R-Estimation in Linear Models with α-stable Errors

Marc Hallin

Département de Mathématique and E.C.A.R.E.S., Université Libre de Bruxelles

London, December 11, 2015

Outline of the talk

Structure

Stable distributions

Stable distributions are extremely attractive from several points of view.

Stochastic properties

(1) Stable distributions are the only nondegenerate distributions with a *domain of attraction* : non-trivial limits of normalized sums of independent identically distributed terms are *necessarily* stable.

(2) Stable families are quite flexible : four parameters

$$\boldsymbol{\theta} := (\alpha, b, c, \delta) \in \boldsymbol{\Theta} = (0, 2] \times [-1, 1] \times R^{+} \times \mathbb{R}$$

(a) δ and c are *location-scale* parameters :

$$f_{(\alpha,b,c,\delta)}(x) = f_{\alpha,b,1,0}\left(\frac{x-\delta}{c}\right)/c.$$

(b) α and b are *shape parameters* :

- α, the *characteristic exponent* is a *tail index* : the smaller α, the heavier the tails
- b is a *skewness* parameter : f is symmetric if $b = 0$, totally skewed if $|b| = 1$.

(3) some well-known stable densities

(a) $\alpha = 2$ (any $b \in [-1, 1]$) : Gaussian distribution, $f(x) = \frac{1}{\sqrt{4\pi}} e^{-x^2/4}$.

(b) $\alpha = 1$ and $b = 0$: Cauchy distribution, $f(x) = \frac{1}{\pi(1+x^2)}$.

(c) $\alpha = 1/2$ and $b = 1$: Lévy distribution $f(x) = \sqrt{\frac{1}{2\pi}} \frac{e^{-1/2x}}{x^{3/2}}$.

Stochastic modelling

Empirical evidence of non-Gaussian stable behavior is present in a variety of fields, among which economics, insurance, finance, signal processing, teletraffic engineering, ...

... where neglecting heavy tails and asymmetry results in underestimated risks , reckless decision making, and quite severe losses.

Student families (generally with three degrees of freedom or more) therefore are quite popular in such areas—but Student distributions are symmetric, and Student tails with three or five degrees of freedom often are still too light :

Only stable tails provide a reasonable account for a number of stylized facts

Moreover, ...

Statistical inference : Who's afraid of heavy tails ?

Stable families : a statistician's dream ?

Contrary to a widespread opinion, statistical experiments involving stable noise are extremely well-behaved.

In this talk, we concentrate on linear (regression) models driven by stable errors.

We show below that linear models with i.i.d. stable errors are Locally Asymptotically Normal (LAN, and even ULAN, with traditional root-n contiguity rates)—a most comfortable situation, under which *all* inference problems, *in principle*, can be solved in a locally asymptotically optimal way.

... although, as a rule, the traditional Gaussian procedures are not valid anymore

Statistical inference : Who's afraid of heavy tails ?

Stable families : a statistician's nightmare ?

No closed form for stable densities ! ! (except for the Gaussian, Cauchy and Lévy densities).

No finite moments of order p for any $p \geq \alpha$! (except for the Gaussian).

No standard central-limit behavior of traditional (Gaussian) statistics

Hence,

1. no closed-form likelihoods, even less for MLEs
2. no closed forms for optimal scores (log-derivatives of the densities)
3. no closed forms for *central sequences* (in the LAN framework) ...

Statistical inference : Who's afraid of heavy tails ?

A very rich literature exists on algorithmic methods trying to palliate the lack of explicit forms. That literature, it seems, remains largely underexploited by practitioners.

However,
• specifying or estimating the tail parameter α reamins difficult/risky
• assuming that appropriate "stable-likelihood-based" procedures can be worked out, they are likely to be sensitive to violations of the stability assumption : while traditional Gaussian methods notoriously break down under stable densities and infinite variances, the converse is likely to hold as well : stable likelihood-based (stable quasi-likelihood) methods are likely to run into problems under non-stable conditions.

Rank-based inference/Rank tests and R-estimation

Rank-based methods, thanks to distribution-freeness, appear as a simple and quite natural alternative to stable quasi-likelihood procedures.

Moreover, as we shall see, rank-based methods (in the context of linear models) achieve parametric efficiency at stable reference densities.

Surprisingly, ranks seldom (never ?) have been considered in the stable context. Several delicate questions indeed remain open.

Under stable densities or stable noise,

1. which rank tests/R-estimators should we use ?
2. what are the performances of those tests ?/ the asymptotic variances of those estimators ?
3. feasability ? (computational problems in relation with the absence of explicit densities/scores ...)

Those are the issues we plan to investigate here in the familiar context of linear regression.

Structure

Hypothesis testing in linear models with stable noise

Denote by $\mathcal{H}_f^{(n)}(\boldsymbol{\beta})$ the hypothesis under which the vector of observations $(X_1^{(n)}, \ldots, X_n^{(n)})'$ satisfies the equation

$$X_i^{(n)} = a + \sum_{l=1}^{K} c_{il}^{(n)} \beta_l + \varepsilon_i^{(n)}, \;\; i = 1, \ldots, n,$$

where

- $\mathbf{c}_i^{(n)} := \left(c_{i1}^{(n)}, \ldots, c_{iK}^{(n)} \right)'$ are regression constants, satisfying the usual conditions ;
- the intercept a is a nuisance, the regression parameter $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)'$ is the parameter of interest ;
- $a + \varepsilon_i^{(n)}$, $i = 1, \ldots, n$ is a sequence of i.i.d. random variables with density f.

Under $\mathcal{H}_f^{(n)}(\boldsymbol{\beta})$, the residuals $Z_i^{(n)}(\boldsymbol{\beta}) = X_i^{(n)} - \sum_{k=1}^{K} c_{ik}^{(n)} \beta_k$ $(i = 1, \ldots, n)$ then are i.i.d. with density f.

(i) If the errors are Gaussian, optimal testing procedures are well-known : optimal tests are based on the Student statistic T_n, which is asymptotically standard normal ; OLS estimators are optimal.

(ii) If the errors are non-normal α-stable, the optimal testing/estimation problem is a non-standard one, but LAN, as we shall see, in principle, provides asymptotically optimal solutions.

We concentrate on testing null hypotheses of the form $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, but linear restrictions on $\boldsymbol{\beta}$ could be considered as well, leading to the same efficiency conclusions and comments.

ULAN for general linear model with stable errors

The main theoretical tools throughout are *Local Asymptotic Normality* (LAN, actually ULAN) and *Le Cam's third Lemma*.

Let $\bar{c}_k^{(n)} := n^{-1}\sum_{i=1}^n c_{ik}^{(n)}$, $\mathbf{c}_i^{(n)} := (c_{i1}^{(n)}, \ldots, c_{iK}^{(n)})'$, $\mathbb{C}^{(n)} := n^{-1}\sum_{i=1}^n \mathbf{c}_i^{(n)}\mathbf{c}_i^{(n)\prime}$, and $\mathbb{K}^{(n)} := \left(\mathbb{C}^{(n)}\right)^{-1/2}$.

Assumption (A1) For all $n \in \mathbb{N}$, $\mathbb{C}^{(n)}$ is positive definite and converges, as $n \to \infty$, to a positive definite $\mathbb{K}^{-2}$.

Assumption (A2) (Noether conditions) For all $k = 1, \ldots, K$, one has

$$\lim_{n\to\infty}\left[\max_{1\leq t\leq n}\left(c_{tk}^{(n)} - \bar{c}_k^{(n)}\right)^2 \Big/ \sum_{t=1}^n \left(c_{tk}^{(n)} - \bar{c}_k^{(n)}\right)^2\right] = 0.$$

Denote by $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\beta}}$ the probability distribution of $\mathbf{X}^{(n)}$ under parameter values $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$. Let

$$Z_i^{(n)}(\boldsymbol{\beta}) := X_i^{(n)} - \sum_{k=1}^{K} c_{ik}^{(n)} \beta_k \ (i = 1, \ldots, n), \quad i = 1, \ldots, n$$

be the residual associated with $\boldsymbol{\beta}$. Under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},\boldsymbol{\beta}}$, the residuals $Z_i^{(n)}(\boldsymbol{\beta})$ coincide with $a + \varepsilon_i^{(n)}$: they are i.i.d. with density $f_{\boldsymbol{\theta}} = f_{(\alpha,b,c,a)}$.

The following then holds.

Theorem

(ULAN) Suppose that (A1) and (A2) hold. Let $\boldsymbol{\nu}(n) := n^{-\frac{1}{2}}\mathbb{K}^{(n)}$ *and fix* $\boldsymbol{\theta} = (\alpha, b, c, a) \in \boldsymbol{\Theta}$. *Then, the regression model with stable errors is ULAN w.r.t.* $\boldsymbol{\beta}$. *More precisely, for all* $\boldsymbol{\beta} \in \mathbb{R}^K$, *all sequence* $\boldsymbol{\beta}^{(n)}$ *such that* $\boldsymbol{\nu}^{-1}(n)(\boldsymbol{\beta}^{(n)} - \boldsymbol{\beta}) = O(1)$ *and all bounded sequence* $\boldsymbol{\tau}^{(n)} \in \mathbb{R}^K$,

$$(i)\ \Lambda^{(n)}_{\boldsymbol{\theta},\boldsymbol{\beta}^{(n)}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}} := \log \frac{d\mathrm{P}^{(n)}_{\boldsymbol{\theta},a,\boldsymbol{\beta}^{(n)}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}}}{d\mathrm{P}^{(n)}_{\boldsymbol{\theta},a,\boldsymbol{\beta}^{(n)}}} = \sum_{t=1}^{n} \log \left[\frac{f_{\boldsymbol{\theta}}\left(Z_t^{(n)}\left(\boldsymbol{\beta}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}\right)\right)}{f_{\boldsymbol{\theta}}\left(Z_t^{(n)}(\boldsymbol{\beta})\right)} \right]$$

$$= \boldsymbol{\tau}^{(n)\prime}\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}}(\boldsymbol{\beta}^{(n)}) - \frac{1}{2}\mathcal{I}(\boldsymbol{\theta})\boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)} + o_P(1)$$

under $\mathcal{H}^{(n)}_{\boldsymbol{\theta}}(\boldsymbol{\beta})$ *as* $n \to \infty$, *where, setting* $\varphi_{\boldsymbol{\theta}}(\cdot) := -\dot{f}_{\boldsymbol{\theta}}(\cdot)/f_{\boldsymbol{\theta}}(\cdot)$,

$$\mathcal{I}(\boldsymbol{\theta}) := \int_{-\infty}^{\infty} \varphi^2_{\boldsymbol{\theta}}(x) f_{\boldsymbol{\theta}}(x)dx$$

($\mathcal{I}(\boldsymbol{\theta})\mathbf{I}$ *is the* information matrix*) and*

$$(ii) \qquad \boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}}(\boldsymbol{\beta}) = n^{-1/2}\left(\mathbb{K}^{(n)}\right)' \sum_{i=1}^{n} \varphi_{\boldsymbol{\theta}}\left(Z_i^{(n)}(\boldsymbol{\beta})\right)\mathbf{c}_i^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta})\mathbf{I}) \qquad (2.1)$$

($\boldsymbol{\Delta}^{(n)}_{\boldsymbol{\theta}}(\boldsymbol{\beta})$ *a* central sequence*).*

Benefits of ULAN

Consequences of ULAN

ULAN allows us to

1. build optimal "parametric" and *rank-based* tests for $\mathcal{H}^{(n)}_{f_{\boldsymbol{\theta}}}(\boldsymbol{\beta})$
   - the validity of the "parametric" tests requires correct specification of f_θ, and/or root-n-consistent estimation of $\boldsymbol{\theta}$, while
   - the validity of *rank-based* tests does not depend on the underlying distribution ; this includes, of course, the stable ones, and $\boldsymbol{\theta}$ thus needs not be estimated ;
2. (via Le Cam's "third Lemma") compute the asymptotic local powers of these (and any other) tests (now, as a function of the actual f, hence, in the stable case, a function of $\boldsymbol{\theta}$) ;
3. compare these tests through AREs ;
4. perform (one-step) R-estimation, with the same ARE values as the corresponding rank tests.

Structure

Rank tests

Consider a *target* or *reference distribution* g at which LAN also holds, with central sequence

$$\boldsymbol{\Delta}_g^{(n)}(\boldsymbol{\beta}) = n^{-\frac{1}{2}} \left(\mathbb{K}^{(n)}\right)' \sum_{i=1}^{n} \varphi_g \left(Z_i^{(n)}(\boldsymbol{\beta}) \right) \mathbf{c}_i^{(n)};$$

define the "rank based central sequence"

$$\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}) = n^{-\frac{1}{2}} \left(\mathbb{K}^{(n)}\right)' \sum_{i=1}^{n} J \left(\frac{R_i^{(n)}}{n+1} \right) \mathbf{c}_i^{(n)},$$

where

(i) $J : (0,1) \mapsto \mathbb{R} : x \to \varphi_g(G^{-1}(x))$,

(ii) $\mathbf{R}^{(n)} = \mathbf{R}^{(n)}(\boldsymbol{\beta}_0) = (R_1^{(n)}, \ldots, R_n^{(n)})$ is the vector of ranks of the residuals $Z_1^{(n)}(\boldsymbol{\beta}_0), \ldots, Z_n^{(n)}(\boldsymbol{\beta}_0)$.

Rank-based tests

Proposition

Set

$$T_J^{(n)}(\boldsymbol{\beta}_0) = \mathcal{J}^{-1}(J)\ \left(\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0)\right)' \left(\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0)\right)$$

Then $T_J^{(n)}(\boldsymbol{\beta}_0)$ is

- asymptotically chi-square under $\bigcup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}_0)$,
- asymptotically chi-square under $\bigcup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}_0 + n^{-1/2}\boldsymbol{\tau})$ with non-centrality parameter

$$\frac{\boldsymbol{\tau}'\boldsymbol{\tau}\mathcal{J}^2(J,\boldsymbol{\theta})}{\mathcal{J}(J)},$$

with $\mathcal{J}(J,\boldsymbol{\theta}) = \int_0^1 J(u)\varphi_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(u))du$ and $\mathcal{J}(J) = \int_0^1 J^2(u)du$.

Tests and AREs

The corresponding tests (at nominal asymptotic level α) consist in rejecting the null hypothesis $\bigcup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \mathcal{H}^{(n)}_{\boldsymbol{\theta}}(\boldsymbol{\beta}_0)$ whenever $T^{(n)}_J(\boldsymbol{\beta}_0)$ exceeds the α-upper quantile of the (central) chi-square distribution with K degrees of freedom.

Let J and $\tilde{J}$ be two score-generating functions, and denote by $\text{ARE}_{\boldsymbol{\theta}}\left(J/\tilde{J}\right)$ the asymptotic relative efficiency, under stable density $f_{\boldsymbol{\theta}}$, of the rank test based on $T^{(n)}_J(\boldsymbol{\beta}_0)$ with respect to the rank test based on $\tilde{T}^{(n)}_{\tilde{J}}(\boldsymbol{\beta}_0)$. Then,

$$\text{ARE}_{\boldsymbol{\theta}}\left(J/\tilde{J}\right) = \frac{\mathcal{J}^2(J,\boldsymbol{\theta})}{\mathcal{J}^2(\tilde{J},\boldsymbol{\theta})}\frac{\mathcal{J}(\tilde{J})}{\mathcal{J}(J)}.$$

Standard tests ...

Standard tests

We applied those results to the following standard tests :

1. *van der Waerden scores :*

$$J(u) = \Phi^{-1}(u),$$

2. *Wilcoxon scores :*

$$J(u) = \frac{\pi}{\sqrt{3}}(2u - 1),$$

3. *Laplace scores :*

$$J(u) = \sqrt{2} sign(F^{-1}(u)),$$

with $F(\cdot)$ cdf of standardized double-exponential.

and less standard ones ...

new rank tests based on "stable scores"

... but also to some non standard tests, based on stable scores :

1. *Cauchy scores :*
$$J(u) = \sin(2\pi(u - 1/2)),$$
2. *Lévy scores :*
$$J(u) = \sqrt{2}\,\left(\Phi^{-1}((u+1)/2)\right)^2 (3 - 2\sqrt{2}\,\left(\Phi^{-1}((u+1)/2\right)^2),$$
3. *general stable scores :* $J(u) = \varphi_f(F^{-1}(u))$.

Remarks on AREs

Recall that

Theorem (Chernoff-Savage, 1958)

$$\inf_g ARE_g(vdW/Student) = 1$$

Theorem (Hodges-Lehmann, 1956)

$$\inf_g ARE_g(W/Student) = 0.864$$

In the present context, however, since Student tests are not valid, we rather take the van der Waerden tests (which uniformly dominate the Student ones) as a reference for ARE computations.

AREs : Wilcoxon, Laplace, Cauchy vs van der Waerden

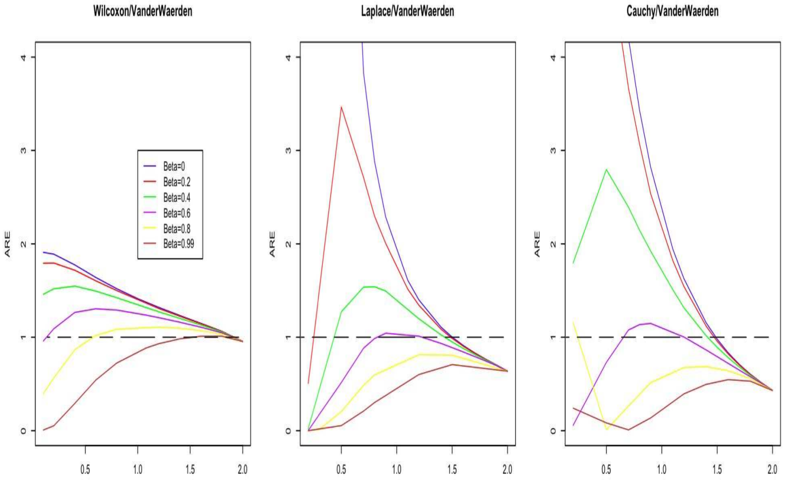

Figure: AREs of Wilcoxon, Laplace and Cauchy with respect to van der Waerden as functions of the tail index α, for various values of the skewness parameter b.

A remark on AREs

Theorem

$$\sup_g ARE_g(W/vdW) = \frac{6}{\pi} \approx 1.910$$

where the supremum is taken over all g with finite Fisher information for location (which includes stable densities).

The supremum is attained by the limiting version of symmetric heavy tailed laws with infinitely fat tails, e.g. α-stables with $\alpha \to 0$ or Student t_n with $n \to 0$.

AREs : Optimal stable vs van der Waerden

$\lvert b \rvert$	α =1.6			α =1.7			α =1.8			α =1.9		
	0	0.2	0.4	0	0.2	0.4	0	0.2	0.4	0	0.2	0.4
α =1.6												
0	1.2127	1.2045	1.1787	1.1332	1.1269	1.1075	1.0446	1.0407	1.0288	0.9444	0.9429	0.9386
0.2	1.2043	1.2129	1.2039	1.1277	1.1333	1.1256	1.0416	1.0447	1.0396	0.9433	0.9445	0.9428
0.4	1.1779	1.2033	1.2135	1.1100	1.12811	1.1337	1.0320	1.0425	1.0450	0.9396	0.9438	0.9450
α = 1.7												
0	1.2010	1.1954	1.1772	1.1442	1.1394	1.1241	1.0756	1.0721	1.0615	0.9893	0.9879	0.9834
0.2	1.1942	1.2011	1.1962	1.1393	1.1444	1.1392	1.0727	1.0757	1.0714	0.9882	0.9894	0.9876
0.4	1.1731	1.1925	1.2017	1.1236	1.1387	1.1448	1.0631	1.0730	1.0759	0.9845	0.9886	0.9897
α =1.8												
0	1.1645	1.1614	1.1511	1.1313	1.1284	1.1188	1.0878	1.0852	1.0771	1.0240	1.0226	1.0182
0.2	1.1600	1.1647	1.1628	1.1276	1.1315	1.1291	1.0852	1.0879	1.0852	1.0228	1.0240	1.0222
0.4	1.1465	1.1587	1.1654	1.1161	1.1267	1.1319	1.0767	1.0849	1.0881	1.0190	1.0231	1.0242
α = 1.9												
0	1.1005	1.0994	1.0957	1.0878	1.0867	1.0831	1.0704	1.0693	1.0656	1.0405	1.0394	1.0361
0.2	1.0988	1.1008	1.1006	1.0862	1.0880	1.0876	1.0689	1.0706	1.0698	1.0394	1.0405	1.0394
0.4	1.0937	1.0988	1.1019	1.0813	1.0860	1.0887	1.0643	1.0686	1.0709	1.0360	1.0393	1.0406

Table: AREs for tests based on stable scores with respect to van der Waerden's. Rows correspond to scores, columns to the (stable) densities under which AREs are computed. For instance, row 1 contains the AREs with respect to van der Waerden of the test based on stable scores for $\alpha = 1.6$, $b = 0$, under stable densities with tail parameter $\alpha = 1.6$ and skewness b ranging from 0 through 0.4.

Monte Carlo ...

We generated $N = 2500$ samples from the regression models

$$Y_i^{(l)} = ((l/20))c_i + \epsilon_i, \quad i = 1, \ldots, n = 100, \ \ l = 0,\ 1,\ 2,\ 3, \tag{3.2}$$

where the the ϵ_i's are i.i.d. with centered alpha-stable distribution. The regression constants c_i ($i = 1, \ldots, 100$) (the same ones across the 2500 replications) were drawn from the uniform distribution on $[-5, 5]$.

Observations $Y_i^{(0)}$ thus are generated under the null, $Y_i^{(1)}$, $Y_i^{(2)}$ and $Y_i^{(3)}$ under increasing alternatives of the form $\beta = l/20$, $l = 1, 2$ and 3.

We performed the various tests at nominal level 5% for the null hypothesis $\beta = 0$.

Critical values were computed from asymptotic distributions.

	density	l				density	l			
test		0	1	2	3		0	1	2	3
ϕ_{vdW}		.0416	.1728	.3968	.5824		.0420	.2324	.6116	.8728
ϕ_W		.0488	.2600	.5712	.7724		.0484	.3176	.7700	.9564
ϕ_L	$\alpha = .5$	.0500	.5992	.9032	.9780	$\alpha = .85$	.0476	.4600	.9084	.9908
ϕ_C	$\beta = 0$	.0496	.5304	.8576	.9500	$\beta = 0$	.0472	.4304	.8744	.9740
$\phi_{1.6;0}$		.0532	.2916	.6180	.8120		.0516	.3568	.8224	.9720
ϕ_t		.0164	.0244	.0240	.0204		.0288	.0344	.0516	.0872
ϕ_{vdW}		.0416	.2004	.4204	.6164		.0408	.2484	.6452	.8836
ϕ_W		.0500	.2784	.5784	.7752		.0428	.3388	.7752	.9580
ϕ_L	$\alpha = .5$	.0484	.3236	.6980	.8812	$\alpha = .85$	.0472	.3520	.8136	.9716
ϕ_C	$\beta = .4$	.0480	.1856	.3956	.5480	$\beta = .4$	.0492	.1932	.5124	.7632
$\phi_{1.6;0}$		.0508	.3124	.6444	.8244		.0476	.3744	.8300	.9772
ϕ_t		.0196	.0224	.0196	.0212		.0360	.0420	.0528	.0764
ϕ_{vdW}		.0396	.3472	.6668	.8176		.0424	.3488	.8020	.9604
ϕ_W		.0448	.2732	.5992	.7684		.0476	.3216	.7748	.9524
ϕ_L	$\alpha = .5$	.0448	.1028	.2224	.4036	$\alpha = .85$	0496	.1784	.5248	.8188
ϕ_C	$\beta = .99$	.0420	.1880	.2120	.1776	$\beta = .99$	.0480	.0500	.0544	.0780
$\phi_{1.6;0}$		.0428	.2280	.5396	.7336		.0488	.2884	.7444	.9452
ϕ_t		.0136	.0224	.0192	.0252		.0328	.0376	.0424	.0688

Table: Rejection frequencies (out of 2,500 replications), under the null ($l = 0$) and under alternatives ($l = 1, 2, 3$), of the van der Waerden test ϕ_{vdW}, the Wilcoxon test ϕ_W, the Laplace test (the sign test) ϕ_L, the Cauchy test ϕ_C, the test $\phi_{1.6;/0}$ (optimal at the stable distribution with $\alpha = 1.6$ and $\beta = 0$) and the Student test ϕ_t. Underlying stable densities with $\alpha = .5$ and $.85$.

	density	l				density	l			
test		**0**	**1**	**2**	**3**		**0**	**1**	**2**	**3**
ϕ_{vdW}	$\alpha = 1.6$, $\beta = 0$	.0424	.3964	.9208	.9968	$\alpha = 1.8$, $\beta = 0$	.0340	.4488	.9556	.9996
ϕ_{W}		.0512	.4580	.9540	.9992		.0416	.4848	.9660	.9996
ϕ_{L}		.0516	.3724	.9004	.9972		.0428	.3740	.9028	.9984
ϕ_{C}		.0488	.2788	.7400	.9624		.0440	.2392	.7044	.9520
$\phi_{1.6;0}$		.0580	.4864	.9624	.9996		.0432	.4880	.9680	.9996
ϕ_{t}		.0436	.2700	.6948	.8700		.0468	.4052	.8720	.9692
ϕ_{vdW}	$\alpha = 1.6$, $\beta = .4$	.0396	.3972	.9208	.9988	$\alpha = 1.8$, $\beta = .4$	.0364	.4436	.9600	1.000
ϕ_{W}		.0440	.4512	.9548	1.000		.0444	.4860	.9724	1.000
ϕ_{L}		.0492	.3568	.8952	.9956		.0508	.3832	.9024	1.000
ϕ_{C}		.0552	.2164	.6476	.9228		.0536	.2120	.6616	.9312
$\phi_{1.6;0}$		.0460	.4676	.9628	1.000		.0468	.4944	.9752	1.000
ϕ_{t}		.0464	.2836	.6848	.8748		.0468	.4064	.8844	.9664
ϕ_{vdW}	$\alpha = 1.6$, $\beta = .99$	.0392	.4404	.9504	.9992	$\alpha = 1.8$, $\beta = .99$	.0372	.4408	.9728	1.000
ϕ_{W}		.0492	.4584	.9532	.9992		.0400	.4624	.9768	1.000
ϕ_{L}		.0524	.3332	.8684	.9948		.0480	.3440	.8976	.9976
ϕ_{C}		.0500	.1352	.4172	.7512		.0464	.1736	.5440	.8708
$\phi_{1.6;0}$		.0552	.4392	.9472	.9988		.0408	.4608	.9676	1.000
ϕ_{t}		.0440	.2824	.7120	.8664		.0496	.3916	.8800	.9696

Table: Rejection frequencies (out of 2,500 replications), under the null ($l = 0$) and under alternatives ($l = 1, 2, 3$), of the van der Waerden test ϕ_{vdW}, the Wilcoxon test ϕ_{W}, the Laplace test (the sign test) ϕ_{L}, the Cauchy test ϕ_{C}, the test $\phi_{1.6;/0}$ (optimal at the stable distribution with $\alpha = 1.6$ and $\beta = 0$) and the Student test ϕ_t. Underlying stable densities with $\alpha = 1.6$ and 1.8 .

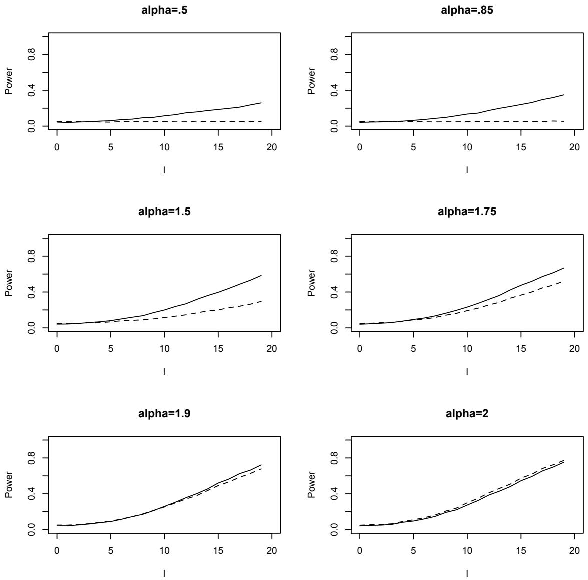

Figure: **Power curves of the van der Waerden (solid line) and Student (dotted line) tests computed from 10,000 replications for various symmetric stable errors. Sample size is $n = 100$ and regression constants are drawn from the uniform distribution on $[-5, 5]$.**

Structure

Classical estimation methods in the presence of heavy tails fail to provide satisfactory solutions.

(a) *OLS estimators :* consistency rate depends on the tail index α (Samorodnitsky *et al.* 2007) ; that rate is strictly less than the optimal root-n rate.

(b) *Stable MLEs :* problem of the absence of closed form likelihoods and the information matrix, moreover, is not block-diagonal.

(c) *Linear unbiased estimators :* consistency rates again crucially depend on α and are strictly less than the optimal root-n ones ; asymptotic covariances depend on α as well.

(d) *LAD (Least Absolute Deviations) estimators :* achieve rate-optimal consistency at arbitrary stable densities. But LAD estimators are optimal under (light-tailed) double-exponential noise, and cannot be efficient under any heavy-tailed stable densities.

Alternative : estimation based on ranks !

Hodges-Lehmann R-estimation

Estimation methods based on ranks—in short, R-estimation—go back to Hodges and Lehmann (1963) (one-sample and two-sample location models, based on the Wilcoxon and van der Waerden (signed) rank statistic) ; extension to regression was made possible by Jurečková (1971) and Koul (1971)

Under classical Argmin form, the Hodges-Lehmann R-estimator $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL}}$ of $\boldsymbol{\beta}$ is defined as

$$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL}} := \mathrm{argmin}_{\mathbf{t}\in\mathbb{R}^{\boldsymbol{K}}} \, | \, \underset{\sim}{Q}^{(n)}(\mathbf{R}^{(n)}(\mathbf{t}))|,$$

where $\underset{\sim}{Q}^{(n)}(\mathbf{R}^{(n)}(\boldsymbol{\beta}))$ is a (signed)-rank test statistic for the $\mathcal{H}_0 : \boldsymbol{\beta} = \boldsymbol{\beta}$ (two-sided test).

Main advantages of $\underset{\sim}{\beta}^{(n)}_{\rm HL}$ over usual M-estimators (under parameter value β and error density f, and standard root-n consistency conditions) :

- $n^{1/2}(\underset{\sim}{\beta}^{(n)}_{\rm HL} - \beta)$ is asymptotically equivalent to a function which depends on the unknown actual density f but is measurable with respect to the ranks $\mathbf{R}^{(n)}(\beta)$ of the unobservable noise ($\sim$ "equivariance" under monotone continuous transformations of the errors)
- the asymptotic relative efficiencies (AREs) of the R-estimator $\underset{\sim}{\beta}^{(n)}_{\rm HL}$ with respect to other R-estimators, or with respect to its Gaussian competitor (OLS or Gaussian MLE, when root-n consistent) are the same as the AREs of the corresponding rank tests with respect to their Gaussian competitors

Main advantages of $\underset{\sim}{\beta}^{(n)}_{\mathrm{HL}}$ over usual M-estimators (under parameter value β and error density f, and standard root-n consistency conditions) :

- $n^{1/2}(\underset{\sim}{\beta}^{(n)}_{\mathrm{HL}} - \beta)$ is asymptotically equivalent to a function which depends on the unknown actual density f but is measurable with respect to the ranks $\mathbf{R}^{(n)}(\beta)$ of the unobservable noise ($\sim$ "equivariance" under monotone continuous transformations of the errors)
- the asymptotic relative efficiencies (AREs) of the R-estimator $\underset{\sim}{\beta}^{(n)}_{\mathrm{HL}}$ with respect to other R-estimators, or with respect to its Gaussian competitor (OLS or Gaussian MLE, when root-n consistent) are the same as the AREs of the corresponding rank tests with respect to their Gaussian competitors

Main disadvantage of $\underset{\sim}{\vartheta}^{(n)}_{\mathrm{HL}}$ in the regression context

- the Argmin becomes rapidly impractical as the dimension of β increases (optimization over a K-dimensional grid)
- even for small K, a grid method involving stable scores is computationally infeasible

One-step R-estimators.

We therefore rather recommend a linearized form of the definition of the form

"preliminary root-n consistent" + rank based improvement.

Here the preliminary root-n consistent estimator will be the LAD estimator $(\hat{a}^{(n)}_{\mathrm{LAD}}, \hat{\boldsymbol{\beta}}^{(n)\prime}_{\mathrm{LAD}})'$ of $(a, \boldsymbol{\beta}')'$, obtained by minimizing the L_1-objective function

$$(\hat{a}^{(n)}_{\mathrm{LAD}}, \hat{\boldsymbol{\beta}}^{(n)\prime}_{\mathrm{LAD}})' := \mathrm{argmin}_{(a,\boldsymbol{\beta})\in\mathbb{R}^{K+1}} \sum_{i=1}^{n} |Z_i^{(n)}(\boldsymbol{\beta})|.$$

For the rank-based improvement, consider (again)

$$\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\boldsymbol{\beta}) := n^{-\frac{1}{2}} \mathbb{K}^{(n)\prime} \sum_{i=1}^{n} J\left(\frac{R_i^{(n)}}{n+1}\right) \mathbf{c}_i^{(n)}, \qquad (4.3)$$

which satisfies an asymptotic linearity property

$$\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\boldsymbol{\beta} + \boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}) - \underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\boldsymbol{\beta}) = -\mathcal{J}(J, g)\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1).$$

In principle, the one-step R-estimator of $\boldsymbol{\beta}$ should then take the very simple form

$$\underset{\sim}{\tilde{\boldsymbol{\beta}}}_J^{(n)} := \hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)} + \boldsymbol{\nu}^{(n)} \mathcal{J}^{-1}(J, g)\, \underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)})$$

From the asymptotic linearity of $\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}$, we get $\boldsymbol{\nu}^{-1}(n)(\underset{\sim}{\tilde{\boldsymbol{\beta}}}_J^{(n)} - \boldsymbol{\beta})$ is asymptotically $\mathcal{N}(\mathbf{0}, (\mathcal{J}(J)/\mathcal{J}^2(J, g))\mathbf{I}_K)$ under $\mathrm{P}_{g,a,\boldsymbol{\beta}}^{(n)}$.

This in turn implies that $\boldsymbol{\nu}^{-1}(n)(\underset{\sim}{\tilde{\boldsymbol{\beta}}}_J^{(n)} - \boldsymbol{\beta})$, for $J(u) = \varphi_f(F^{-1}(u))$, is asymptotically $\mathcal{N}(\mathbf{0}, \mathcal{J}^{-1}(J)\mathbf{I}_K)$ under $\mathrm{P}_{f,a,\boldsymbol{\beta}}^{(n)}$, i.e.

$\underset{\sim}{\tilde{\boldsymbol{\beta}}}_J^{(n)}$ reaches parametric efficiency at correctly specified density $f = g$.

Unfortunately, the scalar *cross-information quantity* $\mathcal{J}(J, g)$ is not known.

Thus, $\tilde{\underset{\sim}{\beta}}_J^{(n)}$ is not a "genuine" estimator.

That cross-information quantity $\mathcal{J}(J, g)$ has to be consistently estimated.

To obtain such a consistent estimator, we adopt here an idea first developed in Hallin, Oja and Paindaveine (2006) and generalized in Cassart, Hallin and Paindaveine (2010).

The *one-step R-estimator*

$$\underset{\sim}{\boldsymbol{\beta}}_J^{(n)} := \tilde{\boldsymbol{\beta}}^{(n)}(\widehat{\mathcal{J}}^{-1}(J,g)) = \hat{\boldsymbol{\beta}}_{\rm LAD}^{(n)} + \boldsymbol{\nu}^{(n)}\widehat{\mathcal{J}}^{-1}(J,g)\,\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\hat{\boldsymbol{\beta}}_{\rm LAD}^{(n)})$$

is such that

• $n^{1/2}(\underset{\sim}{\boldsymbol{\beta}}_J^{(n)} - \boldsymbol{\beta})$ is asymptotically normal with mean zero and covariance matrix $(\mathcal{J}(J)/\mathcal{J}^2(J,g))\mathbb{K}^2$ under $\mathrm{P}_{g,a,\boldsymbol{\beta}}^{(n)}$ with $g \in \mathcal{F}$

• letting $J(u) = \varphi_f(F^{-1}(u))$, $\underset{\sim}{\boldsymbol{\beta}}_J^{(n)}$ achieves the parametric efficiency bound under $\mathrm{P}_{f,a,\boldsymbol{\beta}}^{(n)}$

• the asymptotic relative efficiences of R-estimators clearly coincide with those of the corresponding rank tests

Table: AREs of R-estimators with respect to LAD estimators

Estimators	Underlying stable density			
	$\alpha = 2\,;\ b = 0$	$\alpha = 1.8\,;\ b = 0$	$\alpha = 1.8\,;\ b = 0.5$	$\alpha = 0.5\,;\ b = 0.5$
$\underset{\sim}{\beta}^{(n)}_{J_{\rm W}}/\hat{\beta}^{(n)}_{\rm LAD}$	1.4999	1.3888	1.3984	1.7776
$\underset{\sim}{\beta}^{(n)}_{J_{\rm vdW}}/\hat{\beta}^{(n)}_{\rm LAD}$	1.5708	1.3056	1.3285	1.251
$\underset{\sim}{\beta}^{(n)}_{J_{\rm C}}/\hat{\beta}^{(n)}_{\rm LAD}$	0.6759	0.7880	0.7769	2.007
$\underset{\sim}{\beta}^{(n)}_{J_{1.8;0}}/\hat{\beta}^{(n)}_{\rm LAD}$	1.4459	1.4183	1.4222	1.6453
$\underset{\sim}{\beta}^{(n)}_{J_{1.8;.5}}/\hat{\beta}^{(n)}_{\rm LAD}$	1.4452	1.3969	1.4459	1.4432
$\underset{\sim}{\beta}^{(n)}_{J_{.5;.5}}/\hat{\beta}^{(n)}_{\rm LAD}$	0.0925	0.1099	0.1175	21.2364

AREs for R-estimators based on various scores with respect to the LAD estimator. Columns correspond to the (stable) densities under which AREs are computed, rows to the scores considered : Wilcoxon ($J_{\rm W}$), van der Waerden ($J_{\rm vdW}$), Cauchy ($J_{\rm C}$), and three ($\delta = 0$, $\gamma = 1$) stable scores ($J_{\alpha;b}$); recall that the R-estimator based on Laplace scores asymptotically coincides with the LAD estimator.

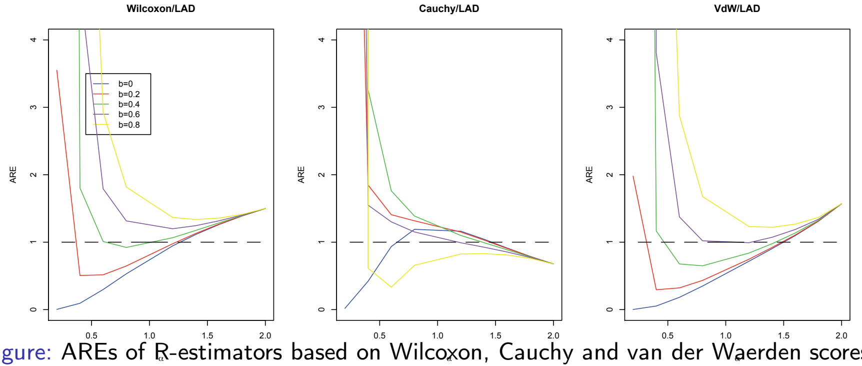

Figure: AREs of R-estimators based on Wilcoxon, Cauchy and van der Waerden scores, with respect to the LAD estimator, as a function of α and for various values of b.

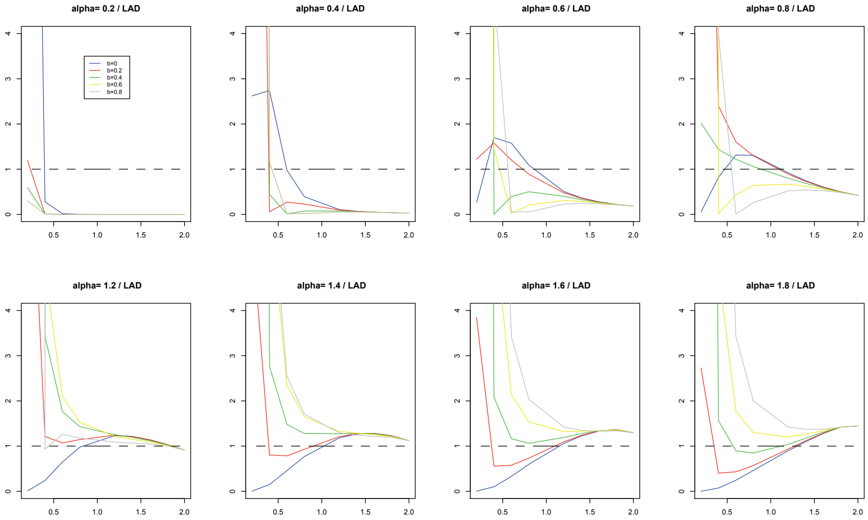

Figure: AREs under stable distributions of R-estimators based on various stable scores with respect to the LAD estimator, as a functions of α and b.

We generated $M = 1000$ samples from two multiple regression models,

$$Y_i^{(1)} = c_{i1} + c_{i2} + \epsilon_i, \quad i = 1, \ldots, n = 100, \tag{4.4}$$

with two regressors, and

$$Y_i^{(2)} = c_{i1} + c_{i2} + c_{i3} + c_{i4} + \epsilon_i, \quad i = 1, \ldots, n = 100, \tag{4.5}$$

with four regressors, both with alpha-stable i.i.d. ϵ_i's. The regression constants c_{ij} (the same ones across the 1000 replications) were drawn (independently) from the uniform distribution on $[-1, 1]^2$ and $[-1, 1]^4$, respectively. Letting $\mathbf{1}_K := (1, 1, \ldots, 1) \in \mathbb{R}^K$, the true values of the regression parameters are thus $\boldsymbol{\beta} = \mathbf{1}_2$ in model (4.4) and $\boldsymbol{\beta} = \mathbf{1}_4$ in model (4.5).

Table: Empirical bias and MSE for various estimators of $\boldsymbol{\beta}$ in (4.4) (2 regressors)

Estimator		Underlying stable density (α/b)				
		$\alpha=2/b=0$	$\alpha=1.8/b=0$	$\alpha=1.8/b=0.5$	$\alpha=1.2/b=0$	$\alpha=1.2/b=0.5$
$\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LS}}$	(Bias)	.00193	-.00134	.01385	.18680.	-.19255
	(MSE)	.06770	.19459	.27336	124.46	88.070
$\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LAD}}$	(Bias)	.00167	-.00087.	.00502	.02995	.00646
	(MSE)	.10674	.10411	.11638	.11560	.13396
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\mathrm{vdW}}}$	(Bias)	.00256	-.00136	.00694	.03376.	-.00243
	(MSE)	.06878	.07694.	.08545	.15165.	.14499
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\mathrm{W}}}$	(Bias)	.00076	.00015	.00920	.02957	-.00147
	(MSE)	.07234	.07454	.08366	.12060	.12219
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\mathrm{L}}}$	(Bias)	.00167	-.00087	.00502	.02995	.00646
	(MSE)	.10674	.10411.	.11638	.11560	.13396
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.8/0}}$	(Bias)	.00250	.00063	.00883	.03046	.00068
	(MSE)	.07088	.07457.	.08310	.12976.	.12820
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.8/.5}}$	(Bias)	.00187	-.00119	.01057	.03284	-.00037
	(MSE)	.07104	.07683	.08139	.13562	.12398
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.2/0}}$	(Bias)	.00424	.00353	.01373	.02155	-.00363
	(MSE)	.11613	.09812	.11040	.09641	.10971
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.2/.5}}$	(Bias)	.00670	-.00418	.01609	.02735	.00310
	(MSE)	.11416	.10382	.10822	.11455	.08917
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{.5/.5}}$	(Bias)	.01070	.03350	.00357	.04768	-.01671
	(MSE)	.22575	.28311	.24386	.35926	.18999
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL};\mathrm{vdW}}$	(Bias)	-.01668	-.01040	-.00253	.04306	-.01664
	(MSE)	.07936	.08958	.09508	.20227	.20441
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL};\mathrm{W}}$	(Bias)	-.00672	-.02019	-.01113	-.01052	-.03408
	(MSE)	.08225	.09071	.09702	.16290	.14918
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL};1.8/0}$	(Bias)	-.02274	-.02834	-.01923	-.01504	-.05129
	(MSE)	.09066	.10291	.10488	.18247	.19072

Empirical bias and MSE of the LSE $\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LS}}$, the LAD $\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LAD}}$ and various rank-based estimators computed from 1000 replications of model (4.4) with sample size n=100, under various stable error distributions.

Table: Empirical bias and MSE for various estimators of β in model (4.5) (4 regressors)

Estimator		Underlying stable density (α/b)				
		$\alpha=2/b=0$	$\alpha=1.8/b=0$	$\alpha=1.8/b=0.5$	$\alpha=1.2/b=0$	$\alpha=1.2/b=0.5$
$\hat{\beta}^{(n)}_{\rm LS}$	(Bias)	.00314	.01367	-.01945	-4.09468	-.09272
	(MSE)	.06339	.30161	.12752	15818.91	39.45292
$\hat{\beta}^{(n)}_{\rm LAD}$	(Bias)	.00693	.00880	-.00774	-.00652	.00352
	(MSE)	.09995	.09992	.09548	.08495	.09984
$\underset{\sim}{\beta}^{(n)}_{J_{\rm vdW}}$	(Bias)	.00378	.00638	-.01177	-.00763	-.01262
	(MSE)	.06463	.06964	.07238	.11369	.11015
$\underset{\sim}{\beta}^{(n)}_{J_{\rm W}}$	(Bias)	.00542	.00579	-.01236	-.00624	-.00774
	(MSE)	.06811	.06847	.06988	.09038	.09127
$\underset{\sim}{\beta}^{(n)}_{J_{\rm L}}$	(Bias)	.00693	.00880	-.00774	-.00652	.00352
	(MSE)	.09995	.09992	.09548	.08495	.09984
$\underset{\sim}{\beta}^{(n)}_{J_{1.8/0}}$	(Bias)	.00499	.00531	-.01221	-.00445	-.00980
	(MSE)	.06755	.06735	.07021	.09908	.09562
$\underset{\sim}{\beta}^{(n)}_{J_{1.8/.5}}$	(Bias)	.00339	.00526	-.01109	-.00438	-.01151
	(MSE)	.06686	.06914	.06977	.10095	.09397
$\underset{\sim}{\beta}^{(n)}_{J_{1.2/0}}$	(Bias)	.00802	.00608	-.01297	.00682	.00404
	(MSE)	.10763	.09229	.08986	.07061	.08406
$\underset{\sim}{\beta}^{(n)}_{J_{1.2/.5}}$	(Bias)	.00291	.00024	-.01401	.00396	-.00231
	(MSE)	.10332	.09233	.08567	.09036	.07037
$\underset{\sim}{\beta}^{(n)}_{J_{.5/.5}}$	(Bias)	.03400	.03653	-.02823	-.05925	-.00469
	(MSE)	.30150	.35030	.28818	.43049	.18807
$\underset{\sim}{\beta}^{(n)}_{\mathbf{HL};\mathbf{vdW}}$	(Bias)	.00401	.00634	-.01208	-.00704	-.01234
	(MSE)	.06513	.06968	.07266	.11310	.10956
$\underset{\sim}{\beta}^{(n)}_{\mathbf{HL};\mathbf{W}}$	(Bias)	.00513	.00623	-.01285	-.00547	-.00755
	(MSE)	.06854	.06855	.07006	.09010	0.09100
$\underset{\sim}{\beta}^{(n)}_{\mathbf{HL};1.8/0}$	(Bias)	.00494	.00582	-.01245	-.00396	-.01081
	(MSE)	.06783	.06753	.07037	.09854	.09594

Empirical bias and MSE of the LSE $\hat{\beta}^{(n)}_{\rm LS}$, the LAD $\hat{\beta}^{(n)}_{\rm LAD}$ and various rank-based estimators computed from 1000 replications of model (4.5) with sample size n=100, under various stable error distributions.

Table: One-step R-estimation versus Argmin

Estimator		Underlying stable density (α / b)					
		$\alpha = 2/b = 0$	$\alpha = 1.8/b = 0$	$\alpha = 1.8/b = 0.5$	$\alpha = 1.2/b = 0$	$\alpha = 1.2/b = 0.5$	$\alpha = 0.5/b =$
$K = 6$							
$\underset{\sim}{\beta}^{(n)}_{J_{\mathrm{vdW}}}$	(Bias)	-.01991	-.00485	.01084	-.01890	.02246	.00162
	(MSE)	.07707	.08821	.08935	.16485	.15258	.61554
$\underset{\sim}{\beta}^{(n)}_{\mathbf{HL}\,;\,\mathbf{vdW}}$	(Bias)	-.19519	-.19834	-.19202	-.36809	-.30435	-.59222
	(MSE)	.24257	.27483	.27461	.58981	.52245	2.51344
$K = 10$							
$\underset{\sim}{\beta}^{(n)}_{J_{\mathrm{vdW}}}$	(Bias)	-.00877	.00607	.00187	-.00807	-.01376	.06003
	(MSE)	.07834	.09133	.08641	.16835	.15545	1.4346
$\underset{\sim}{\beta}^{(n)}_{\mathbf{HL}\,;\,\mathbf{vdW}}$	(Bias)	-.91080	-.89626	-.92196	-1.00979	-.99976	-.97662
	(MSE)	1.04321	1.07289	1.09949	1.50269	1.43327	3.23870
$K = 15$							
$\underset{\sim}{\beta}^{(n)}_{J_{\mathrm{vdW}}}$	(Bias)	-.00374	-.01421	-.00575	.02479	0.00271	.01123
	(MSE)	.08894	.10969	.10539	.20918	.19621	2.00335
$\underset{\sim}{\beta}^{(n)}_{\mathbf{HL}\,;\,\mathbf{vdW}}$	(Bias)	-1.07573	-1.11915	-1.11057	-1.23107	-1.21492	-1.31910
	(MSE)	1.19685	1.33319	1.32890	1.91879	1.88120	4.32374

Empirical bias and MSE of the one-step and Argmin versions $\underset{\sim}{\beta}^{(n)}_{J_{\mathrm{vdW}}}$ and $\underset{\sim}{\beta}^{(n)}_{\mathbf{HL}\,;\,\mathbf{vdW}}$ of the van der Waerden R-estimator computed (via the Nelder-Mead (1965) algorithm for the Hodges-Lehmann case) from 1000 replications of model (4.5) with $K = 6, 10, 15$, sample size n=100 and various stable error distributions.

Conclusions

1. contrary to common belief, regression experiments with stable errors are LAN, with traditional root-n rate ;
2. traditional testing/estimation methods, however, as a rule, are rate-suboptimal in the presence of stable errors
3. One exception is the LAD estimator, along with the related median-test or Laplace score tests which are rate-optimal—but far from efficient
4. Rank-based methods allow for testing and estimation methods that remain valid irrespective of the tail index and skewness parameter ...
5. but can be tuned in order to reach parametric efficiency at given stable distribution
6. ... and, for adequate scores, yield uniformly better performance than LAD estimators and Laplace score tests over the class of stable densities with $\alpha \geq 1$ or $\alpha \leq 1$ ...
7. Finally, the one-step form of R-estimation significantly outperforms the Argmin form in finite sample

References

Hallin, M., Y. Swan, T. Verdebout and D. Veredas (2011) Rank-based testing in linear models with stable errors. *Journal of Nonparametric Statistics*, to appear.

Hallin, M., Y. Swan, T. Verdebout and D. Veredas (2011) One-step R-estimation in linear models with stable errors. *Journal of Econometrics*, to appear.