INSTITUTE OF ECONOMIC STUDIES, FACULTY OF SOCIAL SCIENCES
CHARLES UNIVERSITY IN PRAGUE (*established 1348*)

ROBUST STATISTICS AND ECONOMETRICS

INSTITUTE OF ECONOMIC STUDIES
FACULTY OF SOCIAL SCIENCES
CHARLES UNIVERSITY IN PRAGUE

JAN ÁMOS VÍŠEK

Week 8

Content of lecture

1. At the beginning of any lecture let us repeat .....
   - Algorithm for *M*-estimators
   - Rousseeuw proposals of estimators and their algorithms
2. Our algorithms
   - Boček-Lachout algorithm for LMS and its comparison with exact LTS
   - Algorithm for LTS
   - Diagnostics by robust methods with high breakdown point
   - Algorithm for LWS

Computing M-estimate of regression coefficients

We have considered the extremal problem

$$\hat{\beta}^{(GM,n,w)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{n} w_i \rho\left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right).$$

Write it as

$$\hat{\beta}^{(M,n)} = \arg\min_{\beta \in R^p} \sum_{i:(Y_i - X_i'\beta)\neq 0} w_i \left[\rho\left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right) \cdot \left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right)^{-2}\right] \left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right)^2$$

$$= \arg\min_{\beta \in R^p} \sum_{i=1}^{n} \tilde{w}_i \cdot \left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right)^2$$

where $\tilde{w}_i = w_i \rho\left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right) \cdot \left(\frac{Y_i - X_i'\beta}{\hat{\sigma}}\right)^{-2}$ for $i : (Y_i - X_i'\beta) \neq 0$,

otherwise $\tilde{w}_i = 0$.

Computing M-estimate of regression coefficients

Then

$$\hat{\beta}^{(GM,n,w)} = \left(X'\tilde{W}X \right)^{-1} X'\tilde{W}Y$$

where $\tilde{W} = \mathrm{diag}(\tilde{w}_1, \tilde{w}_2, ..., \tilde{w}_n)$.

And an iterative computation, starting with a "guess" of

$$\hat{\beta}^{(GM,n,w)}_{(starting)},$$

lead usually after several tens or hundreds of cycles to the desired estimate.

Antoch, J., J. Á. Víšek (1991):
Robust estimation in linear models and its computational aspects.
Contributions to Statistics: Computational Aspects of Model Choice, Springer Verlag, (1992), ed. J. Antoch, 39 - 104.

A pursuit for highly robust estimator of regression coefficients

Prior to continuing let us make an agreement:

For any $\beta \in R^p$

$$r_i(\beta) = Y_i - X_i'\beta \qquad \text{not only} \qquad \cancel{r_i(\hat{\beta}) = Y_i - X_i'\hat{\beta}}$$

Order statistics

$$r^2_{(1)}(\beta), r^2_{(2)}(\beta), \ldots, r^2_{(n)}(\beta),$$

some texts alternatively employ

$$r^2_{(1:n)}(\beta), r^2_{(2:n)}(\beta), \ldots, r^2_{(n:n)}(\beta).$$

A pursuit for highly robust estimator of regression coefficients

Regression quantiles

Koenker,R., G. Bassett (1978): Regression quantiles.
Econometrica, 46, 33-50.

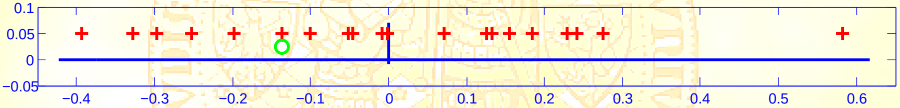
$$\hat{\beta}^{(\alpha)} = \arg\min_{\beta \in R^p} \left\{ \sum_{i=1}^{n} \left[\alpha \cdot |r_i(\beta)| \cdot I\{r_i(\beta) < 0\} + (1-\alpha) \cdot |r_i(\beta)| \cdot I\{r_i(\beta) > 0\} \right] \right\}$$

$$\hat{\beta}^{(L,n)} = \sum_{\ell=1}^{K} c_\ell \hat{\beta}^{(\alpha_\ell)}$$

$\hat{\beta}^{(\alpha)}$ is *M*- and simultaneously *L*-estimator

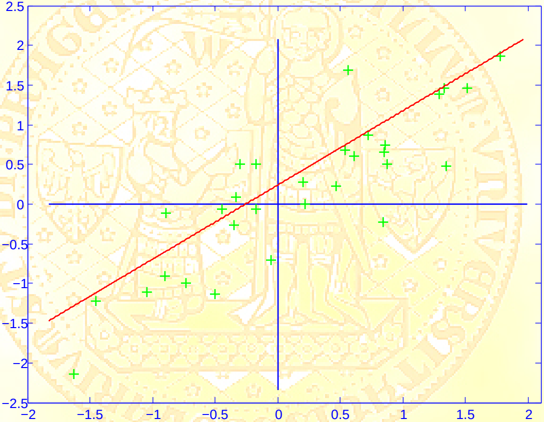
Classical quantiles

30% location quantile



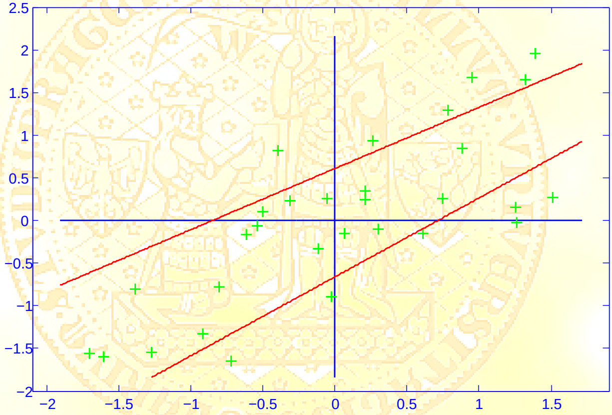
Regression quantiles

20% regression quantile



Regression quantiles

Two regression quantiles, 20% and 89%, say



A pursuit for highly robust estimator of regression coefficients

The trimmed least squares (TLS)

Ruppert, D., R. J. Carroll (1980):

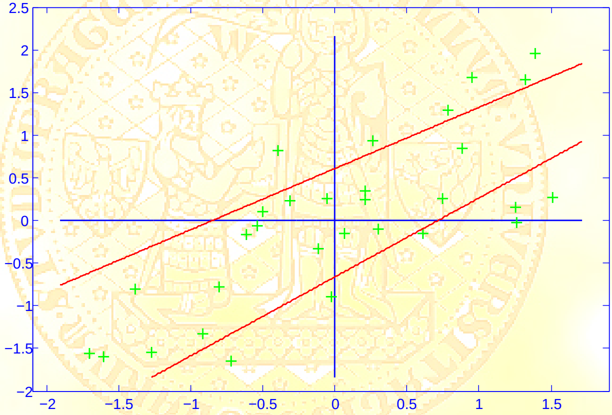
Trimmed least squares estimation in linear model.

J. Americal Statist. Ass., 75 (372), 828–838.

Trimming by $\left[x' \cdot \hat{\beta}^{(\alpha_1)}, x' \cdot \hat{\beta}^{(\alpha_2)}\right] \quad 0 \le \alpha_1 < \alpha_2 \le 1 \quad \rightarrow \quad \hat{\beta}^{(TLS,n)_{(\alpha_1,\alpha_2)}}$

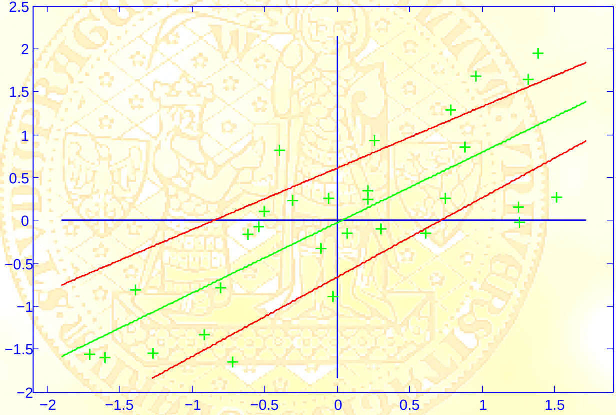
The trimmed least squares

Two regression quantiles



The trimmed least squares

Two regression quantiles with OLS for trimmed data



We have studied LMS

Rousseeuw, P. J. (1983): Least median of square regression.
Journal of Amer. Statist. Association 79, pp. 871-880.

the Least Median of Squares

$$\hat{\beta}^{(LMS,n,h)} = \arg\min_{\beta \in R^p} r^2_{(h)}(\beta) \qquad \frac{n}{2} < h \leq n,$$

Many advantages - mainly

1. *breakdown point equal to* $([\frac{n-p}{2}] + 1)n^{-1}$ *if* $h = [\frac{n}{2}] + [\frac{p+1}{2}]$
2. *scale- and regression equivariant*

(without any studentization of residuals).

Main disadvantage

$$\sqrt[3]{n}\left(\hat{\beta}^{(LMS,n,h)} - \beta^0\right) = \mathcal{O}_p(1)$$

(Cernobyl)

Peter Rousseeuw proposed the algorithm:

1. Select randomly an elemental set of p points and fit a regression plane to them.
2. Compute all squared residuals and find the h-th smallest.
3. Repeat it “10 000” times and select that model (among these “10 000”) with smallest h-th squared residual.

An improvement of the algorithm - a geometric characterization

Joss, J., A. Marazzi (1990):
Probabilistic algorithms for LMS regression.
Computational Statistics & Data Analysis 9, 123-134.

The geometric characterization
of exact solution of LMS extremal problem:

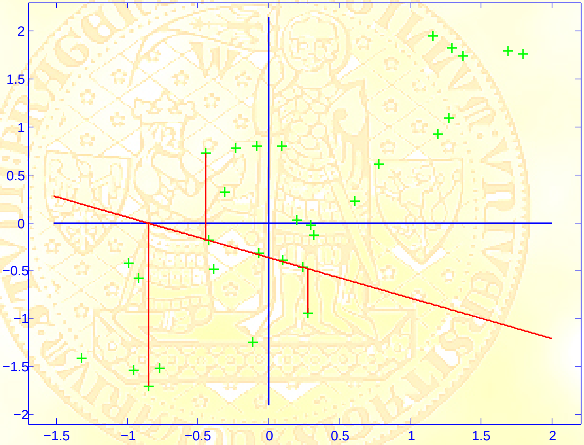
The exact solution has at least
$p+1$ residuals of the same (absolute) value.

An improvement of the algorithm - a geometric characterization

1. Select randomly an elemental set of p points and fit a regression plane to them.
2. Perform (repeatedly) its shift and rotation to decrease the value of the h-th squared residual and to reach the geometric representation.
3. Repeat it “10 000” times and select that model (among these “10 000”) with smallest h-th squared residual.

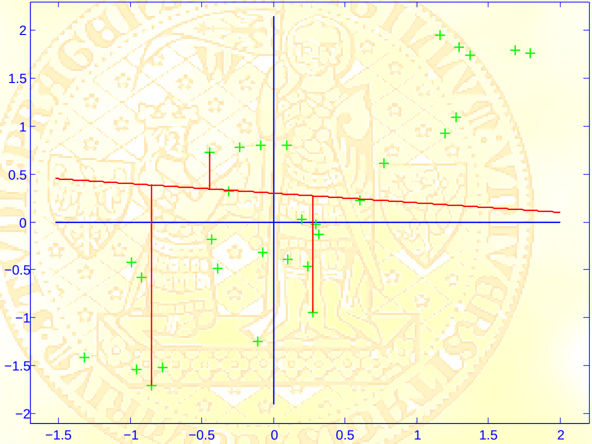
A geometric characterization

Unlucky selection of starting points



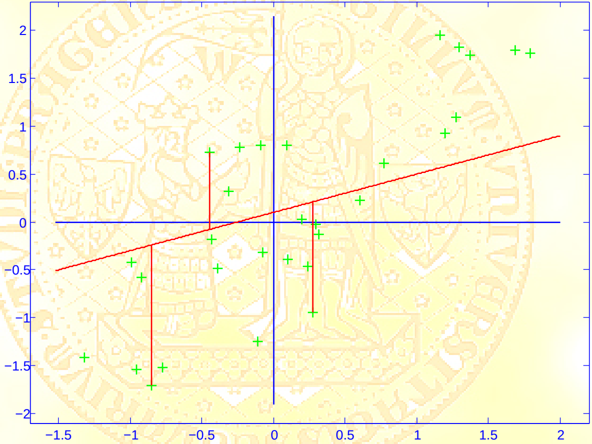
A geometric characterization

Starting shifting and spinning the line



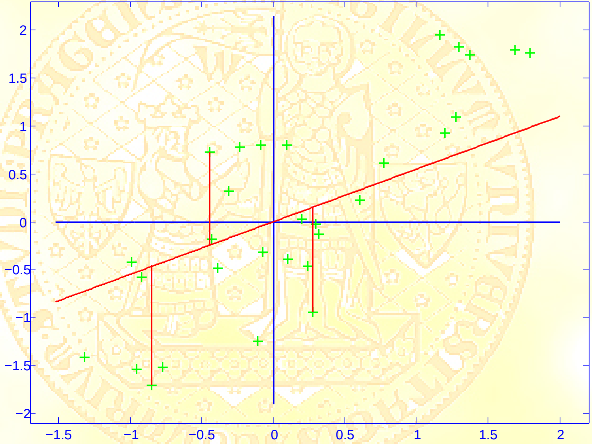
A geometric characterization

Continuing shifting and spinning the line



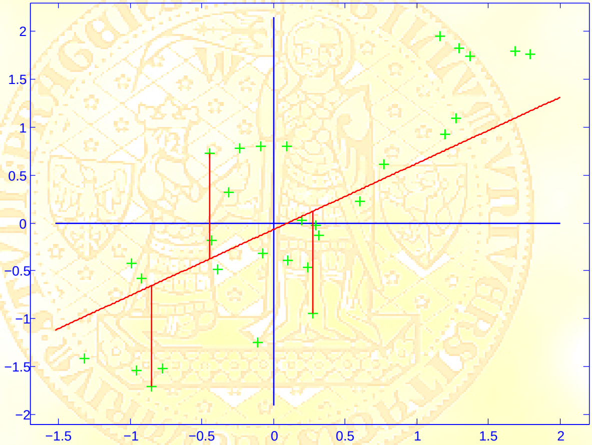
A geometric characterization

Nearly reaching the geometric characterization



A geometric characterization

Reaching the geometric characterization



A shock and frustration - Engine Knock Data

Hettmansperger, T. P., S. J. Sheather (1992):
A Cautionary Note on the Method of Least Median Squares.
The American Statistician 46, 79–83.

Engine Knock Data ($n = 16$, $p = 4$, $h = 11$)

c	x_1	x_2	x_3	x_4	y
1	13.3	13.9	31	697	84.4
2	13.3	14.1	30	697	84.1
3	13.4	15.2 15.1	32	700	88.4
4	12.7	13.8	31	669	84.2
16	12.7	15.9	37	696	93.1

In fact they worked with two data sets.

Let's call these data "Correct".
Let's call these data "Damaged".

x_1 is spark timing x_2 air/fuel ratio
x_3 intake temperature x_4 exhaust temperature
y engine knock number

A shock and frustration

Hettmansperger, T. P., S. J. Sheather (1992):
A Cautionary Note on the Method of Least Median Squares.
The American Statistician 46, 79–83.

Engine Knock Data $(n = 16, p = 4, h = 11)$

The values of $\hat{\beta}^{(LMS,n,h)}$ by "elemental" algorithm !
(still included in some packages - see the next slide)

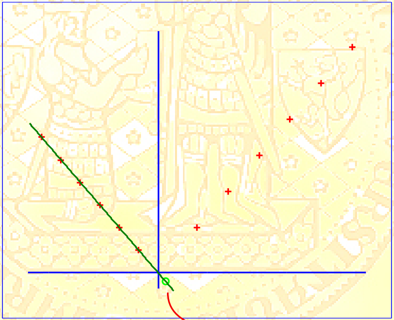
Data	Interc.	SPARK	AIR	INTK	EXHS.
Correct data ($x_{22} = 14.1$)	30.08	0.21	2.90	0.56	-0.01
Damaged data($x_{22} = 15.1$)	-86.50	4.59	1.21	1.47	0.07

x_1 is spark timing x_2 air/fuel ratio
x_3 intake temperature x_4 exhaust temperature
y engine knock number

An (academic) explanation by a shift of “inlier”

SENSITIVITY OF ANY HIGH-BREAKDOWN-POINT ESTIMATOR
TO A SMALL CHANGE OF DATA

Model for the majority of data

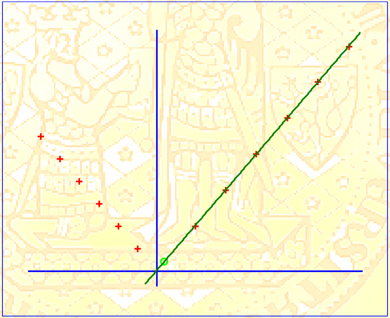


We are going to shift up this point “ o ”.

An (academic) explanation by a shift of “inlier”

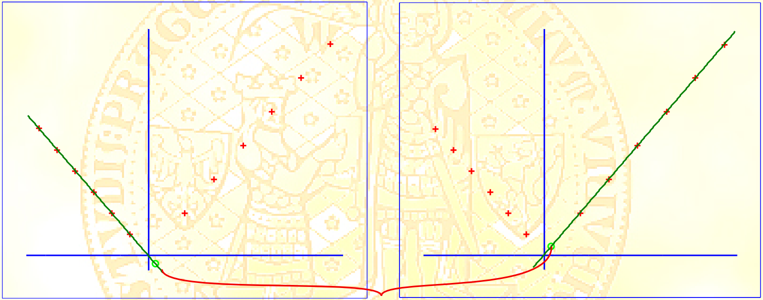
SENSITIVITY OF ANY HIGH-BREAKDOWN-POINT ESTIMATOR TO A SMALL CHANGE OF DATA

Again model for the majority of data



An (academic) explanation by a shift of “inlier”

In both cases the model is for the majority of data



Notice: *The closer* the point (“ o ”) is to the *y*-axe,
the smaller shift causes the “switch” of the model.

A substantial improvement of the algorithm

- an employment of simplex method

Boček, P., P. Lachout (1993):
Linear programming approach to LMS-estimation.
Memorial volume of Comput. Statist. & Data Analysis 19(1995), 129 - 134.

A description is a bit complicated - it requires
to be familiar with a dual form of simplex method.

Boček-Lachout algorithm

First of all, the algorithm gave:

1. much smaller 11^{th} squared residual
   than the algorithm used by Hettmansperger & Sheather,

	11^{th} order statistics	
Method	PRO-*LMS*	Bo-La-*LMS*
Correct data ($x_{22} = 14.1$)	0.322	0.227
Damaged data ($x_{22} = 15.1$)	0.573	0.451

2. in a much shorter time.

Boček-Lachout algorithm

Hettmansperger, T. P., S. J. Sheather (1992):
A Cautionary Note on the Method of Least Median Squares.
The American Statistician 46, 79–83.

Engine Knock Data $(n = 16, p = 4, h = 11)$

c	x_1	x_2	x_3	x_4	y

The value of $\hat{\beta}^{(LMS,n,h)}$ by Boček-Lachout algorithm.

Data	Interc.	SPARK	AIR	INTK	EXHS.
Correct data ($x_{22} = 14.1$)	30.04	0.14	3.08	0.46	-0.01
Damaged data($x_{22} = 15.1$)	48.38	-0.73	3.36	0.23	-0.01

The difference between these two models is much lower.
So, the effect announced by H-S was a consequence of the bad algorithm.

x_1 is spark timing x_2 air/fuel ratio
x_3 intake temperature x_4 exhaust temperature
y engine knock number

Hettmansperger, T. P., S. J. Sheather (1992):
A Cautionary Note on the Method of Least Median Squares.
The American Statistician 46, 79–83.

Engine Knock Data $(n = 16, p = 4, h = 11)$

c	x_1	x_2	x_3	x_4	y

The value of $\hat{\beta}^{(LMS,n,h)}$ by Boček-Lachout algorithm.

Data	Interc.	SPARK	AIR	INTK	EXHS.
Correct data ($x_{22} = 14.1$)	30.04	0.14	3.08	0.46	-0.01

BUT THIS CONCLUSION - ALTHOUGH TRUE - WAS MISLEADING.

x_1 is spark timing x_2 air/fuel ratio
x_3 intake temperature x_4 exhaust temperature
y engine knock number

We have seen: A shock and frustration

Hettmansperger, T. P., S. J. Sheather (1992):
A Cautionary Note on the Method of Least Median Squares.
The American Statistician 46, 79–83.

Engine Knock Data $(n = 16, p = 4, h = 11)$

c	x_1	x_2	x_3	x_4	y
1	13.3	13.9	31	697	84.4
					84.1
					88.4
					84.2
⋮	⋮	⋮	⋮	⋮	⋮
14	12.7	16.1	35	649	93.0
15	12.9	15.1	36	721	93.3
16	12.7	15.9	37	696	93.1

Let's verify once again that
$h = [\frac{n}{2}] + [\frac{p+1}{2}] = 11.$

x_1 is spark timing x_2 air/fuel ratio
x_3 intake temperature x_4 exhaust temperature
y engine knock number

Hettmansperger, T. P., S. J. Sheather (1992):
A Cautionary Note on the Method of Least Median Squares.
The American Statistician 46, 79–83.

Engine Knock Data ($n = 16$, $p = 4$, $h = 11$)

Realize that $\binom{16}{11} = 4368$, so that we can compute $\hat{\beta}^{(LTS,16,11)}$ exactly, just computing $\hat{\beta}^{(OLS,11)}$ for all subsamples of size 11 and select the "best" one.

This is the exact value of $\hat{\beta}^{(LTS,n,h)}$!

Data	Interc.	SPARK	AIR	INTK	EXHS.
Correct data ($x_{22} = 14.1$)	35.11	-0.028	2.949	0.477	-0.009
Damaged data ($x_{22} = 15.1$)	-88.7	4.72	1.06	1.57	0.068

x_1 is spark timing x_2 air/fuel ratio

Víšek, J.Á (1994): A cautionary note on the method
of Least Median of Squares reconsidered.
Transactions of the Twelfth Prague Conference 1994, 254 - 259.

Correct data

Engine Knock Data (Air/Fuel 14.1, $n = 16$, $p = 4$)
(p is dimension of data including intercept)

Method	PRO-*LMS*	Bo-La-*LMS*	Exact *LTS*	Iterative *LTS*
11^{th} order stat.	0.3221	0.22783	0.3092	0.3092
Sum of squares	0.4239	0.3575	0.2707	0.2707

Damaged data

Engine Knock Data (Air/Fuel 15.1, $n = 16$, $p = 4$)
(p is dimension of data including intercept)

Method	PRO-*LMS*	Bo-La-*LMS*	Exact *LTS*	Iterative *LTS*
11^{th} order stat.	0.5729	0.4506	0.5392	0.5392
Sum of squares	1.0481	1.432	0.7283	0.7283

Another benchmark

Stackloss Data ($n = 21$, $p = 4$)
(p is dimension of data including intercept)

Brownlee, K.A. (1965):
Statistical Theory and Methodology in Science and Engineering, Wiley, NY.
Rousseew, P. J., A. M. Leroy (1987):
Robust Regression and Outlier Dectection, Wiley, NY.

Operational data of a plant for the oxidation of ammonia to nitric acid.

X1 - Air Flow X2 - Temperature X3 - Acid Concentration Y - Stackloss

Case	X1	X2	X3	Y
1	80	27	89	42
2	80	27	88	37
3	75	25	90	37
4	62	24	87	28
5	62	22	87	18
6	62	23	87	18
7	62	24	93	19

Case	X1	X2	X3	Y
8	62	24	93	20
9	58	23	87	15
10	58	18	80	14
11	58	18	89	14
12	58	17	88	13
13	58	18	82	11
14	58	19	93	12

Case	X1	X2	X3	Y
15	50	18	89	8
16	50	18	86	7
17	50	19	72	8
18	50	19	79	8
19	50	20	80	9
20	56	20	82	15
21	70	20	91	15

Another benchmark

Stackloss Data ($n = 21$, $p = 4$)
(p is dimension of data including intercept)

Brownlee, K.A. (1965):
Statistical Theory and Methodology in Science and Engineering, Wiley, NY.
Rousseew, P. J., A. M. Leroy (1987):
Robust Regression and Outlier Dectection, Wiley, NY.

Method	PRO-*LMS*	Bo-La-*LMS*	Exact *LTS*	Iterative *LTS*
12^{th} order stat.	0.6640	0.5321	0.7014	0.7014
Sum of squares	2.4441	1.9358	1.6371	1.6371

Another benchmark

Demographical Data $(n = 49, p = 7)$

Gunst, R. F., and Mason, R. L. (1980):
Regression Analysis and Its Application: A Data-Oriented Approach.
New York: Marcel Dekker.

Chatterjee, S., Hadi, A. S. (1988):
Sensitivity Analysis in Linear Regression. New York: J. Wiley & Sons.

1. X1 - Infant deaths per 1000 live birth
2. X2 - Number of inhabitants per physician
3. X3 - Population per square kilometer
4. X4 - Population per 1000 hectares of agricultural land
5. X5 - Percentage literate of population aged 15 years and over
6. X6 - Number of students enrolled in higher education per 100000 population
7. Y - Gross national product per capita in 1957 in $ (U.S.)

An algorith for LTS

LTS - ALGORITHM

A

Find the plane through $p+1$ randomly selected observations.

Evaluate squared residuals of all observations and order them increasingly. Then sum up the h smallest squares of residuals and the sum denote $S(\hat{\beta}_{present})$.

Is $S(\hat{\beta}_{present})$ less than $S(\hat{\beta}_{past})$?

no → B

yes

Establish *new* $\hat{\beta}_{present}$ just applying *OLS* on the h observations with the smallest squared residuals.

An algorithm for LTS

LTS - ALGORITHM$_{(continued)}$

B

Was ℓ-times found the same model with minimal value of $S(\beta)$?

yes

no

Was already k-times repeated outer cycle ?

no

A

yes

As $\hat{\beta}^{(LTS,n,w)}$ we will assume $\beta \in R^p$ for which the functional $S(\beta)$ attained - through just described iterations - minimal value.

Another benchmark

Demographical Data $(n = 49, p = 7)$

Gunst, R. F., and Mason, R. L. (1980):
Regression Analysis and Its Application: A Data-Oriented Approach.
New York: Marcel Dekker.
Chatterjee, S., Hadi, A. S. (1988):
Sensitivity Analysis in Linear Regression. New York: J. Wiley & Sons.

1. X1 - Infant deaths per 1000 live birth
2. X2 - Number of inhabitants per physician
3. X3 - Population per square kilometer
4. X4 - Population per 1000 hectares of agricultural land
5. X5 - Percentage literate of population aged 15 years and over
6. X6 - Number of students enrolled in higher education per 100000 population
7. Y - Gross national product per capita in 1957 in $ (U.S.)

Another benchmark

Demographical Data $(n = 49, p = 7)$

Gunst, R. F., and Mason, R. L. (1980):
Regression Analysis and Its Application: A Data-Oriented Approach.
New York: Marcel Dekker.

Chatterjee, S., Hadi, A. S. (1988):
Sensitivity Analysis in Linear Regression. New York: J. Wiley & Sons.

Method	PRO-*LMS*	Bo-La-*LMS*	Iterative *LTS*
28^{th} order stat.	131.50	95.38	104.20
Sum of squares	134260	132340	64159

Another benchmark

Educational Data $(n = 50, p = 4)$

Rousseeuw, P. J., Leroy, A. M. (1987):
Robust Regression and Outlier Detection. New York: J.Wiley & Sons.

Chatterjee, S., Hadi, A. S. (1988):
Sensitivity Analysis in Linear Regression. New York: J. Wiley & Sons.

1. X1 - Number of residents (per 1000) residing in urban areas in 1970
2. X2 - Personal income per capita in 1973
(i. e. sum of personal incomes divided by number of inhabitants)
3. X3 - Number of residents per thousand under 18 years of age in 1974
4. Y - Education expenditure for public education per capita in 1975

Another benchmark

Educational Data $(n = 50, p = 4)$

Rousseeuw, P. J., Leroy, A. M. (1987):
Robust Regression and Outlier Detection. New York: J.Wiley & Sons.

Chatterjee, S., Hadi, A. S. (1988):
Sensitivity Analysis in Linear Regression. New York: J. Wiley & Sons.

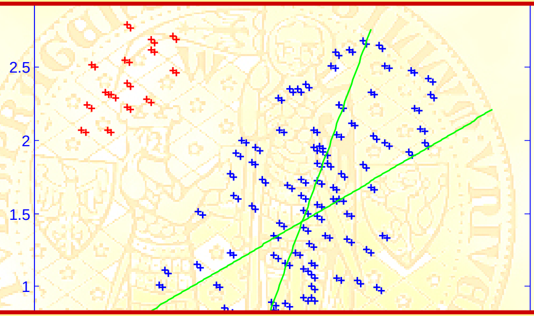
Method	PRO-*LMS*	Bo-La-*LMS*	Iterative *LTS*
27^{th} order stat.	19.3562	16.63511	19.0378
Sum of squares	3605.5	3728.6	3414.5

Diagnostics by LTS

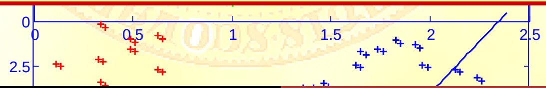
THE PROBLEM IS HOW LARGE h WE SHOULD SELECT FOR LTS.

We may start with $h \approx \frac{n}{2}$ and increase it step by step. It works as follows.

FOR $h << k$ WE OBTAIN ONE OF GREEN LINES,
AND ESTIMATES OF COEFFS (ETC.) MODESTLY VARY.



FOR $h \leq k$ BUT NEAR TO k WE OBTAIN BLUE LINE,
POPULATIONS ARE NESTED
AND ESTIMATES OF COEFFS (ETC.) ARE STABLE.



Diagnostics by LTS

ANALYSIS OF THE EXPORT FROM THE CZECH REPUBLIC TO EU IN 1994 BY MEANS OF THE *least trimmed squares*.

Benáček, V., J. Á Víšek (2002):
Impacts of the EU opening-up on small open economy:
Czech exports and imports.
In *Karadeloglou P. (ed.): Enlarging the EU - The Trade Balance Effects Palgrave/Macmillan, New York, 2002, 3 - 29.*

Atkinson, A. C., M. Riani, A. Cerioli (2004):
Exploring multivariate data with the forward search.
Springer, NY, Berlin, Heidelberg.

green is $\hat{\beta}_3^{(LTS,n,h)}$, the red is $\hat{\beta}_4^{(LTS,n,h)}$ and the light blue (the lowest curve) is $\hat{\beta}_6^{(LTS,n,h)}$ (down-scaled again by $\frac{1}{10}$). There is an evident break at 54.

Analysis of the export from the Czech Republic to EU in 1994 by means of the *least trimmed squares*

has found:

Main subgroup

with number of industries 54 and model

$$\frac{X_\ell}{S_\ell} = 4.64 - 0.032 \cdot \frac{US_\ell}{VA_\ell} - 0.022 \cdot \frac{HS_\ell}{VA_\ell} - 0.124 \cdot \frac{K_\ell}{VA_\ell} + 1.035 \cdot CR_\ell$$
$$-3.199 \cdot TFPW_\ell + 1.048 \cdot BAL_\ell + 0.452 \cdot DP_\ell + \varepsilon_\ell$$

X_ℓ	-	*export from i-th industry,*
US_ℓ	-	*number of university-passed employees in the i-th industry,*
HS_ℓ	-	*nuber of high school-passed employees in the i-th industry,*
VA_ℓ	-	*value added in the i-th industry,*
K_ℓ	-	*capital in the i-th industry,*
CR_ℓ	-	*percentage of market occupied by 3 largest producers,*
$TFPW_\ell$	-	*by wages normed productvity in the i-th industry,*
Bal_ℓ	-	*Balasa index in the i-th industry„*
DP_ℓ	-	*cost discontinuity in 1993 in the i-th industry*

with coefficient of determination 0.97 and stable submodels

ANALYSIS OF THE EXPORT FROM THE CZECH REPUBLIC TO EU IN 1994 BY MEANS OF THE *least trimmed squares*

has found:

COMPLEMENTARY SUBGROUP

with number of industries 33 and model

$$\frac{X_\ell}{S_\ell} = -0.634 + 0.089 \cdot \frac{US_\ell}{VA_\ell} + 0.235 \cdot \frac{HS_\ell}{VA_\ell} + 0.249 \cdot \frac{K_\ell}{VA_\ell} + 1.174 \cdot CR_\ell$$
$$+0.690 \cdot TFPW_\ell + 2.691 \cdot BAL_\ell - 0.051 \cdot DP_\ell + \varepsilon_\ell$$

X_ℓ	-	*export from i-th industry,*
US_ℓ	-	*number of university-passed employees in the i-th industry,*
HS_ℓ	-	*nuber of high school-passed employees in the i-th industry,*
VA_ℓ	-	*value added in the i-th industry,*
K_ℓ	-	*capital in the i-th industry,*
CR_ℓ	-	*percentage of market occupied by 3 largest producers,*
$TFPW_\ell$	-	*by wages normed productvity in the i-th industry,*
Bal_ℓ	-	*Balasa index in the i-th industry,,*
DP_ℓ	-	*cost discontinuity in 1993 in the i-th industry*

with coefficient of determination 0.93 and stable submodels

An algorith for LWS

LWS - ALGORITHM

A

Find the plane through $p+1$ randomly selected observations.

↓

Evaluate squared residuals of all observations. Then sum up the products of the weights and of the order statistics of squared residuals and the sum denote $S(\hat{\beta}_{present})$.

↓

Is $S(\hat{\beta}_{present})$ less than $S(\hat{\beta}_{past})$?

no → B

↓ yes

Establish *new* $\hat{\beta}_{present}$ just applying *WLS* on the reordered observations (reoredered according to the squared residuals).

An algorith for LWS

LWS - ALGORITHM(*continued*)

B

Was ℓ-times found the same model with minimal value of $S(\beta)$?

yes

no

Was already k-times repeated outer cycle ?

no

A

yes

As $\hat{\beta}^{(LTS,n,w)}$ we will assume $\beta \in R^p$ for which the functional $S(\beta)$ attained - through just described iterations - minimal value.

Víšek, J. Á. (1990): Empirical study of estimators of coefficients of linear regression model.
Technical report of Institute of Information Theory and Automation,
Czechoslovak Academy of Sciences (1990), number 1699.

Antoch, J., J. Á. Víšek: Robust estimation in linear models and its computational aspects.
Contributions to Statistics: Computational Aspects of Model Choice,
Springer Verlag, (1992), 39 - 104.

Hawkins, D. M. (1994):
The feasible solution algorithm for least trimmed squares regression.
Computational Statistics and Data Analysis 17, 185 - 196.

Hawkins, D. M., D. J. Olive (1999):
Improved feasible solution algorithms for breakdown estimation.
Computational Statistics & Data Analysis 30, 1 - 12.

Antoch, J., J. Á. Víšek: Robust estimation in linear models and its computational aspects.
Contributions to Statistics: Computational Aspects of Model Choice,
Springer Verlag, (1992), 39 - 104.

Hawkins, D. M. (1994):
The feasible solution algorithm for least trimmed squares regression.
Computational Statistics and Data Analysis 17, 185 - 196.

Hawkins, D. M., D. J. Olive (1999):
Improved feasible solution algorithms for breakdown estimation.
Computational Statistics & Data Analysis 30, 1 - 12.

Čížek, P., J. Á. Víšek (2000): Least trimmed squares.
XPLORE, Application Guide, 49 - 64. Springer Verlag, (2000), Berlin,
eds. W. Hărdle, Z. Hlávka, S. Klinke.

Hawkins, D. M., D. J. Olive (2003): Inconsistency of resampling algorithm for high breakdown regression estimation and a new algorithm. *Journal of the American Statistical Association 97, 136-159.*

Rousseeuw, P.J., K. van Driessen (2006): *Computing LTS regression for large data sets.* *Data Mining and Knowledge Discovery 12, 29 - 45.*

Van Huffel, S., J. Vandewalle (1988): The partial total least squares algorithm. *Journal of Computational and Applied Mathematics 21, 333 - 341.*

Salibian-Barrera, M., V. Yohai (2006): A fast algorithm for *S*-regression estimates. *Journal of Computational and Graphical Statistics 15, 414-427.*

THANKS FOR ATTENTION