNA splicing in human embryonic stem cells and their transition to neural stem cells. Xu et al[12] used robust linear regression and analyzed the association between DNA copy number and gene expression in tumor cells from metastatic lymph nodes in patients with oral squamous cell carcinoma. In computer vision, robust regression methods have been used extensively to estimate surface model parameters in small image regions and imaging geometry of multiple cameras (see Stewart[13] and the more recent work of Zaharescu and Horaud[14]). In the analysis of gene microarrays, Ahdesmäki et al[15] proposed use of robust M estimation of regression when dealing with biological time series models. Karan et al[16] assessed the use of visual aids when obtaining informed consent for cataract surgery, using linear regression model as their model of choice. The study of van Vliet et al[17] used

robust linear regression to demonstrate that low plasma apolipoprotein E gene levels in midlife are associated with increased risk of Alzheimer's disease in later life. In a large-scale study, Pinheiro et al[18] used robust linear regression to study the association between race and hormone levels in breast cancer patients. Floyd et al[19] used multiple linear regressions to conclude that cerebral blood flow increases after cardiac surgery. Mircean et al[20] used robust linear regression to obtain an accurate estimate of protein expression in different samples on the lysate microarray data. Kocak et al[21] used linear regression to predict breast tissue resection weights for reduction mammoplasty. Cepeda and Carr[22] used a robust linear regression to determine whether there was a difference in the opioid requirement between men and women. These cases are representative of the prevalence of use of linear regression in biomedical research, and clearly there is a need for simple and effective methods for finding an accurate estimate in the presence of outliers.

The goal of this paper is to introduce the TELBS robust linear regression estimator to researchers in every field, including biomedicine. This new method for robust regression has a bounded influence and high breakdown point and asymptotic efficiency under normal distribution and is able to estimate the parameters of linear regression in a way that is close to the parameter estimates we would have estimated with the absence of outliers in the data. The next section introduces the TELBS robust linear regression model, followed by an algorithm describing its implementation. We then apply this new model to analyze a problem involving magnetic resonance imaging of cerebral blood flow and a second problem involving drug absorption in rat livers. In addition, Monte Carlo simulations are performed to evaluate the robustness of the TELBS method, in comparison with ordinary least squares, M and the MM methods. Computer programs for application data sets are given in the Appendix.

TELBS robust linear regression model

Suppose $y_1, y_2, \ldots, y_n$ is a sample of n observations. In this paper, we consider the standard linear regression model of the form

$$y_i = x_i^t \beta + \varepsilon_i,$$

where $i = 1, 2, \ldots, n$, and the parameter vector $\beta^t = (\beta_0, \beta_1, \ldots, \beta_k)$. The errors $\varepsilon_i's$ are random variables. In a designed experiment, $x_{ij}'s$ are fixed but when $x_{ij}'s$ are observational, they are random variables. The predictor can be

fixed, random, or mixed. The ordinary least squares estimate of parameter vector β is found by

$$\hat{\beta}_L = \arg\min_{\beta \in \mathrm{R}^{k+1}} \sum_{i=1}^{n} \varepsilon_i^2,$$

where

$$x_i^t = (x_{i0}, x_{i1}, ..., x_{ik}); \ 1 \le i \le n$$

with $x_{i0} \equiv 1$ and k being the number of predictor variables in the model.

The TELBS estimate of the parameter vector β is found by

$$\hat{\beta} = \arg\min_{\beta \in \mathrm{R}^{k+1}} \sum_{i=1}^{n} \frac{\rho_\omega(t_i)}{L_i}, \tag{1}$$

where

$$\rho_\omega(x) = 1 - Sech(\omega x),$$

and the positive real number ω is called the tuning constant. The function $Sech(\cdot)$ is the hyperbolic secant function and t_i's are defined by

$$t_i = \frac{(y_i - x_i^t \beta)(1 - h_{ii})}{\sigma}, \tag{2}$$

where σ is the error standard deviation, and h_{ii}'s are the diagonal elements of the hat matrix of the form

$$H = X(X^t X)^{-1} X^t,$$

and the design matrix X is

$$X = \begin{pmatrix} x_{10} & x_{11} & ... & x_{1k} \\ x_{20} & x_{21} & ... & x_{2k} \\ x_{n0} & x_{n1} & ... & x_{nk} \end{pmatrix} = (X_0 X_1 \ ... \ X_k)$$

For $j = 1, 2, ..., k$, we define

$$M_j = Median\left\{ |x_{1j}|, |x_{2j}|, ..., |x_{nj}| \right\}$$

and for $i = 1, 2, ..., n$, we define

$$L_i = \sum_{j=1}^{k} Max\left\{ M_j, |x_{ij}| \right\}$$

If σ is unknown, one may use one of the following two estimators of σ which were proposed by Rousseeuw and Croux[23]

$$\hat{\sigma} = 1.1926 \, Median_{\{i:1 \le i \le n\}} \left(Median_{\{j:1 \le j \le n\}} |r_i - r_j| \right) \tag{3}$$

or

$$\hat{\sigma} = 2.2219 \left\{ |r_i - r_j|; i < j, i, j = 1, ..., n \right\}_{(p)} \tag{4}$$

where $r_i = y_i - x_i^t \hat{\beta}$ and $p = \begin{pmatrix} [n/2]+1 \\ 2 \end{pmatrix}$ is the binomial coefficient and $\{.\}_{(p)}$ is the pth order statistic.

The abovementioned estimators of σ have very high breakdown points. Under the assumption of normality, (3) and (4) have higher efficiency than the median absolute deviation. In this paper, all of our computations are performed using formula (3).

The function $\rho_\omega : \mathrm{R} \to \mathrm{R}$ is a differentiable function satisfying the following properties:

   i. ρ_ω is bounded,
   ii. $\forall x \in \mathrm{R}, \rho_\omega(x) \ge 0,$
   iii. $\rho_\omega(0) = 0,$
   iv. $\forall x \in \mathrm{R}, \rho_\omega(x) = \rho_\omega(-x),$
   v. $\forall a, b \in \mathrm{R}, |a| > |b| \Rightarrow \rho_\omega(a) \ge \rho_\omega(b),$
   vi. $\lim_{x \to \infty} \rho_\omega(x) = \lim_{x \to -\infty} \rho_\omega(x) = 1,$
   vii. $\forall \kappa > 0, \lim_{x \to \infty} \rho_\omega(\kappa x)/\rho_\omega(x) = 1,$
   viii. $\lim_{|x| \to \infty} d\rho_\omega(x)/dx = 0.$

Since ρ_ω satisfies property (viii), it is called a redescending function. The function ρ_ω is also a slow variation function because it satisfies property (vii). Taking the partial derivatives of (1) with respect to parameters and setting them equal to zero results in the following system of equations

$$\sum_{i=1}^{n} \frac{\psi_\omega(t_i)}{L_i} \frac{\partial t_i}{\partial \beta_j} = 0 \tag{5}$$

where ψ_ω is the derivative of ρ_ω which is equal to $\psi_\omega(x) = \omega Sech(\omega x) Tanh(\omega x)$.

Define the weights w_i as

$$w_i = \frac{\psi(t_i)(1 - h_{ii})}{\sigma(y_i - x_i^t \beta)L_i} \tag{6}$$

Then the equation (5) can be written as

$$\sum_{i=1}^{n} w_i(y_i - x_i^t \beta)x_i = 0.$$

The matrix of weights, W is a diagonal matrix, the elements of which on the main diagonals are $w_1, w_2, ..., w_n$, and the estimator of the parameter vector β is given by

$$\hat{\beta}(X, y) = (X^t W X)^{-1} X^t W y.$$

Theorem 1

Let M be an invertible matrix of size $k + 1$. The TELBS estimator $\hat{\beta}(X, y)$ has the following properties:

a. $\forall \alpha \in \mathbb{R}^{k+1}, \hat{\beta}(X, y + X\alpha) = \hat{\beta}(X, y) + \alpha$ (regression equivariance)

b. $\forall \gamma \in \mathbb{R}, \hat{\beta}(X, \gamma y) = \gamma \hat{\beta}(X, y)$ (Scale equivariance)

c. $\forall M, \hat{\beta}(XM, y) = M^{-1}\hat{\beta}(X, y)$. (affine equivariance)

Asymptotically, $\hat{\beta}$ has a normal distribution with mean β and variance-covariance matrix

$$V = \frac{\sigma^2 E(\psi_\omega^2(t))}{[E(\psi_\omega'(t))]^2} E((X^t X)^{-1}),$$

where

$$E[\psi_\omega'(t)] = \int_{-\infty}^{\infty} \psi_\omega'(t) \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt$$

and

$$E[\psi_\omega^2(t)] = \int_{-\infty}^{\infty} \psi_\omega^2(t) \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt.$$

The function $\psi_\omega'(t)$ is defined as

$$\psi_\omega'(t) = \omega^2[Sech^3(\omega t) - Sech(\omega t)Tanh^2(\omega t)]$$

Under the assumption of normality for the underlying distribution, the asymptotic efficiency A_{eff} is defined as

$$Aeff = \frac{(E[\psi_\omega'(t)])^2}{E[\psi_\omega^2(t)]}. \tag{7}$$

The tuning constant ω can be calculated by solving equation (7) for ω. For TELBS model, the numerical values for ω at the asymptotic efficiency levels of 0.70, 0.75, 0.80, 0.85, 0.90, and 0.95 are approximately 0.901, 0.8118, 0.721, 0.628, 0.525, and 0.405 respectively. Although the choice for the tuning constant ω is left for the investigator to decide, we do recommend an efficiency of 0.85 which corresponds to $\omega = 0.628$. The solid curves in Figure 1 show the graph of ρ_ω and the dashed curves are the graphs of ψ_ω functions with tuning constant ω values set to 0.901, and 0.405. The three-dimensional graphs of $\rho_\omega(x)$ and $\psi_\omega(x)$ as a function of ω and x are shown in Figures 2 and 3, respectively.

An estimate for the variance-covariance matrix is given by

$$\hat{V} = \frac{n^2 \hat{\sigma}^2 \sum_{i=1}^{n} \psi_\omega^2(t_i)}{(n-p)\left(\sum_{i=1}^{n} \psi_\omega'(t_i)\right)^2}(X^t X)^{-1}.$$

The robust deviance is defined as

$$D = 2\hat{\sigma}^2 \sum_{i=1}^{n} \frac{1 - Sech(\omega t_i)}{L_i}.$$

The deviance plays a major role in model fitting. A smaller value of deviance is preferred to the larger value. Following the Akaike information criterion[24] and Ronchetti,[25] the robust equivalence of the Akaike information criterion (AICR) is given by
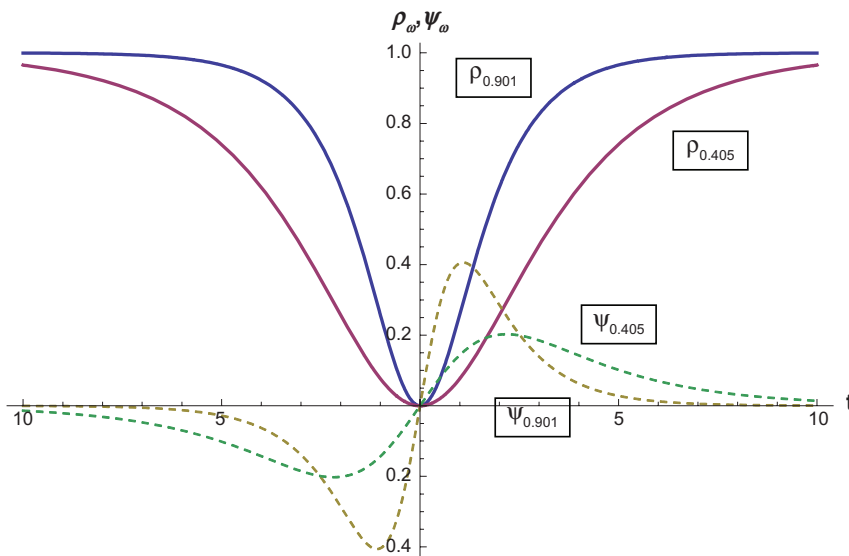

Figure 1 Graph of ρ_ω and of ψ_ω functions for $\omega = 0.901$ and $\omega = 0.405$.
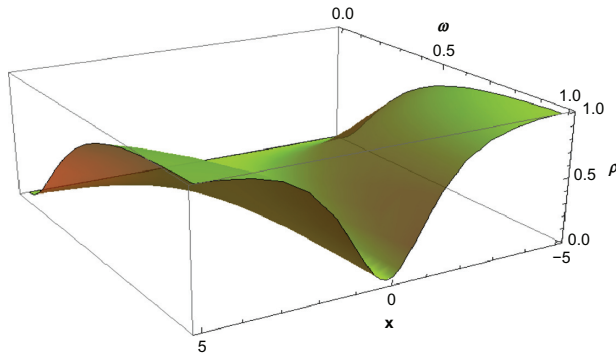
Figure 2 Three-dimensional graph of ρ_ω function.

$$AICR = \frac{D}{\hat{\sigma}^2} + 2p\frac{E[\psi_\omega^2(t)]}{E[\psi_\omega'(t)]},$$

and the robust Schwarz information criterion (BICR) is given by

$$BICR = \frac{D}{\hat{\sigma}^2} + 2p\ln(n)$$

The following is a robust version of the coefficient of determination

$$R^2 = 1 - \left(\frac{\underset{\{i:1\leq i\leq n\}}{Median}|r_i|}{\underset{\{i:1\leq i\leq n\}}{Median}\left|y_i - \underset{\{j:1\leq j\leq n\}}{Median}(y_j)\right|}\right)^2.$$

If no intercept is used in the regression equation, then

$$R^2 = 1 - \left(\frac{\underset{\{i:1\leq i\leq n\}}{Median}|r_i|}{\underset{\{i:1\leq i\leq n\}}{Median}|y_i|}\right)^2.$$
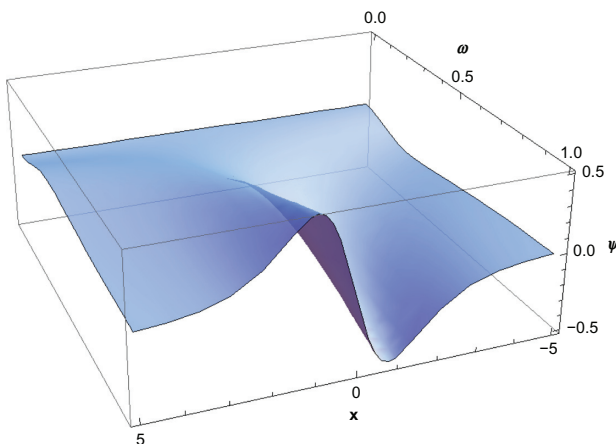

Figure 3 Three-dimensional graph of ψ_ω function.

For more details, see Rosseeuw and Leroy.[1]

To measure the overall effect of the covariates on the response variable, one can use an F type statistic of the form

$$F = \frac{2E[\psi_\omega'(t)]\left(\sum_{i=1}^{n}\rho_\omega(ConstantOnlyModel)_i - \sum_{i=1}^{n}\rho_\omega(FullModel)_i\right)}{(p-1)E[\psi_\omega^2(t)]}.$$

There are numerous variable selection techniques available in the literature. One may use the stepwise procedure involving forward selection or backward elimination. For each set $S \subseteq \{x_1, x_2, \ldots, x_p\}$ of explanatory variables, the robust final predicted error, as defined by Maronna et al,[26] is denoted by $RFPE(S)$ and is defined as

$$RFPE(S) = \frac{\sum_{i=1}^{n}\rho_\omega(t_i)}{n} + \frac{\#(S)\sum_{i=1}^{n}\psi_\omega^2(t_i)}{n\sum_{i=1}^{n}\psi_\omega'(t_i)},$$

where $\#(S)$ is the number of elements in the set S. In the forward or backward elimination, choose the one for which inclusion or deletion results in smallest value of $RFPE$.

Based on RFPE, we introduce a new coefficient of determination called R_{RFPR}^2 which is defined as

$$R_{RFPR}^2 = 1 - \left(\frac{RFPE(FullModel)}{RFPE(ConstantOnlyModel)}\right)^2.$$

This new statistic measures the overall performance of the model and is based on the value of $RFPE(FullModel)$ which is the RFPE when the final model is selected and $RFPE(ConstantOnlyModel)$ when only constant parameter is included in the linear regression model. The range for the values of R_{RFPR}^2 is from a low of zero up to a maximum of one.

To perform hypothesis testing, we let $\Omega \subseteq \mathbf{R}^{k+1}$ be the parameter space and $\{\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_q}\}$ be a subset of $\{\beta_0, \beta_1, \ldots, \beta_k\}$.

Define

$$\Omega_0 = \{\beta \in \Omega : \beta_{j_1} = \beta_{j_2} = \ldots = \beta_{j_q} = 0\},$$

and the function $f(\beta)$ as

$$f(\beta) = \sum_{i=1}^{n}\frac{\rho_\omega(t_i)}{L_i}.$$

Then a robust likelihood ratio type test statistic for testing the null hypothesis $H_0:\beta \in \Omega_0$ against the alternative $H_1:\beta \in \Omega_0^c$ is

$$S_n^2 = 2(\sup_{\beta \in \Omega_0} f(\beta) - \sup_{\beta \in \Omega} f(\beta))/q$$

For more information, see Hampel et al.[27] Asymptotically under the null hypothesis $\{E[\psi_\omega'(t)]/E[\psi_\omega^2(t)]\}S_n^2$ has a Chi-square distribution with q degrees of freedom, where the Wald-type test statistic is defined as

$$W_n^2 = n\left(\hat{\beta}_{j_1}, \hat{\beta}_{j_2}, ..., \hat{\beta}_{j_q}\right)V_q^{-1}\left(\hat{\beta}_{j_1}, \hat{\beta}_{j_2}, ..., \hat{\beta}_{j_q}\right)^t,$$

where $(1/n)V_q$ is the asymptotic variance-covariance matrix for the given vector $\left(\hat{\beta}_{j_1}, \hat{\beta}_{j_2}, ..., \hat{\beta}_{j_q}\right)$. The null distribution of the statistic W_n^2 is asymptotically a Chi-square distribution with q degrees of freedom.

To identify the outliers, influential observations, or leverage points, we recommend that the following diagnostic measures along with graphical measures be part of any regression diagnostic efforts. The first measure we recommend is the robust Studentized residual using TELBS robust estimates of parameters. This statistic has the form

$$SR_i(TELBS) = \frac{t_i}{\hat{\sigma}\sqrt{1 - h_{ii}}},$$

where most preferably $\hat{\sigma}$ should be calculated using formula (3) or (4). The next diagnostic measure that needs to be calculated is the robust Cook's distance using TELBS estimated parameters. The robust Cook's distance has the form

$$CD_i(TELBS) = \frac{h_{ii}t_i^2}{p(1 - h_{ii})^4}$$

We recommend that in addition to considering the elements of the main diagonal of the hat matrix h_{ii}'s, one should use the following influence measure called $S_h(i)$ which we define as

$$S_h(i) = \frac{h_{ii} - \underset{\{i:1 \le i \le n\}}{Median}(h_{ii})}{\hat{\sigma}_h},$$

where

$$\hat{\sigma}_h = 1.1926 \underset{\{i:1 \le i \le n\}}{Median}\left(\underset{\{j:1 \le j \le n\}}{Median}\left|h_{ii} - h_{jj}\right|\right).$$

Large values of $|S_h(i)|$ indicates the presence of an influential observation. This statistic seems to be a very good measure in identifying the leverage points.

Algorithm

The following algorithm uses the TELBS method to estimate the parameter vector β of the linear regression model.

1. Set $j = 0$ and $\hat{\sigma}^0 = 1$. Calculate an initial estimate of vector β by
    minimizing the function $f(\beta) = \sum_{i=1}^{n} \frac{\rho_\omega(t_i)}{L_i}$ and call this estimate $\hat{\beta}^{(0)}$.
2. Set $j = j + 1$. Calculate $\hat{\sigma}^{(j)}$ and weights $w_i^{(j)}$ for $1 \le i \le n$ and use them to calculate the weights matrix $W^{(j)}$.
3. Use information from step 2 to calculate $\hat{\beta}^{(j)} = (X^tW^{(j)}X)^{-1}X^tW^{(j)}y$.
4. If convergence occurs, stop. Otherwise go to step 2 and continue the process.

    The convergence occurs when $\left\|\hat{\beta}^{(j)} - \hat{\beta}^{(j-1)}\right\| \cong 0$. The convergence criterion is to stop the algorithm when $\left\|\hat{\beta}^{(j)} - \hat{\beta}^{(j-1)}\right\| \le \varepsilon$ or $\left[\left\|\hat{\beta}^{(j)} - \hat{\beta}^{(j-1)}\right\|/\left\|\hat{\beta}^{(j)}\right\|\right] \le \varepsilon$. One may choose the value of $\varepsilon = 0.00001$.

Applications of TELBS robust linear regression

Cerebral blood flow

Arterial spin labeling and dynamic susceptibility contrast are used to measure cerebral blood flow. This technique can help patients with brain tumors. Warmuth et al[28] implemented an arterial spin labeling method which allows assessment of microvascular perfusion and is capable of making a distinction between low-grade and high-grade gliomas in patients with brain tumors. They also compared the arterial spin labeling method with dynamic susceptibility-weighted contrast material-enhanced magnetic resonance imaging for evaluation of blood flow. The authors investigated the linear association between arterial spin labeling and dynamic susceptibility-weighted contrast magnetic resonance imaging in the tumor region of interest. They believe that either method can identify high-grade and low-grade gliomas with a reasonable degree of accuracy. They conclude that arterial spin-labeling is a suitable method for assessment of microvascular perfusion and can identify high-grade and low-grade gliomas. Table 1 gives the values of three diagnostic measures SR_i, $S_h(i)$, and CD_i, using TELBS as a robust estimator of regression parameters. Considering all three measures, the strongest outliers appear to be observations 30 and 14 followed by observation 16. Table 2 shows the parameter estimates, standard errors, and one-sided P values using ordinary least squares estimator, ordi-

Table 1 Diagnostic measures SR_i, $S_h(i)$ and CD_i for cerebral blood flow data

Observation	SR_i	$S_h(i)$	CD_i
1	−0.305	2.018	0.004
2	−0.470	−0.701	0.003
3	−0.229	−0.166	0.001
4	0.367	0.297	0.003
5	0.472	−0.654	0.004
6	−0.118	−0.410	0.000
7	−0.726	−0.805	0.008
8	−1.325	−0.544	0.034
9	−0.354	−0.654	0.002
10	0.060	0.478	0.000
11	0.634	−0.838	0.006
12	0.580	−0.480	0.006
13	1.026	−0.025	0.027
14	6.214	4.260	3.114
15	−0.287	1.253	0.003
16	4.242	1.663	0.829
17	−0.755	4.018	0.044
18	−0.535	0.672	0.009
19	−0.219	0.739	0.001
20	0.031	0.356	0.000
21	−1.966	0.606	0.128
22	−1.361	−0.410	0.039
23	0.728	0.951	0.019
24	−0.577	−0.253	0.007
25	−0.648	−0.838	0.006
26	0.605	0.025	0.009
27	1.597	4.260	0.205
28	1.642	−0.073	0.067
29	0.574	−0.678	0.006
30	10.616	5.282	10.768
31	0.528	−0.678	0.005
32	0.264	1.253	0.002

nary least squares-estimator after removal of some influential observations, M estimator, MM estimator, and the TELBS estimators. In Figure 4, we have drawn the scatter plot of the ratio of the predictor variable (bolus tracking) versus the response variable (spin-labeling) and the graph of the fitted regression line to these data using the TELBS robust regression method. In this graph, observations 14, 16, and 30 are highlighted as potential outliers. The Mathematica program used to perform the computations for TELBS robust linear regression for this data set is provided in the Appendix. For this example, TELBS is an effective regression tool in estimating model parameters in the presence of outliers. Furthermore, we see that the diagnostic measure $S_h(i)$ helps to identify outliers.

Rat liver experiment

The following is a classic example of a case where the influence of a single observation can significantly change the outcome of the linear regression model. These data have appeared in Weisberg[29] as rat liver data. Nineteen rats were randomly selected and their body weights were measured. They were then placed under anesthesia and given an oral dose of 40 mg of a drug, in the amount of 40 mg per kilogram of body weight. After a fixed period of time, each rat was sacrificed and the liver weight and percent of the dose in the liver were determined. The goal of this study was to predict the percentage of drug dose absorbed by the rat liver using a linear function of body weight, liver weight, and relative drug dose. The investigator expected that the percentage of drug dose absorbed by the rat liver to be independent of dose and body weight. The ordinary least squares results point to the significance of both dose and body weight. On the other hand, diagnostic examinations in Table 3 reveal that observation 3 is a highly influential observation which has a high leverage. All three diagnostic measures point toward observation number 3 as a highly influential one. It is possible that rat number 3 received a significantly higher dose relative to its body weight. Table 4 shows the parameter estimates, standard errors, and one-sided P values using ordinary least squares estimator, ordinary least squares estimator after removal of influential observation 3, M estimator, MM estimator, and the TELBS estimator. When the influence of observation 3 is removed, then the ordinary least square estimator points to the nonsignificance of both dose and body weight as predictors of percentage of drug dose absorbed by rat liver. The least squares estimator, M estimator, and MM estimator all failed to give the correct conclusion. Only the TELBS estimator was in agreement with the nonsignificance of both drug dose and body weight.

Simulation

To evaluate the performance of the TELBS estimator in comparison with the least squares, M, and MM-estimators, we performed a simulation. We first generated 1000 samples of size n = 20 where the errors $\varepsilon_i' s$ were independent and identically distributed normal variables with a mean of 0 and a standard deviation of 1. The single explanatory variable x was then generated using a standard normal distribution with a mean of 0 and a standard deviation of 1. The response variable y is generated using the following equation

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

where the parameters are $\beta_0 = 1$ and $\beta_1 = 3$. Additionally, we generated 1000 samples of size n = 100 where the errors $\varepsilon_i' s$

Table 2 Parameter estimates, standard errors, and P values for the cerebral blood flow data

Method	Coefficients		Standard errors	
	Constant	Bolus tracking	Constant	Bolus tracking
LS	−0.0582	1.2997	0.1578	0.1685
One-sided P values	0.3575	0.0000		
LS OBS 30 removed	0.0950	1.0430	0.1170	0.0000
One-sided P values	0.0210	0.0000		
LS OB 14, 30 removed	0.1990	0.8640	0.0940	0.1090
One-sided P values	0.0215	0.0000		
LS OBS 14, 16, 30 removed	0.2612	0.7417	0.0716	0.0856
One-sided P values	0.0005	0.0000		
M	0.1442	0.9580	0.0801	0.0855
One-sided P values	0.0005	0.0000		
MM	0.2599	0.7445	0.0720	0.0768
One-sided P values	0.0003	0.0000		
TELBS efficiency 70%	0.2603	0.7449	0.0278	0.0297
One-sided P values	0.0000	0.0000		
TELBS efficiency 75%	0.2596	0.7470	0.0342	0.0365
One-sided P values	0.0000	0.0000		
TELBS efficiency 80%	0.2582	0.7504	0.0428	0.0457
One-sided P values	0.0000	0.0000		
TELBS efficiency 85%	0.2546	0.782	0.0570	0.0608
One-sided P values	0.0004	0.0000		
TELBS efficiency 90%	0.2462	0.7752	0.0784	0.0838
One-sided P values	0.0017	0.0000		
TELBS efficiency 95%	0.2244	0.8177	0.1149	0.1227
One-sided P values	0.0508	0.0000		

Abbreviations: LS, least squares; OBS, observation; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

were independent and identically distributed normal variables with a mean of 0 and a standard deviation of 1. The explanatory variables x_1 and x_2 were then generated using a standard normal distribution with a mean of 0 and a standard deviation of 1. The response variables $y_i's$ were calculated using the equation

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$$

where $\beta_0 = 1$ and $\beta_1 = 3$ and $\beta_2 = 0.50$. Mathematica software and R were used in the simulation process. To evaluate the robustness of these estimators, we randomly chose 5%, 25%, and 40% of the simulated observations and contaminated the selected data by magnifying their size by a factor
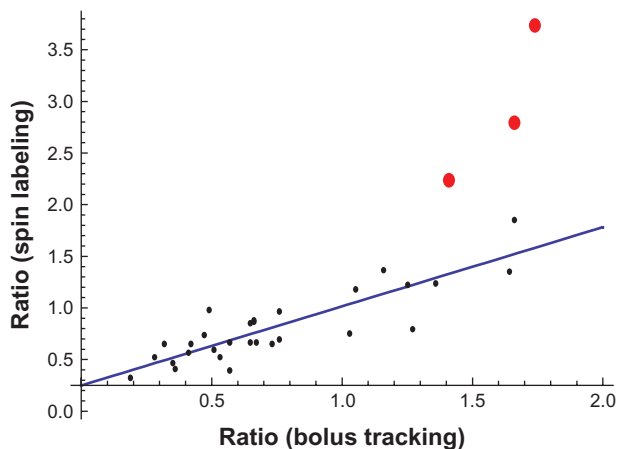
Figure 4 Scatter plot and fitted regression line using TELBS linear regression method.
Abbreviation: TELBS, Tabatabai, Eby, Li, Bae, and Singh.

Table 3 Diagnostic measures SR_i, $S_h(i)$ and CD_i for rat liver data

Observation	SR_i	$S_h(i)$	CD_i
1	1.379	0.088	0.102
2	−1.082	0.110	0.064
3	4.865	10.757	33.780
4	−1.274	−1.027	0.048
5	−1.499	3.474	0.361
6	0.025	−0.178	0.000
7	0.592	−0.563	0.014
8	0.989	1.288	0.083
9	0.075	−1.670	0.000
10	0.446	−0.835	0.006
11	−0.846	−0.838	0.024
12	0.567	0.000	0.016
13	−1.639	2.279	0.310
14	−0.477	−0.649	0.008
15	0.248	−1.525	0.001
16	−0.800	0.701	0.044
17	−0.250	0.361	0.003
18	0.669	−0.375	0.019
19	1.867	0.088	0.188

Table 4 Parameter estimates and standard errors for rat liver data

Method	Coefficients				Standard errors			
	Constant	BW	LW	DO	Constant	BW	LW	DO
LS	0.2659	−0.0212	0.0143	4.1781	0.1945	0.0079	0.0172	1.5226
One-sided P values	0.0960	0.0090	0.2095	0.0075				
LS (OBS 3 removed)	0.3114	−0.0078	0.0089	1.4848	0.2051	0.0187	0.0186	3.7131
One-sided P values	0.7550	0.3420	0.3185	0.3475				
M	0.2339	−0.0212	0.0171	4.1806	0.2047	0.0084	0.0181	1.6014
One-sided P values	0.2532	0.0116	0.3448	0.0090				
MM	0.2291	−0.0212	0.0186	4.1777	0.2132	0.0087	0.0189	1.6684
One-sided P values	0.2826	0.0148	0.3251	0.0123				
TELBS efficiency 90%	0.2931	−0.0070	0.0007	1.439	0.1662	0.0068	0.0147	1.3001
One-sided P values	0.0716	0.2907	0.9690	0.2583				
TELBS efficiency 95%	0.3009	−0.0073	0.0013	1.4671	0.1602	0.0067	0.0142	1.2538
One-sided P values	0.0604	0.2688	0.9293	0.2420				

Abbreviations: BW, body weight; DO, relative dose; LW, liver weight; OBS, observation; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

of 1000, first in the direction of explanatory variable(s), response variable(s), and both the response and explanatory variable(s). Finally, we estimated both bias and mean squared errors using the following equation

$$Bias = \left| \frac{\sum_{l=1}^{m} (\hat{\beta}_l)}{m} - \beta \right|$$

where m is the number of iterations in the simulation. The mean squared error was estimated by

$$MSE = \frac{\sum_{l=1}^{m} (\hat{\beta}_l - \beta)^2}{m}$$

Tables 5–10 give the summary of the simulation results for both small and large sample sizes and contamination levels of 5%, 25%, and 40%. The asymptotic efficiency for all of our simulation methods is 95%. By examining the simulation tables, we see that the least squares method performs poorly at all levels of contamination and for both small and large samples. The M method underperforms in all cases, except at the 5% contamination level and only when the contamination is in the direction of the response variable only. The MM method performs better than the M and least squares methods. Overall, the MM method does not seem to perform as well when the contamination level is 25% or 40%, and especially when the ratio of the number of parameters to the sample size becomes large. The TELBS method outperforms all the models under consideration for both small and large samples and at all levels of contamination in the direction of the response variable, the explanatory variable, or both.

Conclusion

In this paper we introduced a new robust estimator of linear regression parameters known as the TELBS estimator. Based on TELBS estimates of regression parameters, we developed

Table 5 Bias and mean square errors with contamination in the x direction

	5%		25%		40%	
	β_0	β_1	β_0	β_1	β_0	β_1
n = 20						
LS						
Bias	0.0062	2.9856	0.0058	2.9969	0.0089	2.9969
MS	0.5159	8.9283	0.4225	8.9818	0.3652	8.9819
M						
Bias	0.0081	2.9668	0.0097	2.9969	0.0004	2.9969
MS	0.5611	8.8713	0.4251	8.9818	0.2827	8.9821
MM						
Bias	0.0082	0.0008	0.0123	0.2185	0.0289	2.7200
MS	0.0593	0.0739	0.0927	0.7732	0.2965	8.1722
TELBS						
Bias	0.0058	0.0031	0.0064	0.0024	0.0043	0.0099
MS	0.0649	0.0725	0.0723	0.0983	0.0710	0.0995
n = 100						
LS						
Bias	0.0033	2.9969	0.0010	2.9969	0.0025	2.9969
MS	0.1066	8.9814	0.0796	8.9819	0.0645	8.9819
M						
Bias	0.0205	2.9969	0.0051	2.9970	0.0004	2.9969
MS	0.0178	8.9816	0.0769	8.9820	0.0500	8.9819
MM						
Bias	0.0024	0.0022	0.0003	0.0030	0.0063	2.9267
MS	0.0110	0.0109	0.0144	0.0143	0.0512	8.7760
TELBS						
Bias	0.0011	0.0009	0.0042	0.0002	0.0007	0.0038
MS	0.0111	0.0142	0.0126	0.0151	0.0172	0.0165

Abbreviations: LS, least squares; MS, mean squares; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

Table 6 Bias and mean square errors with contamination in the response y direction

	5%		25%		40%	
	β_0	β_1	β_0	β_1	β_0	β_1
n = 20						
LS						
Bias	43.6473	154.4103	249.7034	749.7920	414.7960	1185.640
MS	28576.56	72966.19	156694.1	742500.3	300353.1	1643825
M						
Bias	0.0201	0.0777	15.5802	79.3839	263.8728	832.6838
MS	0.0745	0.0786	13866.84	95715.33	197651	1098011
MM						
Bias	0.0134	0.0109	0.0020	0.0115	0.0076	0.0016
MS	0.0728	0.0884	0.0920	0.1245	0.1026	0.1094
TELBS						
Bias	0.0068	0.0002	0.0113	0.0160	0.0015	0.0081
MS	0.0648	0.0862	0.0891	0.1135	0.0939	0.1319
n = 100						
LS						
Bias	45.2973	145.2830	252.3375	751.2341	405.7585	1199.109
MS	6506.701	30715.70	81952.11	602530.9	189135.5	1485956
M						
Bias	0.0251	0.0699	0.2283	0.7080	232.6861	707.8531
MS	0.0140	0.0171	0.1238	0.6627	80513.89	631891.0
MM						
Bias	0.0002	0.0014	0.0043	0.0019	0.0017	0.0045
MS	0.0113	0.0116	0.0132	0.0134	0.0174	0.0185
TELBS						
Bias	0.0034	0.0000	0.0006	0.0017	0.0006	0.0056
MS	0.0123	0.0140	0.0145	0.0179	0.0164	0.0192

Abbreviations: LS, least squares; MS, mean squares; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

Table 7 Bias and mean square errors with contamination in both x and y directions

	5%		25%		40%	
	β_0	β_1	β_0	β_1	β_0	β_1
n = 20						
LS						
Bias	0.2076	1.1573	212.1027	0.0091	363.1680	0.0125
MS	45.9164	382.2538	58563.03	0.6532	152871.5	0.2283
M						
Bias	00647	0.2297	0.5398	0.0174	74.1418	0.0398
MS	7.9941	111.9223	0.6401	1.0751	22616.97	0.4646
MM						
Bias	0.0092	0.0181	0.0115	0.0215	0.0025	0.0121
MS	0.0637	0.1116	0.0831	0.1997	0.1082	0.2319
TELBS						
Bias	0.0016	0.0002	0.0173	0.0005	0.0052	0.0089
MS	0.0553	0.0760	0.0756	0.0821	0.0805	0.1065
n = 100						
LS						
Bias	39.9438	0.0231	240.4222	0.0027	393.5501	0.0083
MS	2096.474	0.6392	60462.25	0.0761	159192	0.0372
M						
Bias	0.0560	0.0202	0.5216	0.0039	35.1305	0.0017
MS	0.0309	1.1145	0.3135	0.1665	5043.506	0.0874
MM						
Bias	0.0020	0.0042	0.0002	0.0026	0.0012	0.0004
MS	0.0116	0.0231	0.0144	0.0351	0.0167	0.0360
TELBS						
Bias	0.0034	0.0019	0.0021	0.0015	0.0010	0.0001
MS	0.0125	0.0130	0.0140	0.0144	0.0150	0.0170

Abbreviations: LS, least squares; MS, mean squares; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

Table 8 Bias and mean square errors with contamination in x_1, x_2 direction

	5%			25%			40%		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
n = 20									
LS									
Bias	0.0172	1.4867	0.4673	0.0476	2.9969	0.5000	0.0445	2.9969	0.5000
MS	0.5778	3.4604	1.5548	0.6964	8.9819	0.2500	0.6501	8.9818	0.2500
M									
Bias	0.0093	1.5391	0.5099	0.0168	2.9969	0.4999	0.0213	2.9969	0.4999
MS	0.6182	3.5974	1.6077	0.7245	8.9819	0.2499	0.5762	8.9819	0.2499
MM									
Bias	0.0088	0.0208	0.5131	0.0135	0.1752	0.5144	0.0691	2.3432	0.4878
MS	0.3175	0.0832	0.3668	0.3678	0.5797	0.4602	0.5557	7.0571	0.2887
TELBS									
Bias	0.0312	0.0020	0.0048	0.0045	0.0021	0.0037	0.0047	0.0093	0.0007
MS	0.0671	0.0816	0.0839	0.0774	0.0870	0.0907	0.0965	0.1071	0.1196
n = 100									
LS									
Bias	0.0212	2.9968	0.4999	0.0196	2.9969	0.5000	0.0274	2.9969	0.5000
MS	0.3444	8.9812	0.2499	0.3274	8.9819	0.2500	0.3211	8.9819	0.2500
M									
Bias	0.0075	2.9968	0.5000	0.0064	2.9969	0.5000	0.0218	2.9969	0.5000
MS	0.3553	8.9813	0.2500	0.3529	8.9819	0.2500	0.2961	8.9819	0.2500
MM									
Bias	0.0007	0.0057	0.4986	0.0264	0.0179	0.4972	0.0275	2.8732	0.5004
MS	0.2588	0.0117	0.2686	0.2613	0.0150	0.2802	0.2999	8.6144	0.2522
TELBS									
Bias	0.0009	0.0017	0.0033	0.0033	0.0039	0.0007	0.0026	0.0003	0.0053
MS	0.0121	0.0122	0.0123	0.0151	0.0156	0.0141	0.0160	0.0164	0.0177

Abbreviations: LS, least squares; MS, mean squares; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

Table 9 Bias and mean square errors with contamination in y direction

	5%			25%			40%		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
n = 20									
LS									
Bias	47.5324	150.7233	7.9741	237.2123	757.2887	15.6698	423.8958	1231.093	13.3813
MS	27539.16	75026.31	30514.76	174783.8	763935.4	136845.6	370305	1801726	178241.1
M									
Bias	0.0221	0.0661	0.5003	27.9257	136.7052	15.0851	304.2819	942.7024	23.5915
MS	0.3266	0.0848	0.3317	25830.51	146578.4	36127.98	271497.9	1325407	146333.8
MM									
Bias	0.0035	0.0085	0.5037	0.0157	0.0201	0.5174	0.0036	0.0066	0.5025
MS	0.3248	0.0726	0.3193	0.3140	0.0938	0.3632	0.3781	0.1139	0.4065
TELBS									
Bias	0.0080	0.0011	0.0138	0.0014	0.0000	0.0172	0.0014	0.0000	0.0172
MS	0.0688	0.0939	0.0850	0.0800	0.1001	0.1114	0.0800	0.1005	0.1114
n = 100									
LS									
Bias	48.1726	147.6965	0.3145	245.1727	746.1551	6.7841	405.2485	1179.175	7.9694
MS	7693.723	30459.01	5102.882	96704.02	594204.2	22831.46	227561.4	1440607	30060.28
M									
Bias	0.0489	0.0630	0.5002	0.2612	0.7666	0.5475	236.0896	725.6934	1.4953
MS	0.2889	0.0158	0.2633	0.5174	2.9772	1.7309	100135.2	654402.2	18713.87
MM									
Bias	0.0056	0.0026	0.5005	0.0284	0.0014	0.5066	0.0041	0.0033	0.5060
MS	0.2736	0.0122	0.2628	0.2756	0.0155	0.2709	0.2572	0.0153	0.2731
TELBS									
Bias	0.0019	0.0053	0.0021	0.0031	0.0016	0.0022	0.0019	0.0014	0.0014
MS	0.0126	0.0134	0.0132	0.0157	0.0160	0.0159	0.0126	0.0150	0.0134

Abbreviations: LS, least squares; MS, mean squares; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

Table 10 Bias and mean square errors with contamination in both x_1, x_2 and y direction

	5%			25%			40%		
	β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
n = 20									
LS									
Bias	0.0647	0.1164	0.5577	157.9964	0.0587	0.5099	328.3148	0.0166	0.5061
MS	0.5506	5.3699	3.8677	46224.97	0.8650	1.0943	159787.1	0.3129	0.5967
M									
Bias	0.0066	0.0507	0.4944	0.5344	0.0284	0.5113	48.1713	0.0161	0.5265
MS	0.6885	2.4124	2.8125	1.4986	1.7021	1.8582	18968.09	0.6473	0.9399
MM									
Bias	0.0018	0.0014	0.4827	0.0364	0.0182	0.5090	0.0135	0.0108	0.4768
MS	0.3186	0.1018	0.3426	0.3503	0.2253	0.4639	0.3509	0.2846	0.4742
TELBS									
Bias	0.0033	0.0053	0.0164	0.0139	0.0120	0.0111	0.0217	0.0058	0.0026
MS	0.0688	0.0745	0.0758	0.0758	0.0889	0.0834	0.0914	0.1086	0.1086
n = 100									
LS									
Bias	30.1399	0.0031	0.4799	240.3774	0.0108	0.4982	391.0141	0.0014	0.5022
MS	1639.960	0.9444	1.3195	73844.49	0.0782	0.3289	193011.8	0.0416	0.2949
M									
Bias	0.0465	0.0088	0.5183	0.4865	0.0243	0.4845	71.7581	2.6559	0.4951
MS	0.3300	1.6034	1.7357	0.8035	0.1945	0.4367	12853.47	7.2023	0.2693
MM									
Bias	0.0168	0.0002	0.5069	0.0018	0.0029	0.4954	0.0035	0.0014	0.4979
MS	0.2606	0.0174	0.2755	0.2522	0.0352	0.2800	0.2711	0.0452	0.2935
TELBS									
Bias	0.0024	0.0075	0.0054	0.0033	0.0036	0.0057	0.0064	0.0001	0.0009
MS	0.0122	0.0124	0.0119	0.1391	0.0155	0.0144	0.0157	0.0162	0.0169

Abbreviations: LS, least squares; MS, mean squares; TELBS, Tabatabai, Eby, Li, Bae, and Singh.

robust measures of goodness of fit and robust method of variable selection. In addition, robust testing of the hypothesis concerning model parameters were introduced and a new diagnostic statistic for identifying influential observations was introduced. Moreover, an algorithm was developed to perform the TELBS estimation of model parameters, and in the Appendix we have provided computer programs using both R and Mathematica software. Either program would perform the calculations necessary for regression analysis of a data set by minor adjustments for the number of explanatory variables. The TELBS method was also applied to two sets of real data and a computer simulation revealed the robustness of this new estimator. For the rat liver data set, the TELBS method was the only method producing the correct conclusion. In linear regression it is of utmost significance to identify outliers, leverage points, and influential observations. These observations may provide clues about the model and the research under study. They may also reveal some interesting aspects of the data and the cause of such behavior. The TELBS estimator of robust linear regression can bound the influence of such observations.

Robust techniques have not been widely used by biomedical researchers mainly because of their computational complexity. The TELBS method provides a simple but powerful alternative to such a problem. The TELBS estimator of linear regression parameters has influence functions bounded in both the explanatory variables and the response variable direction. It has high asymptotic breakdown point and efficiency. The simulation results clearly indicates that the TELBS method is worthy of consideration by scientists and medical practitioners when it comes to linear regression.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Rousseeuw PJ, Leroy AM. *Robust Regression and Outlier Detection*. New York, NY: Wiley Interscience; 1987.
2. Edgeworth FY. On observations relating to several quantities. *Hermathena*. 1887;6:279–285.
3. Huber PJ. Robust regression: asymptotics, conjectures, and Monte Carlo. *Ann Stat*. 1973;1:799–821.
4. Rousseeuw PJ. Least median of squares regression. *J Am Stat Assoc*. 1984;79:871–880.

5. Rousseeuw PJ, Yohai VJ. Robust regression by means of S estimators. In: Franke J, Härdle W, Martin RD, editors. *Robust and Non-Linear Time Series Analysis*. New York, NY: Springer-Verlag; 1984.

6. Yohai VJ, Zamar RH. High breakdown point estimates of regression by means of the minimization of an efficient scale. *J Am Stat Assoc*. 1988;83:406–413.

7. Tabatabai MA, Argyros IK. Robust estimation and testing for general nonlinear regression models. *Appl Math Comput*. 1993;58:85–101.

8. Fasano MV. Robust estimation for linear regression models. PhD thesis. La Plata, Argentina: Department of Mathematics, University of La Plata; 2009.

9. Yohai VJ. High breakdown point and high efficiency. *Ann Stat*. 1987;15:642–656.

10. Yang X, Beason-Held L, Resnick SM, Landman BA. Robust biological parametric mapping: an improved technique for multimodal brain image analysis. *Proc Soc Photo Opt Instrum Eng*. 2011;7962:79623X.

11. Yeo GW, Xu X, Liang TY, et al. Alternative splicing events identified in human embryonic stem cells and neural progenitors. *PLoS Comput Biol*. 2007;3:1951–1967.

12. Xu C, Liu Y, Wang P, et al. Integrative analysis of DNA copy number and gene expression in metastatic oral squamous cell carcinoma identifies genes associated with poor survival. *Mol Cancer*. 2010;9:143.

13. Stewart CV. Robust parameter estimation in computer vision. *SIAM Review*. 1999;41:513–537.

14. Zaharescu A, Horaud R. Robust factorization methods using a Gaussian/uniform mixture model. *Int J Comput Vis*. 2009;81:240–257.

15. Ahdesmäki M, Lähdesmäki H, Gracey A, Shumelvich I, Yli-Harja O. Robust regression for periodicity detection in non-uniformly sampled time-course gene expression data. *BMC Bioinformatics*. 2007;8:233.

16. Karan A, Campbell D, Mayer H. The effect of a visual aid on the comprehension of cataract surgery in a rural, indigent South Indian population. *Digital Journal of Ophthalmology*. 2011;17(3):16–22.

17. van Vliet P, Westendorp RP, Eidelenboom P, et al. Parental history of Alzheimer disease associated with lower plasma apolipoprotein E levels. *Neurology*. 2009;73:681–687.

18. Pinheiro SP, Holmes MD, Pollak MN, Barbieri RL, Hankinson SE. Racial differences in premenopausal endogenous hormones. *Cancer Epidemiol Biomarkers Prev*. 2006;14:2147–2153.

19. Floyd TF, McGarvey M, Ochroch EA, et al. Perioperative changes in cerebral blood flow after cardiac surgery: influence of anemia and aging. *Ann Thorac Surg*. 2003;76:2037–2042.

20. Mircean C, Shmulevich I, Cogdell D, et al. Robust estimation of protein expression ratios with lysate microarray. *Bioinformatics*. 2005;21:1935–1942.

21. Kocak E, Carruthers KH, McMahan JD. A reliable method for the preoperative estimation of tissue to be removed during reduction mammoplasty. *Plast Reconstr Surg*. 2011;127:1059–1064.

22. Cepeda MS, Carr DB. Women experience more pain and require more morphine than men to achieve a similar degree of analgesia. *Anesth Analg*. 2003;97:1464–1468.

23. Rousseeuw PJ, Croux C. Alternatives to the median absolute deviation. *J Am Stat Assoc*. 1993;88:1273–1283.

24. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19:716–723.

25. Ronchetti E. Robust model selection in regression. *Stat Probab Lett*. 1985;3:21–23.

26. Maronna RA, Martin RD, Yohai VJ. *Robust Statistics: Theory and Methods*. New York, NY: Wiley; 2006.

27. Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA. *Robust Statistics: The Approach Based on Influence Functions*. New York, NY: Wiley; 1986.

28. Warmuth C, Günther M, Zimmer C. Quantification of blood flow in brain tumors: comparison of arterial spin labeling and dynamic susceptibility-weighted contrast-enhanced MR imaging. *Radiology*. 2003;228:523–532.

29. Weisberg S. *Applied Linear Regression*. 2nd ed. New York, NY: Wiley; 1985.

Appendix: Mathematica and R programs

R program for rat liver data:

"Tlmrobrat1.R"

```
lmauto2=function(m){
data=read.table("rat.txt")
code1=0; sig=1;
cs=0.405; y=data[,4]
x1=data[,1]; x2=data[,2];
x3=data[,3]
n=dim(data)[1]
int=array(0,c(m,1)); se=array(0,c(m,1));
slope1=array(0,c(m,1)); slope2=array(0,c(m,1))
slope3=array(0,c(m,1)); des=array(1,c(n,1))
mr=array(0,c(n,1)); ma1=array(0,c(n,1))
x=cbind(x1,x2,x3)
a=abs(x1); b=abs(x2)
c=abs(x3)
wd=cbind(des,x)
w=solve(t(wd)%*%wd)
diff1=10; diff2=10;
diff3=10; diff4=10;
j=1;
while(abs(diff1).0.001 | abs(diff2)>0.001 |abs(diff3)>0.001|abs(diff4)>0.001 )
{
j=j+1
lof=function(par){
stdev=sig
prob=array(0,c(n,1))
num=array(0,c(n,1))
sm=wd%*%solve(t(wd)%*%wd)%*%t(wd)
mf=diag(diag(sm))%*%des
  for (i in 1:n){
  prob[i]=par[1]+par[2]*x1[i]+par[3]*x2[i]+par[4]*x3[i]
  ma1[i]=max(median(a), abs(x1[i]))+max(median(b), abs(x2[i]))
    +max(median(c), abs(x3[i]))
  }
num=1-mf
losf=sum( (1-1/cosh(cs*(1/stdev)*(y-prob)*num)) /ma 1)
return(losf)
}
result=optim(par=c(0.2,-0.02,0.01,4),fn=lof)
a=result$par[1]
b=result$par[2]
c=result$par[3]; d=result$par[4];
r=y-(a+b*x1+c*x2+d*x3) ## r is a vector of residuals
for (k in 1:n){
mr[k]=median(abs(r[k]-r))
}
```

```
  sig=1.1926*median(mr)
  code=result$convergence
  code1=code1+code
  int[j]=result$par[1]
  slope1[j]=result$par[2]
  slope2[j]=result$par[3]
  slope3[j]=result$par[4]
  se[j]=sig
  diff1=int[j]-int[j-1]
  diff2=slope1[j]-slope1[j-1]
  diff3=slope2[j]-slope2[j-1]
  diff4=slope3[j]-slope3[j-1]
  #print(diff2)
  }
 out=array(0,c(5,1))
 out[1]=int[j]
 out[2]=slope1[j]
 out[3]=slope2[j]
 out[4]=slope3[j]
 out[5]=se[j]
 out[6]=median(abs(r))
 return(out)
}
a=lmauto2(100)[1]
b=lmauto2(100)[2]
c=lmauto2(100)[3]
d=lmauto2(100)[4]
sighat=lmauto2(100)[5]
ma=lmauto2(100)[6]
##------------------------------------------------------------------
## apply the robust linear method to rat data with three covariates
## par is initial value for optimization
## cs: tuning constant; cs=0.405
##------------------------------------------------------------------
```

source("Tlmrobrat1.R")

```
ratcov2=function(cs){
data=read.table("rat.txt")
n=dim(data)[1]
x1=data[,1];x2=data[,2]
x3=data[,3]
y=data[,4]
x=cbind(x1,x2,x3)
des=array(1,c(n,1));
mr=array(1,c(n,1)) ; num0=array(0,c(n,1))
```

```
wd=cbind(des,x);
sm=wd%*%solve(t(wd)%*%wd)%*%t(wd)
mf=diag(diag(sm))%*%des
s1=array(0,c(n,1))
s2=array(0,c(n,1))
p=dim(data)[2]
num=1-mf;
for (i in 1:n){
num0[i]=1-1/n}
##compute R-square
 for (k in 1: n){
 mr[k]=median(abs(y[k]-y))
 }
 rsq=1-(ma/median(mr))^2
 ## Estimate the variance covariance matrix
 g1=function(t){
 g=cs*1/cosh(cs*t)*tanh(cs*t) ##first derivative
 return(g)
 }
 g2=function(t){
 g=cs^2*(1/cosh(cs*t))^3-cs^2*(1/cosh(cs*t))*(tanh(cs*t))^2
 return(g) ##second derivative
 }
 for(i in 1:n){
 c1=(y[i]-(a+b*x1[i]+c*x2[i]+d*x3[i]))*(1-mf[i])/sighat
 s1[i]=g1(c1)
 s2[i]=g2(c1)
 }
 wt=solve(t(wd)%*%wd)
 sq=(sum(s2))^2
 variance=sum(s1^2)/((n-p)*sq)
 ##covariance matrix
 cov=(sighat^2)*(n^2)*variance*wt
 ##Standard error
 sea=sqrt(diag(cov))[1]
 seb=sqrt(diag(cov))[2]
 sec=sqrt(diag(cov))[3]
 sed=sqrt(diag(cov))[4]
##compute t-value and Wald chi-square
 t1=a/sea; chsq1=t1^2;
 t2=b/seb; chsq2=t2^2;
 t3=c/sec; chsq3=t3^2;
 t4=d/sed; chsq4=t4^2;
 ##compute p-value
 pvalue1=1-pchisq(chsq1,1)
 pvalue2=1-pchisq(chsq2,1)
 pvalue3=1-pchisq(chsq3,1)
 pvalue4=1-pchisq(chsq4,1)
```

```
## Output parameter estimate, S.E., t-value, variance-covariance matrix,
## Wald chi-square, and p-value for each parameter
 cat("a:", a,"\n");
 cat("b:", b,"\n");
 cat("c:", c,"\n");
 cat("d:", d,"\n");
 cat("SE (a):", sea,"\n");
 cat("SE (b):", seb,"\n");
 cat("SE (c):", sec,"\n");
 cat("SE (d):", sed,"\n");
 cat("t value (a):", a/sea,"\n");
 cat("t value (b):", b/seb,"\n");
 cat("t value (c):", c/sec,"\n");
 cat("t value (d):", d/sed,"\n");
 cat("Var-Cov matrix:", cov,"\n");
 cat("Wald (a):", chsq1,"\n");
 cat("Wald (b):", chsq2,"\n");
 cat("Wald (c):", chsq3,"\n");
 cat("Wald (d):", chsq4,"\n");
 cat("p-value (a):", pvalue1,"\n");
 cat("p-value (b):", pvalue2, "\n");
 cat("p-value (c):", pvalue3, "\n");
 cat("p-value (d):", pvalue4, "\n");
 cat("Rsquare:", rsq,"\n");
 cat("Sigma hat:", sighat,"\n");
}
##----------------------------------------------------
## Robust linear method for rat data
## cs: tuning constant; cs=0.405
##----------------------------------------------------

OUTPUT

a: 0.3009474
b: -0.007261964
c: 0.001258697
d: 1.467131
SE (a): 0.1602341
SE (b): 0.006566557
SE (c): 0.01417771
SE (d): 1.253828
t value (a): 1.878173
t value (b): -1.105901
t value (c): 0.08878001
t value (d): 1.170121
Var-Cov matrix: 0.02567496 -0.0003014961 -0.0002754827 0.03294774 -0.0003014961
4.311967e-05
```

```
-1.127635e-05  -0.0081273  -0.0002754827  -1.127635e-05  0.0002010074  0.0007420215
0.03294774
-0.0081273 0.0007420215 1.572085
Wald (a): 3.527536
Wald (b): 1.223018
Wald (c): 0.00788189
Wald (d): 1.369184
p-value (a): 0.06035744
p-value (b): 0.2687692
p-value (c): 0.9292567
p-value (d): 0.2419522
Rsquare: 0.3162099
Sigma hat: 0.08572506
```

Mathematica program for TELBS robust liner regression

```
Data from 2003 Neuradiology:
SiPrime[t_] := ω² * Sech[ω * t]³ − ω² Sech[ω * t] Tanh[ω * t]²; Rho[t_] := 1 − Sech[ω * t];
Si[t_] := ω * Sech[ω * t] Tanh[ω * t]; v[0] = {0, 0}; p = 2; ω = .525; sighat = 1; n = 32;
y = {.33, .66, .59, .65, .85, .66, .65, .75, .67, 1.22, .97, 1.18, 1.36, 2.8,
    1.23, 2.23, 1.35, .41, .47, .57, .80, .39, .65, .53, .69, .74, 1.85, .98,
    .88, 3.74, .87, .52};
x = {.19, .67, .51, .42, .65, .57, .73, 1.03, .65, 1.25, .76, 1.05, 1.16, 1.66,
    1.36, 1.41, 1.64, .36, .35, .41, 1.27, .57, .32, .53, .76, .47, 1.66, .49,
    .66, 1.74, .66, .28};
WD1 = Table [{1}, {i, 1, n }]; L = Table [Max[Median [Abs[x]], Abs[x[i]]]]. {i, n}];
S1 = Diagonal [WD1.Inverse[Transpose [WD1].WD1].Transpose [WD1]];
WD = Table [{1, x[i]]}, {i , 1, n}];
h = Diagonal[WD.Inverse[Transpose[WD].WD].Transpose[WD]];
```

$$f = \sum_{i=1}^{n}\left(1 - \text{Sech}\left[\omega\left(1 - h[[i]]\right)\left(\frac{y[[i]] - \left(a1 + b1 * x[[i]]\right)}{sighat}\right)\right]\right) \Big/ L[[i]];$$

```
{a, b} = NArgMin[f, {a1, b1}];
v[1] = {a, b};
j = 1; While [Norm[v[j] − v[j−1]] > 0.0000001, r = y −a −b * x;
  sighat = 1.1926 * Median [Table[Median[Abs[r[1]] − r]], {1, n}]];
  g = Table[Si[1-h[[i]])(r[[i]]/sighat)]/(sighat*L[[i]] r[[i]]), {i, n}];
  w = DiagonalMatrix [g];
  v[j + 1] = Inverse[Transpose[WD].w.WD].Transpose[WD].w.y; (a, b) = v[j + 1];
```

$$j{+}{+}; \quad f = \sum_{i=1}^{n}\left(1 - \text{Sech}\left[\omega\left(1 - h[[i]]\right)\left(\frac{r[[i]]}{sight}\right)\right]\right) \Big/ L[[i]];$$

```
Print["TELS parameter estimates at iteration", j]
Print["(a,b) = ", v[j]]
```

$$\text{Print}\left["R^2 =", 1 - \left(\frac{\text{Median}[Abs[r]]}{\text{Median}\left[\text{Table}\left[\text{Median}\left[Abs[y - y[[1]]]\right], (1, n)\right]\right]}\right)^{\wedge}2\right]$$
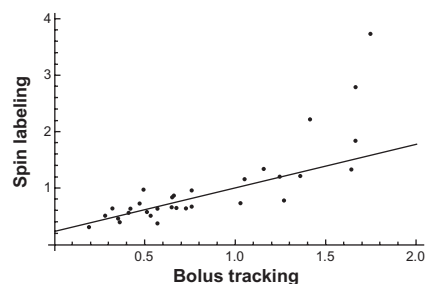
$$\text{"If no intercept presents, use } R^2 = 1 - \left(\frac{\text{Median}[Abs[r]]}{\text{Median}[Abs[Y]]}\right)^{\wedge}2";$$
```
{a, b} = v[j];
```

$$\texttt{Covar} = \texttt{Sighat\^{}2*}\left(\texttt{n\^{}2}\sum_{i=1}^{n}\left(\texttt{Si}\left[\left(\left(\left(\texttt{y[[i]]-Y[[i]]-}\left(\texttt{a+b*x[[i]]}\right)\right)\right)\right)\texttt{/sighat}\right]\right)^{\wedge}2\right)$$

$$/\left(\texttt{(n-p)}\left(\sum_{i=1}^{n}\texttt{SiPrime}\left[\left(\left(\left(\texttt{y[i]-}\left(\texttt{a+b*x[[i]]}\right)\right)\right)\right)\texttt{/sighat}\right)^{\wedge}2\right)\right)_*$$

```
Inverse[Transpose[WD].WD];
```

$$\texttt{11} = \left(\texttt{v[j]}/\sqrt{\texttt{Digonal[Covar]}}\right)^{\wedge}2;$$

$$\texttt{Print}\left[\texttt{"Variance-Covariance Matrix ="}, \texttt{MatrixForm[Covar]}\right]$$

$$\texttt{Print}\left[\texttt{"Standard Error for Parameters = "}, \sqrt{\texttt{Diagonal[Covar]}}\right]$$

$$\texttt{Print}\left[\texttt{"t-values for parameters = "}, \texttt{v[j]}/\sqrt{\texttt{Diagonal[Covar]}}\right]$$

```
Print ["Chi-square-values for parameters = ", (v[j]^2/Diagonal[Covar])]
Needs ["HypothesisTesting'"]
{ChiSquarePValue[11{[1]], 1}, ChiSquarePValue [11[[12]], 1]}
Print ["f = ", f]
```

$$\texttt{Print}\left[\texttt{"Deviance = "}, 2\,\texttt{sighat}\,\^{}2 * \sum_{i=1}^{n}\left(1-\texttt{Sech}\left[\omega\left(1-\texttt{h[[i]]}\left(\frac{\texttt{r[[i]]}}{\texttt{sighat}}\right)\right)\right]\right)/\texttt{L[[i]]}\right]$$

$$\texttt{Print}\left[\texttt{"AICR = "},\left(2*\sum_{i=1}^{n}\left(1-\texttt{Sech}\left[\omega\left(1-\texttt{h[[i]]}\right)\left(\frac{\texttt{r[[i]]}}{\texttt{sighat}}\right)\right]\right)/\texttt{L[[i]]}\right)+\right.$$

$$\left.2*\texttt{P*NIntegrate}\left[\frac{1}{\sqrt{2\text{\AA}}}\left(\texttt{Rho'[m]}\right)\^{}2\,e^{-0.5'm^2},\{m,-\infty,\infty\}\right]/\right.$$

$$\left.\texttt{NIntegrate}\left[\frac{(\texttt{Rho''[m]})\,e^{-0.5'm^2}}{\sqrt{2\text{\AA}}},\{m,-\infty,\infty\}\right]\right]$$

$$\texttt{Print}\left[\texttt{"BICR = "},\left(2*\sum_{i=1}^{n}\left(1-\texttt{Sech}[\omega(1-\texttt{h[[i]]})\left(\frac{\texttt{r[[i]]}}{\texttt{sighat}}\right]\right)/\texttt{L[[i]]}\right)+\texttt{P*Log[n]}\right]$$

```
Show [ListPlot[Table[{{x[[i]], y [i]]}}, {i, 1, n}]], Plot [(a + b * m), {m, 0, 2}],
  PlotRange → {0, 4}, AxesLabel → {"Bolus Tracking", "Spin Labeling"}]
TELS parameter estimates at iteration 17
{a, b} = {0.246233, 0.775229}
```

$R^2 = 0.805433$

$$\texttt{Variance-Covariance Matrix} = \begin{pmatrix} 0.00616112 & -0.00575966 \\ -0.00575966 & 0.00702397 \end{pmatrix}$$

```
Standard Error for Parameters = {0.0784928, 0.0838091}
t-values for parameters = {3.13701, 9.24994}
```

$\{\texttt{OneSidedPValue} \to 0.00170679, \texttt{OneSidedPValue} \to 2.24624 \times 10^{-20}\}$

```
f = 3.71088
Deviance = 0.367665
AICR = 8.16658
BICR = 14.3532
```