LAD Estimation for Time Series Models With Finite and Infinite Variance

Richard A. Davis
Colorado State University

William Dunsmuir
University of New South Wales

LAD Estimation for ARMA Models

- finite variance
- infinite variance

A Real Data Example

- LAD estimation
- unit root problem

Asymptotics in Non-standard Settings

(When Taylor series expansions do not work.)

Applications to:

- LAD estimation
  (Pollard `91, Davis and Dunsmuir `95)
- M-estimation with infinite variance
  (Davis, Knight and Liu `92, Davis `95)
- Unit root problems (AR + MA)
  (Davis and Dunsmuir `95, Davis, Chen and Dunsmuir `95)

Example (The median).

Data. $Z_1, \ldots, Z_n$, IID median 0 and pdf $f(0) > 0$.

Median. $\hat{m}_n = \text{median}(Z_1, \ldots, Z_n)$

(assume m_0 = population median = 0)

Asymptotics:

$\hat{m}_n$ is AN(0,??)

Note: $\hat{m}_n$ minimizes

$$T_n(m) = \sum_{t=1}^{n} \left(|Z_t - m| - |Z_t| \right)$$

Set $u = n^{1/2} m$ and put

$$S_n(u) = T_n(u n^{-1/2})$$

$$= \sum_{t=1}^{n} \left(|Z_t - u n^{-1/2}| - |Z_t|\right).$$

Then

$$\hat{u}_n = n^{1/2} \hat{m}_n .$$

Using the key identity,

$$|z-y| - |z| = -y \,\mathrm{sgn}(z) + 2(y-z) \left(I(0<z<y) - I(y<z<0)\right),$$

we have

$$S_n(u) = -u\, n^{-1/2} \sum_{t=1}^{n} \mathrm{sgn}(Z_t)$$

$$+ 2 \sum_{t=1}^{n} (n^{-1/2} u - Z_t)\{ I(0< Z_t < n^{-1/2} u) - I(n^{-1/2} u < Z_t < 0)\}$$

$$S_n(u) = -u\, n^{-1/2} \sum_{t=1}^{n} \operatorname{sgn}(Z_t)$$

$$+ 2 \sum_{t=1}^{n} (n^{-1/2}u - Z_t)\{ I(0 < Z_t < n^{-1/2}u) - I(n^{-1/2}u < Z_t < 0)\}$$

$$=: A_n + B_n$$

Results: $A_n \xrightarrow{d} -uN, \; N \sim N(0,1)$ (CLT)

$$EB_n = 2nE(n^{-1/2}u - Z_t)\, I(0 < Z_t < n^{-1/2}u) \quad (\text{for } u>0)$$

$$= 2n \int_0^{n^{-1/2}u} (n^{-1/2}u - z)\, F(dz)$$

$$\sim 2n \int_0^{n^{-1/2}u} (n^{-1/2}u - z)\, f(0)dz = u^2 f(0).$$

Conclude :

- $A_n \xrightarrow{d} -uN$
- $B_n \xrightarrow{P} u^2 f(0)$
- $S_n(u) \xrightarrow{d} S(u) := -uN + u^2 f(0)$ [on $C(\mathbb{R})$]
- $\hat{u}_n = n^{1/2}\hat{m}_n \xrightarrow{d} \hat{u}$:= minimizer of $S(u)$.

Solve $S'(u) = -N + 2\,u\,f(0) = 0$, we obtain

$$\hat{u} = N/(2f(0)) \sim N\big(0,1/(4f^2(0))\big)$$

or

$$n^{1/2}\hat{m}_n \xrightarrow{d} N\big(0,1/(4f^2(0))\big)$$

The Paradigm:

- Objective function to be minimized: $T_n(\theta)$
- Reparameterize by setting $u = a_n(\theta - \theta_0)$
  θ_0 = true value, a_n = scaling
- Form new objective function:

$$S_n(u) = T_n(\theta_0 + u/a_n)$$

- Establish weak convergence of $S_n(u)$ to $S(u)$ on $C(R)$.
- Show

$$\hat{u}_n = a_n(\hat{\theta} - \theta_0) \xrightarrow{d} \hat{u} := \operatorname{argmin} S(u)$$

Theorem. Let $\{Y_t\}$ be the linear process

$$Y_t = \sum_{j=0}^{\infty} c_j Z_{t-j}, \qquad \sum_{j=0}^{\infty} |c_j| < \infty,$$

where $\{Z_t\} \sim \mathrm{IID}(0,\sigma^2)$, $\mathrm{median}(Z_1)=0$, and $f(0)>0$.

Then

$$S_n := \sum_{t=1}^{n} \left(|Z_t - n^{-1/2} Y_{t-1}| - |Z_t|\right)$$

$$\xrightarrow{d} \gamma f(0) + N,$$

where $N \sim N(0, \gamma)$ $\quad (\gamma = \mathrm{Var}(Y_t))$.

Key identity:

$$|z-y| - |z| = -y\,\mathrm{sgn}(z) + 2(y-z)\big(I(0<z<y) - I(y<z<0)\big)$$

Using the key identity,

$$|z-y| - |z| = -y\,\mathrm{sgn}(z) + 2(y-z)\big(I(0<z<y) - I(y<z<0)\big),$$

we have

$$S_n = -\sum_{t=1}^{n} n^{-1/2} Y_{t-1}\mathrm{sgn}(Z_t)$$

$$+\, 2 \sum_{t=1}^{n} (n^{-1/2} Y_{t-1} - Z_t)\{ I(0< Z_t < n^{-1/2} Y_{t-1}) - I(n^{-1/2} Y_{t-1}<Z_t<0)\}$$

$$=:\ A_n + B_n$$

Result: $A_n \xrightarrow{d} N$ (MG CLT)

$B_n \xrightarrow{P} \gamma f(0)$ (Ergodic Theorem)

Results :

- $A_n \xrightarrow{d} -uN_1 - v\,N_2$, $(N_1,N_2)^T \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & EX_1 \\ EX_1 & EX_1^{\;2} \end{bmatrix})$.
- $EB_n \longrightarrow EY_1^{\;2} f(0) = (u,v)\,\Sigma\,(u,v)^T f(0)$.
- $B_n \xrightarrow{P} (u,v)\,\Sigma\,(u,v)^T f(0)$.

Conclude :

- $S_n(u,v) \xrightarrow{d} -uN_1 - v\,N_2 + (u,v)\,\Sigma\,(u,v)^T f(0)$.
- $(\hat{u}_n , \hat{v}_n)^T \xrightarrow{d} \Sigma^{-1}(N_1,N_2)^T / 2f(0) \sim N(\mathbf{0}, \Sigma^{-1}/ 4f^2(0))$
- $(n^{1/2} (\hat{\mu}-\mu_0), n^{1/2}(\hat{\phi}-\phi_0))^T \xrightarrow{d} N(\mathbf{0}, \Sigma^{-1}/ 4f^2(0))$

Example (AR(1)).

Data. $X_1, \ldots, X_n$

Model. $X_t = \mu_0 + \phi_0 X_{t-1} + Z_t$,

$|\phi_0| < 1$, $\{Z_t\} \sim \text{IID}(0,\sigma^2)$, $f(0)>0$.

LAD estimation:

Minimize

$$T_n(\mu,\phi) = \sum_{t=1}^{n} \left(|X_t - \mu - \phi X_{t-1}| - |Z_t|\right)$$

$$= \sum_{t=1}^{n} \left(|Z_t - (\mu-\mu_0) - (\phi - \phi_0) X_{t-1}| - |Z_t|\right)$$

Set $u=n^{1/2}(\mu-\mu_0)$, $v= n^{1/2}(\phi-\phi_0)$,

$$S_n(u,v) = \sum_{t=1}^{n} \left(|Z_t - un^{-1/2} - vn^{-1/2} X_{t-1}| - |Z_t|\right)$$

$$S_n(u,v) = \sum_{t=1}^{n} \left(|Z_t - un^{-1/2} - vn^{-1/2} X_{t-1}| - |Z_t|\right)$$

$$= \sum_{t=1}^{n} \left(|Z_t - n^{-1/2} Y_{t-1}| - |Z_t|\right)$$

$$(n^{-1/2} Y_{t-1} = un^{-1/2} + vn^{-1/2} X_{t-1})$$

$$= \sum_{t=1}^{n} - n^{-1/2} Y_{t-1} \operatorname{sgn}(Z_t)$$

$$+ 2 \sum_{t=1}^{n} (n^{-1/2} Y_{t-1} - Z_t)\{ I(0 < Z_t < n^{-1/2} Y_{t-1}) - I(n^{-1/2} Y_{t-1} < Z_t < 0)\}$$

$$=: A_n + B_n$$

Results :

- $A_n \xrightarrow{d} -uN_1 - v\,N_2$, $(N_1,N_2)^T \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & EX_1\\ EX_1 & EX_1^{\,2}\end{bmatrix}\right)$.
- $EB_n \longrightarrow EY_1^{\,2} f(0) = (u,v)\,\Sigma\,(u,v)^T f(0)$.
- $B_n \xrightarrow{P} (u,v)\,\Sigma\,(u,v)^T f(0)$.

Conclude :

- $S_n(u,v) \xrightarrow{d} -uN_1 - v\,N_2 + (u,v)\,\Sigma\,(u,v)^T f(0)$.
- $(\hat{u}_n , \hat{v}_n)^T \xrightarrow{d} \Sigma^{-1}(N_1,N_2)^T / 2f(0) \sim N(\mathbf{0}, \Sigma^{-1}/ 4f^2(0))$
- $(n^{1/2} (\hat{\mu}-\mu_0), n^{1/2}(\hat{\phi}-\phi_0))^T \xrightarrow{d} N(\mathbf{0}, \Sigma^{-1}/ 4f^2(0))$

Example (MA(1)).

Data. $X_1, \ldots, X_n$

Model. $X_t = Z_t + \theta Z_{t-1}$,

$|\theta_0| < 1$, $\{Z_t\} \sim \text{IID}(0,\sigma^2)$, $f(0)>0$.

LAD estimation:

Minimize

$$T_n(\theta) = \sum_{t=1}^{n} \left(|Z_t(\theta)| - |Z_t(\theta_0)|\right)$$

$$= \sum_{t=1}^{n} \left(|X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \cdots (-\theta)^{t-1} X_1| - |Z_t(\theta_0)|\right)$$

Set $u = n^{1/2}(\theta - \theta_0)$,

$$S_n(u) = T_n(\theta_0 + un^{-1/2})$$

$$= \sum_{t=1}^{n} \left(|Z_t(\theta_0 + un^{-1/2})| - |Z_t(\theta_0)| \right)$$

(Not a convex function of u!)

Linearize $Z_t(\theta_0 + un^{-1/2})$ to get

$$S_n(u) \sim \sum_{t=1}^{n} \left(|Z_t(\theta_0) + un^{-1/2} Z_t'(\theta_0)| - |Z_t(\theta_0)| \right)$$

where $-Z_t'(\theta_0)$ is the AR(1) process

$$Y_t = -\theta_0 Y_{t-1} + Z_t .$$

Result : Same limit result as in the AR(1) case, i.e.

$$n^{1/2}(\hat{\theta}_{LAD} - \theta_0) \xrightarrow{d} N / (\mathrm{Var}(Y_t)2f(0))$$

$$\sim N(0, (1-\theta^2) / (\sigma^2 4f^2(0)))$$

Linearized Version :

Initial estimate :

$$\hat{\theta}_0 = \theta_0 + O_p(n^{-1/2})$$

Objective Function:

$$T_n(\theta) = \sum_{t=1}^{n} \left(|Z_t(\hat{\theta}_0) + Z'_t(\hat{\theta}_0)(\theta - \hat{\theta}_0)| - |Z_t(\hat{\theta}_0)| \right).$$

Then

$$n^{1/2}(\hat{\theta}_L - \theta_0) \xrightarrow{d} N\left(0, (1-\theta^2) / (4f^2(0))\right),$$

where $\hat{\theta}_L = \text{argmin } T_n(\theta)$

Extensions :

ARMA Model : $\phi(B)X_t = \theta(B)Z_t$, $\{Z_t\} \sim \text{IID}(0,\sigma^2)$, $f(0)>0$.

Set $\boldsymbol{\beta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)^T$ and

$$\mathbf{v} = n^{1/2}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)$$

Then

(i) $$S_n(\mathbf{v}) := \sum_{t=1}^{n} \left(|Z_t(\boldsymbol{\beta}_0 + \mathbf{v}n^{-1/2})| - |Z_t(\boldsymbol{\beta}_0)|\right)$$

$$\xrightarrow{d} f(0)\mathbf{v}^T \Gamma_Q^{-1/2}\mathbf{v} + \mathbf{v}^T\mathbf{N}, \quad \mathbf{N} \sim N(\mathbf{0}, \Gamma_Q), \text{ (in } C(\mathbf{R}^{p+q}))$$

(ii) $$\hat{\mathbf{v}}_{LAD} = n^{1/2}(\hat{\boldsymbol{\beta}}_{LAD} - \boldsymbol{\beta}_0) \xrightarrow{d} \Gamma_Q^{-1}\mathbf{N}/(2f(0))$$

$$\sim N\left(\mathbf{0}, \Gamma_Q^{-1}/(4f^2(0))\right)$$

Note: $\Gamma_Q^{-1}\sigma^2$ is the limiting covariance matrix in Gaussian case.

ARMA Model With Stable Noise:

(Davis, Knight and Liu `92 for AR case, Davis `95 for ARMA.)

Model:

$\phi(B)X_t = \theta(B)Z_t$, $\{Z_t\}$ ~ IID symmetric stable(α), $0<\alpha<2$.

$$n^{1/\alpha}(\hat{\beta}_{LAD} - \beta_0) \xrightarrow{d} \mathbf{W}$$

In this case **both** A_n and B_n have random limits!

Least squares estimates: $(n / \ln n)^{1/\alpha}(\hat{\beta}_{LS} - \beta_0) \xrightarrow{d} \mathbf{V}$

Simulation results: Cauchy noise

	LS	LAD
AR(1) ϕ=.4	.395(.041)	.399(.015)
MA(1) θ=.8	.795(.049)	.794(.036)
ARMA(1,1) ϕ=.4	.399(.053)	.399(.026)
θ=.8	.781(.046)	.781(.033)

Linear Regression with ARMA Errors :

Model : $Y_t = A_t^T \alpha + X_t$,

where $\{X_t\}$ follows an ARMA process

$$\phi(B)X_t = \theta(B)Z_t\ ,\ \{Z_t\} \sim \text{IID}(0,\sigma^2),\ f(0)>0.$$

Assume $A_t^T = (a_{1t}, \ldots, a_{rt})^T$ satisfies Grenander's conditions:

(a) $|a_{jn}|/b_{jn} \xrightarrow{P} 0$

(b) $B_n^{-1} \sum_{t=1}^{n} A_t A_{t+j}^T B_n^{-1} \xrightarrow{P} \Gamma_A(j)$

(c) $b_{jn} \xrightarrow{P} \infty$

where $b_{jn}^2 = \|A_t\|^2$ and $B_n = \text{diag}(b_{1t}, \ldots, b_{rt})$.

Estimation: Let

$$X_t(\alpha) = \begin{cases} 0, & \text{if } t < 1, \\ Y_t - A_t^T \alpha, & \text{if } t > 0, \end{cases}$$

and for $\tau^T = (\alpha^T, \beta^T)$, define

$$Z_t(\tau) = \begin{cases} 0, & \text{if } t < 1, \\ \phi(B) X_t(\alpha) - \theta_1 Z_{t-1}(\tau) - \cdots - \theta_q Z_{t-q}(\tau), & \text{if } t > 0, \end{cases}$$

Minimize

$$S_n(\mathbf{u}, \mathbf{v}) = \sum_{t=1}^{n} \left(|Z_t(\tau_0 + n(\mathbf{u},\mathbf{v}))| - |Z_t(\tau_0)| \right)$$

where $n(\mathbf{u},\mathbf{v}) = \left((B_n^{-1} \mathbf{u})^T, n^{-1/2} \mathbf{v}^T \right)^T$.

Result :

- $S_n(\mathbf{u}, \mathbf{v}) \xrightarrow{d} S(\mathbf{u}, \mathbf{v})$ on $C(\mathbf{R}^{p+q+r})$

- $B_n(\hat{\alpha}_{LAD} - \alpha_0),\ n^{1/2}(\hat{\beta}_{LAD} - \beta_0) \xrightarrow{d} \Gamma^{-1}\mathbf{N}/(2f(0))$

$$\sim N(\mathbf{0}, \Gamma^{-1}/(4f^2(0))),$$

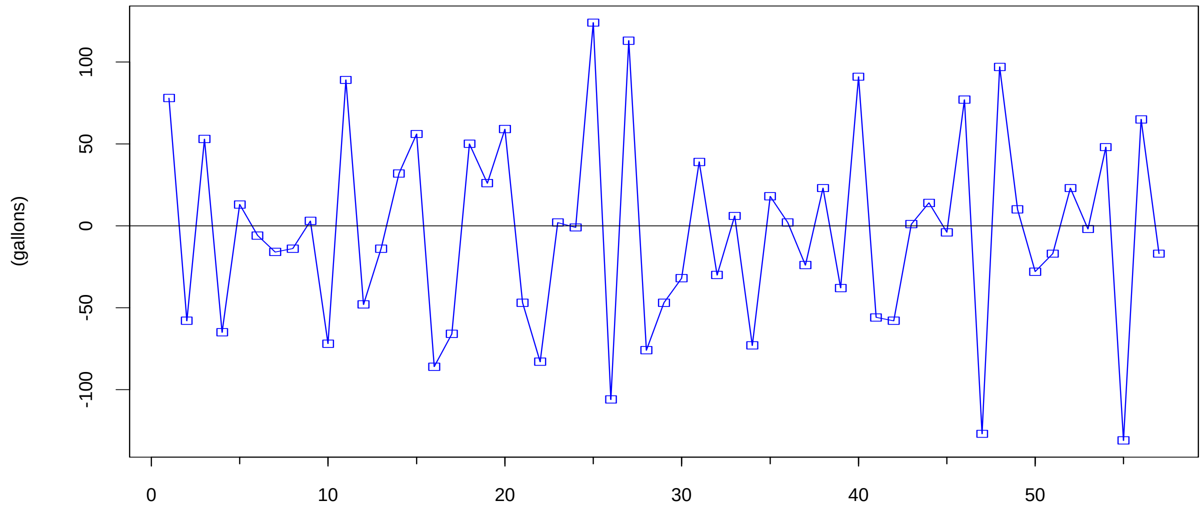
where

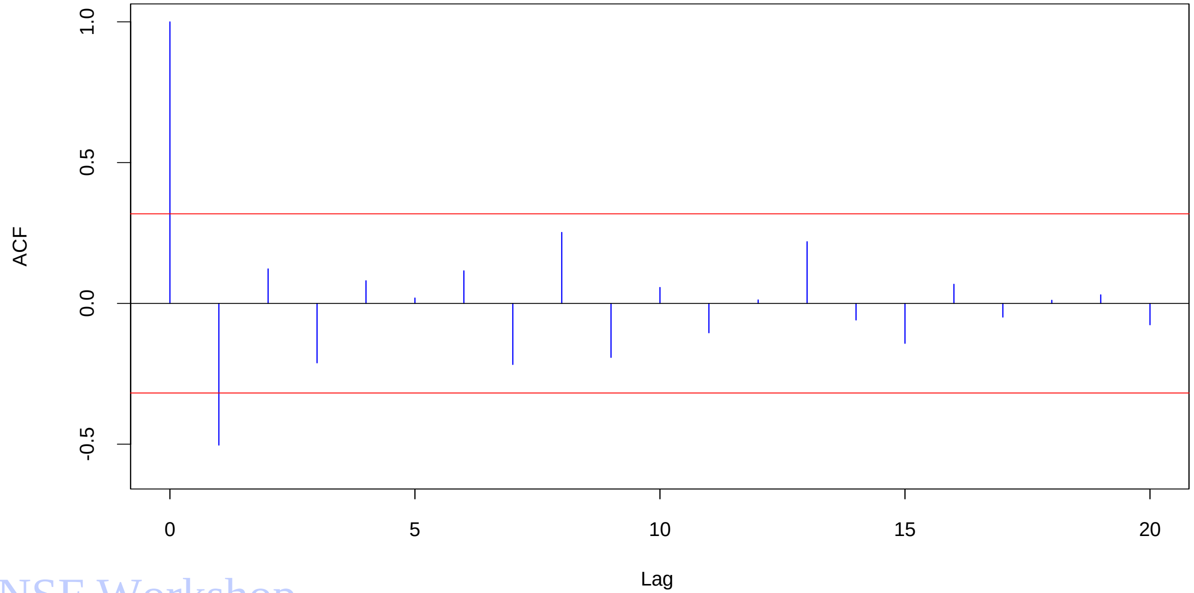
$$\Gamma = \begin{bmatrix} \Gamma_C & 0 \\ 0^T & \Gamma_Q \end{bmatrix}$$

and

$$\Gamma_C = \sum_{j,k} \pi_j \pi_k \Gamma_A(j-k), \quad \pi(B) = \theta_0(B) / \phi_0(B) .$$

An Example : (overshorts $Y_1, \ldots, Y_{57}$ from underground storage tank.)





Model. $Y_t = \mu + Z_t + \theta Z_{t-1}$

Problem. Estimate μ and construct a C.I.?

(Is $\mu < -5$ gallons/day?)

Estimation.

Estimates	Asymptotic Var
$\hat{\mu}_{MLE} = -4.87$	$(1+\theta)^2 \sigma^2/n = (1.408)^2$
$\hat{\theta}_{MLE} = -.849$	$(1-\theta^2)/n = (.070)^2$
$\hat{\mu}_{LAD} = -6.01$	$(1+\theta)^2/(4nf^2(0)) = (2.236)^2$
$\hat{\theta}_{LAD} = -.673$	$(1-\theta^2)/(\sigma^2\, 4n(f^2(0))) = (.106)^2$

Unit Root. If $\theta=1$, then

$\hat{\mu}_{MLE}$ is $AN(\mu_0, 12\sigma^2/n^3)$ (if θ is not estimated)

$\hat{\mu}_{MLE}$ is asymptotically non-normal (if θ is estimated)

Asymptotic distribution of $\hat{\mu}_{LAD}$ and $\hat{\theta}_{LAD}$ when $\theta = 1$?