Maximum Likelihood and R-Estimation for Allpass Time Series Models

Richard A. Davis

Department of Statistics
Colorado State University

http://www.stat.colostate.edu/~rdavis/lectures/

Joint work with
Beth Andrews, Colorado State University
Jay Breidt, Colorado State University

☞ Introduction
- properties of financial time series
- motivating example
- all-pass models and their properties

☞ Estimation
- likelihood approximation
- MLE, R-estimation, and LAD
- asymptotic results
- order selection

☞ Empirical results
- simulation

☞ Noninvertible MA processes
- preliminaries
- a two-step estimation procedure
- Microsoft trading volume

☞ Summary

Financial Time Series

- Log returns, $X_t = 100*(\ln(P_t) - \ln(P_{t-1}))$, of financial assets often exhibit:

  - heavy-tailed marginal distributions

    $$P(|X_1| > x) \sim C\, x^{-\alpha}, \quad 0 < \alpha < 4.$$

  - lack of serial correlation

    $\hat{\rho}_X(h)$ near 0 for all lags h > 0 (MGD sequence)

  - $|X_t|$ and X_t^2 have slowly decaying autocorrelations

    $\hat{\rho}_{|X|}(h)$ and $\hat{\rho}_{X^2}(h)$ converge to 0 *slowly* as $h \to \infty$

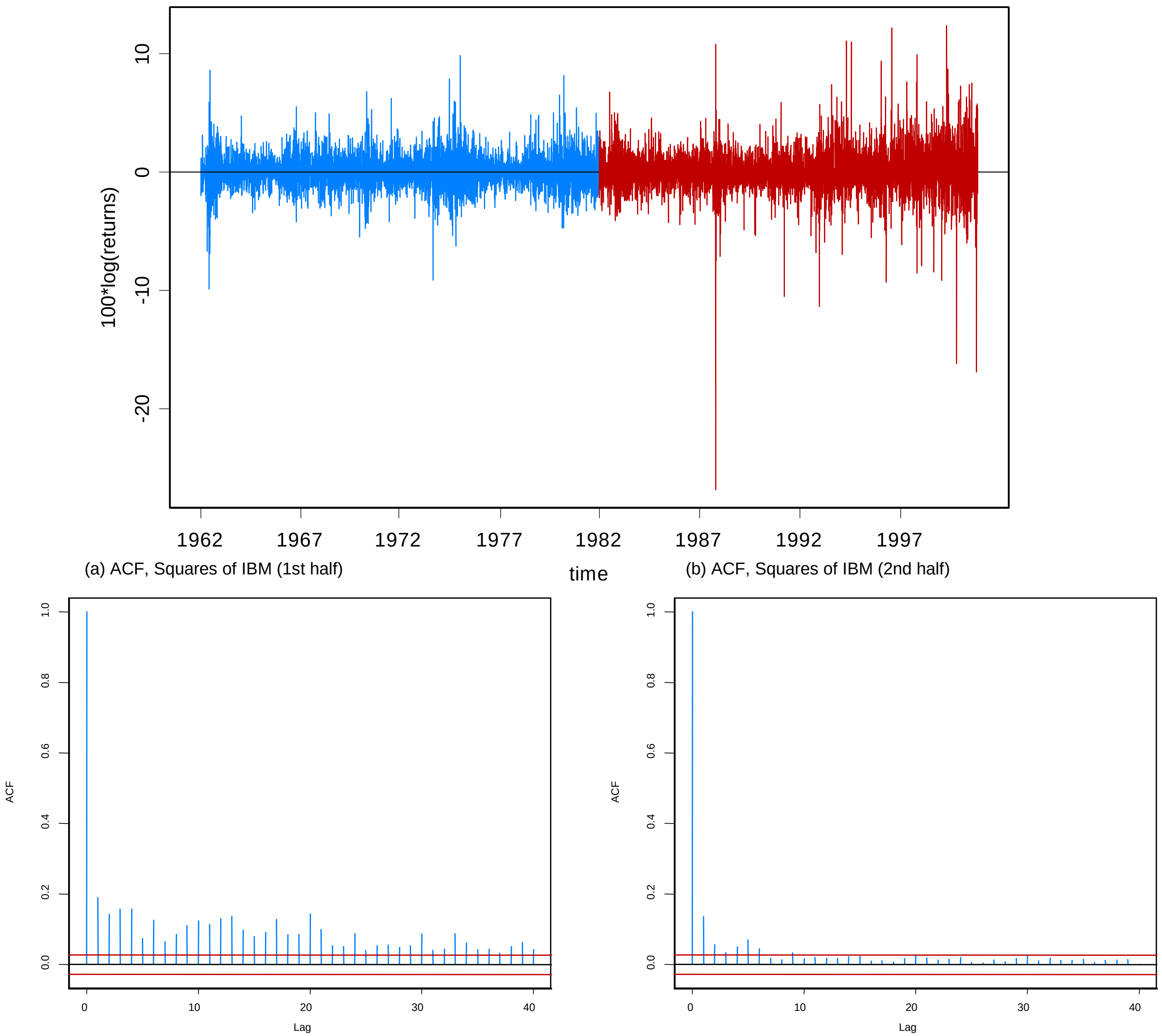
  - process exhibits 'stochastic volatility'

- Nonlinear models $X_t = \sigma_t Z_t$, $\{Z_t\} \sim \text{IID}(0,1)$

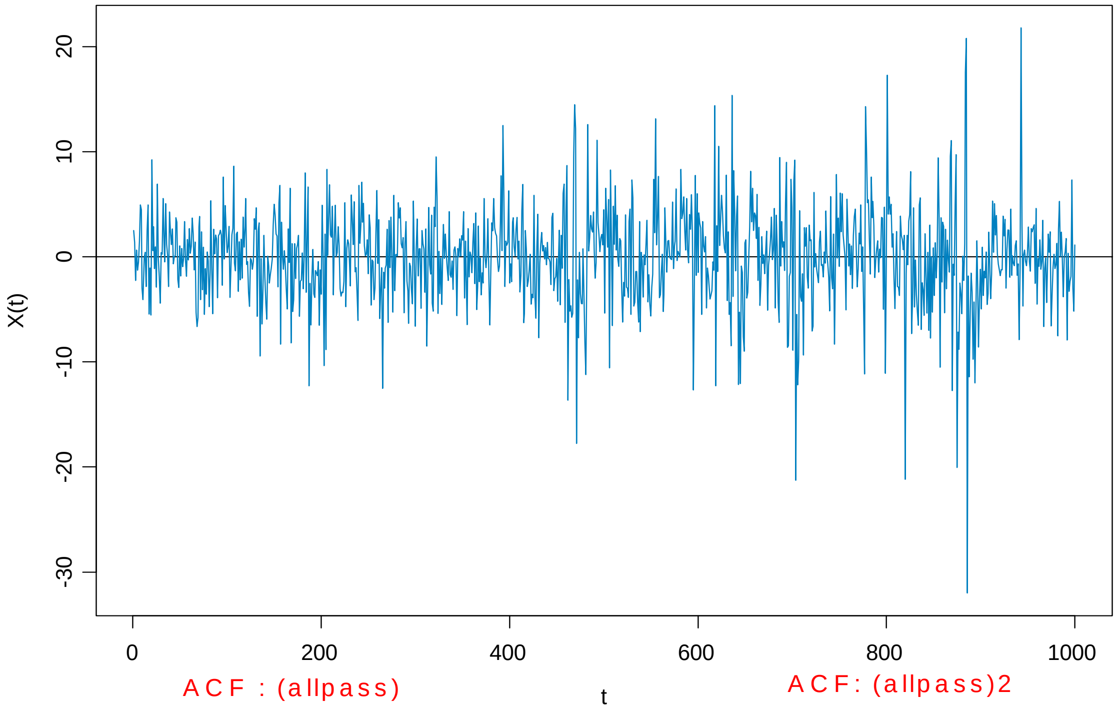
  - ARCH and its variants (Engle `82; Bollerslev, Chou, and Kroner 1992)

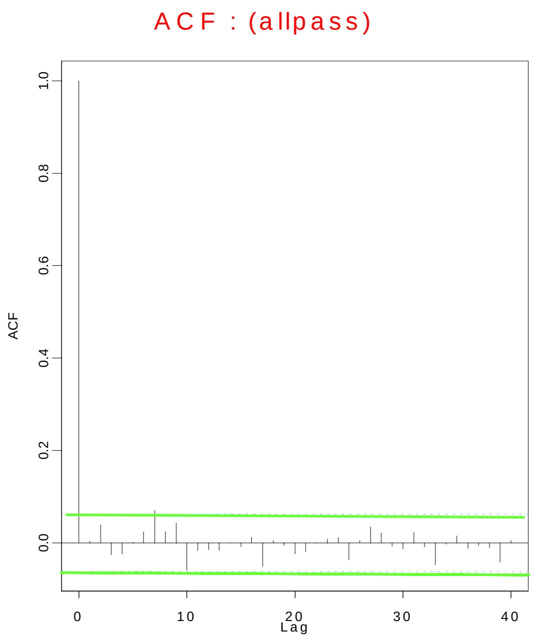
  - Stochastic volatility (Clark 1973; Taylor 1986)

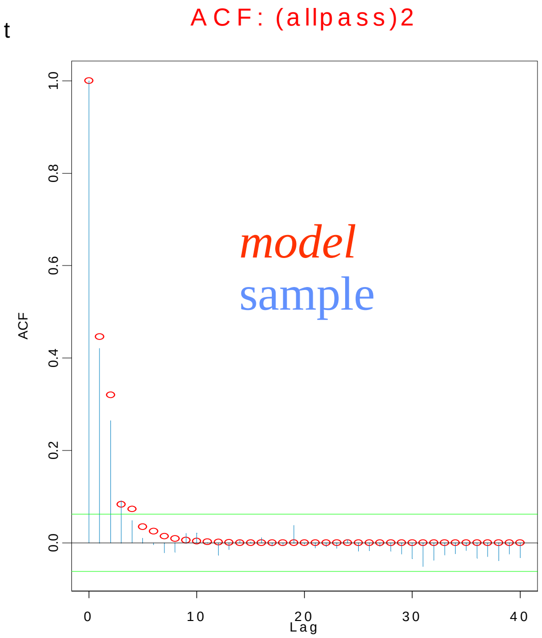
Log returns for IBM 1/3/62 – 11/3/00 (blue = 1961-1981)



All-pass model of order 2 (t3 noise)







All-pass Models

Causal AR polynomial: $\phi(z)=1-\phi_1 z - \cdots - \phi_p z^p$, $\phi(z) \neq 0$ for $|z|\leq 1$.

Define MA polynomial:

$$\theta(z) = -z^p\,\phi(z^{-1})/\phi_p = -(z^p - \phi_1 z^{p-1} - \cdots - \phi_p)/\,\phi_p$$

$\neq 0$ for $|z|\geq 1$ (MA polynomial is non-invertible).

Model for data $\{X_t\}$: $\phi(B)X_t = \theta(B)\,Z_t$, $\{Z_t\} \sim$ IID (non-Gaussian)

$$B^k X_t = X_{t-k}$$

Examples:

All-pass(1): $X_t - \phi\,X_{t-1} = Z_t - \phi^{-1}\,Z_{t-1}$, $|\,\phi\,| < 1$.

All-pass(2): $X_t - \phi_1\,X_{t-1} - \phi_2\,X_{t-2} = Z_t + \phi_1/\,\phi_2\,Z_{t-1} - 1/\,\phi_2\,Z_{t-2}$

Properties:

- causal, non-invertible ARMA with MA representation

$$X_t = \frac{B^p \phi(B^{-1})}{-\phi_p \phi(B)} Z_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

- uncorrelated (flat spectrum)

$$f_X(\omega) = \frac{\left|e^{-ip\omega}\right|^2 \left|\phi(e^{i\omega})\right|^2}{\phi_p^2 \left|\phi(e^{-i\omega})\right|^2} \frac{\sigma^2}{2\pi} = \frac{\sigma^2}{\phi_p^2 2\pi}$$

- zero mean
- data are dependent if noise is non-Gaussian (e.g. Breidt & Davis 1991).
- squares and absolute values are correlated.
- X_t is heavy-tailed if noise is heavy-tailed.

Estimation for All-Pass Models

☞ Second-order moment techniques do not work

- least squares
- Gaussian likelihood

☞ Higher-order cumulant methods

- Giannakis and Swami (1990)
- Chi and Kung (1995)

☞ Non-Gaussian likelihood methods

- likelihood approximation assuming known density
- quasi-likelihood

☞ Other

- LAD- least absolute deviation
- R-estimation (minimum dispersion)

Approximating the likelihood

Data: $(X_1, \ldots, X_n)$

Model: $$X_t = \phi_{01} X_{t-1} + \cdots + \phi_{0p} X_{t-p} - (Z_{t-p} - \phi_{01} Z_{t-p+1} - \cdots - \phi_{0p} Z_t) / \phi_{0r}$$

where ϕ_{0r} is the last non-zero coefficient among the ϕ_{0j}'s.

Noise: $z_{t-p} = \phi_{01} z_{t-p+1} + \cdots + \phi_{0p} z_t - (X_t - \phi_{01} X_{t-1} - \cdots - \phi_{0p} X_{t-p})$,

where $z_t = Z_t / \phi_{0r}$.

More generally define,

$$z_{t-p}(\phi) = \begin{cases} 0, & \text{if } t = n+p, \ldots, n+1, \\ \phi_1 z_{t-p+1}(\phi) + \cdots + \phi_p z_t(\phi) - \phi(B) X_t, & \text{if } t = n, \ldots, p+1. \end{cases}$$

Note: $z_t(\phi_0)$ is a close approximation to z_t (initialization error)

Assume that Z_t has density function f_σ and consider the vector

$$\mathbf{z} = (\underbrace{X_{1-p},\ldots,X_0, z_{1-p}(\phi),\ldots,z_0(\phi)}, \underbrace{z_1(\phi),\ldots,z_{n-p+1}(\phi),\ldots,z_n(\phi)})'$$

independent pieces

Joint density of $\mathbf{z}$:

$$h(\mathbf{z}) = h_1(X_{1-p},\ldots,X_0, z_{1-p}(\phi),\ldots,z_0(\phi))$$

$$\bullet\left(\prod_{t=1}^{n-p} f_\sigma(\phi_q z_t(\phi))\,|\,\phi_q\,|\right) h_2(z_{n-p+1}(\phi),\ldots,z_n(\phi)),$$

and hence the joint density of the data can be approximated by

$$h(\mathbf{x}) = \left(\prod_{t=1}^{n-p} f_\sigma(\phi_q z_t(\phi))\,|\,\phi_q\,|\right)$$

where $q=\max\{0 \le j \le p: \phi_j \ne 0\}$.

Log-likelihood:

$$L(\phi,\sigma) = -(n-p)\ln(\sigma/|\phi_q|) + \sum_{t=1}^{n-p} \ln f(\sigma^{-1}\phi_q z_t(\phi))$$

where $f_\sigma(z) = \sigma^{-1} f(z/\sigma)$.

Least absolute deviations: choose Laplace density

$$f(z) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}\,|z|)$$

and log-likelihood becomes

$$\text{constant} - (n-p)\ln\kappa - \sum_{t=1}^{n-p} \sqrt{2}\,|z_t(\phi)|/\kappa, \quad \kappa = \sigma/|\phi_q|$$

Concentrated Laplacian likelihood

$$l(\phi) = \text{constant} - (n-p)\ln\sum_{t=1}^{n-p} |z_t(\phi)|$$

Maximizing $l(\phi)$ is equivalent to minimizing the absolute deviations

$$m_n(\phi) = \sum_{t=1}^{n-p} |z_t(\phi)|.$$

Assumptions for MLE

☞ Assume $\{Z_t\}$ iid $f_\sigma(z)=\sigma^{-1}f(\sigma^{-1}z)$ with

- σ a scale parameter
- mean 0, variance σ^2
- further smoothness assumptions (integrability, symmetry, etc.) on f
- Fisher information:

$$\tilde{I} = \sigma^{-2}\int (f'(z))^2 / f(z)dz$$

Results

☞ Let $\gamma(h)$ = ACVF of AR model with AR poly $\phi_0(.)$ and

$$\Gamma_p = [\gamma(j-k)]_{j,k=1}^{p}$$

☞ $$\sqrt{n}(\hat{\phi}_{MLE} - \phi_0) \xrightarrow{D} N(0, \frac{1}{2(\sigma^2\tilde{I}-1)}\sigma^2\Gamma_p^{-1})$$

Further comments on MLE

Let $\alpha=(\phi_1, \ldots, \phi_p, \sigma/|\phi_p|, \beta_1, \ldots, \beta_q)$, where $\beta_1, \ldots, \beta_q$ are the parameters of pdf f.

Set

☞ $$\hat{I} = \sigma_0^{-2} \int (f'(z;\beta_0))^2 / f(z;\beta_0)\,dz$$

☞ $$\hat{K} = \alpha_{0,p+1}^{-2}\left\{\int z^2 (f'(z;\beta_0))^2 / f(z;\beta_0)\,dz - 1\right\}$$

☞ $$L = -\alpha_{0,p+1}^{-1} \int z \frac{f'(z;\beta_0)}{f(z;\beta_0)} \frac{\partial f(z;\beta_0)}{\partial \beta_0}\,dz$$

☞ $$I_f(\beta_0) = \int \frac{1}{f(z;\beta_0)} \frac{\partial f(z;\beta_0)}{\partial \beta_0} \frac{\partial f^T(z;\beta_0)}{\partial \beta_0}\,dz$$ (Fisher Information)

Under smoothness conditions on f wrt $\beta_1, \ldots, \beta_q$ we have

$$\sqrt{n}(\hat{\alpha}_{\text{MLE}} - \alpha_0) \xrightarrow{D} N(0, \Sigma^{-1}),$$

where

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{2(\sigma_0^2 \hat{I} - 1)} \sigma^2 \Gamma_p^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\hat{K} - L' I_f^{-1} L)^{-1} & -\hat{K}^{-1} L' (I_f - L \hat{K}^{-1} L')^{-1} \\ \mathbf{0} & -(I_f - L \hat{K}^{-1} L')^{-1} L \hat{K}^{-1} & (I_f - L \hat{K}^{-1} L')^{-1} \end{bmatrix}$$

Note: $\hat{\phi}_{\text{MLE}}$ is asymptotically independent of $\hat{\alpha}_{\text{p+1,MLE}}$ and $\hat{\beta}_{\text{MLE}}$

Asymptotic Covariance Matrix

- For LS estimators of AR(p):

$$\sqrt{n}(\hat{\phi}_{\mathrm{LS}} - \phi_0) \xrightarrow{D} N(0, \sigma^2 \Gamma_p^{-1})$$

- For LAD estimators of AR(p):

$$\sqrt{n}(\hat{\phi}_{\mathrm{LAD}} - \phi_0) \xrightarrow{D} N(0, \frac{1}{4\sigma^2 f^2(0)} \sigma^2 \Gamma_p^{-1})$$

- For LAD estimators of AP(p):

$$\sqrt{n}(\hat{\phi}_{\mathrm{LAD}} - \phi_0) \xrightarrow{D} N(0, \frac{\mathrm{Var}(|Z_1|)}{2(2\sigma^2 f_\sigma(0) - E|Z_1|)^2} \sigma^2 \Gamma_p^{-1})$$

- For MLE estimators of AP(p):

$$\sqrt{n}(\hat{\phi}_{\mathrm{MLE}} - \phi_0) \xrightarrow{D} N(0, \frac{1}{2(\sigma^2 \hat{I} - 1)} \sigma^2 \Gamma_p^{-1})$$

Laplace: (LAD=MLE)

$$\frac{\mathrm{Var}(|Z_1|)}{2(2\sigma^2 f_\sigma(0) - E|Z_1|)^2} = \frac{1}{2} = \frac{1}{2(\sigma^2 \hat{I} - 1)}$$

Students t_ν, $\nu > 2$:

LAD: $$\frac{(\nu-2)}{8\Gamma^2((\nu+1)/2)}(\pi(\nu-1)^2\Gamma^2(\nu/2) - 4(\nu-2)\Gamma^2((\nu+1)/2))$$

MLE: $$\frac{1}{2(\sigma^2\hat{I} - 1)} = \frac{(\nu-2)(\nu+3)}{12}$$

Student's t_3:

LAD: .7337

MLE: 0.5

ARE: .7337/.5=1.4674

R-Estimation:

Minimize the objective function

$$S(\phi) = \sum_{t=1}^{n-p} \varphi\left(\frac{t}{n-p+1}\right) z_{(t)}(\phi)$$

where $\{z_{(t)}(\phi)\}$ are the ordered $\{z_t(\phi)\}$, and the weight function φ satisfies:

- φ is differentiable and nondecreasing on $(0,1)$
- φ' is uniformly continuous
- $\varphi(x) = -\varphi(1-x)$

Remarks:

- $S(\phi) = \sum_{t=1}^{n-p} \varphi\left(\frac{R_t(\phi)}{n-p+1}\right) z_t(\phi)$
- For LAD, take $\varphi(x) = \begin{cases} -1, & 0 < x < 1/2, \\ 1, & 1/2 < x < 1. \end{cases}$

Assumptions for R-estimation

☞ Assume $\{Z_t\}$ iid with density function f (distr F)

- mean 0, variance σ^2

☞ Assume weight function φ is nondecreasing and continuously differentiable with $\varphi(x) = -\varphi(1-x)$

Results

☞ Set

$$\tilde{J} = \int_0^1 \varphi^2(s)ds, \quad \tilde{K} = \int_0^1 F^{-1}(s)\varphi(s)ds, \quad \tilde{L} = \int_0^1 f(F^{-1}(s))\varphi'(s)ds$$

☞ If $\sigma^2\tilde{L} > \tilde{K}$, then

$$\sqrt{n}(\hat{\phi}_R - \phi_0) \xrightarrow{D} N\left(0, \frac{\sigma^2\tilde{J} - \tilde{K}^2}{2(\sigma^2\tilde{L} - \tilde{K})^2}\sigma^2\Gamma_p^{-1}\right)$$

Further comments on R-estimation

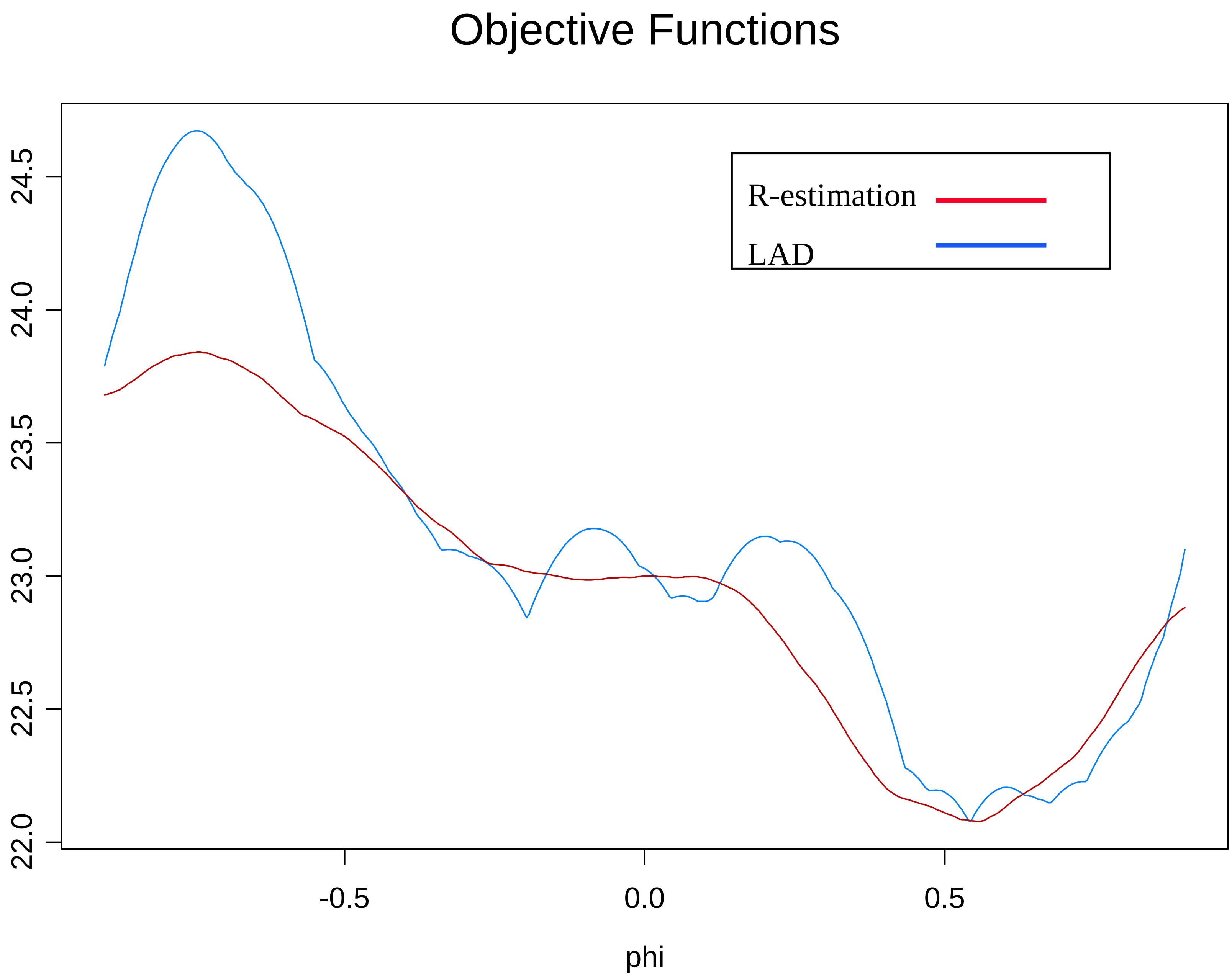
☞ $\varphi(x) = x - 1/2$ is called the Wilcoxon weight function

☞ By formally choosing $\varphi(x) = \begin{cases} -1, & 0 < x < 1/2, \\ 1, & 1/2 < x < 1. \end{cases}$ we obtain

$$\frac{\sigma^2 \tilde{J} - \tilde{K}^2}{2(\sigma^2 \tilde{L} - \tilde{K})^2} \sigma^2 \Gamma_p^{-1} = \frac{\mathrm{Var}(|Z_1|)}{2(2\sigma^2 f_\sigma(0) - E|Z_1|)^2} \sigma^2 \Gamma_p^{-1}.$$

That is R = LAD, asymptotically.

☞ The R-estimation objective function is smoother than the LAD-objective function and hence easier to minimize.



Objective Functions
R-estimation
LAD
24.5
24.0
23.5
23.0
22.5
22.0
-0.5
0.0
0.5
phi

Summary of asymptotics

☞ Maximum likelihood:

$$\sqrt{n}(\hat{\phi}_{\mathrm{MLE}} - \phi_0) \xrightarrow{D} N(0, \frac{1}{2(\sigma^2 \tilde{I} - 1)} \sigma^2 \Gamma_p^{-1})$$

☞ R-estimation

$$\sqrt{n}(\hat{\phi}_{\mathrm{R}} - \phi_0) \xrightarrow{D} N(0, \frac{\sigma^2 \tilde{J} - \tilde{K}^2}{2(\sigma^2 \tilde{L} - \tilde{K})^2} \sigma^2 \Gamma_p^{-1})$$

☞ Least absolute deviations:

$$\sqrt{n}(\hat{\phi}_{\mathrm{LAD}} - \phi_0) \xrightarrow{D} N(0, \frac{\mathrm{Var}(|Z_1|)}{2(2\sigma^2 f_\sigma(0) - E|Z_1|)^2} \sigma^2 \Gamma_p^{-1})$$

Laplace: (LAD=MLE)

R: $$\frac{\sigma^2 \tilde{J} - \tilde{K}^2}{2(\sigma^2 \tilde{L} - \tilde{K})^2} = \frac{5}{6}$$ (using $\varphi(x) = x - 1/2$, Wilcoxon)

LAD=MLE: 1/2

Students t_ν:

ν	LAD	**R**	MLE	**LAD/R**	MLE/R
3	.733	**.520**	.500	**1.411**	.962
6	6.22	**3.01**	3.00	**2.068**	.997
9	16.8	**7.15**	7.00	**2.354**	.980
12	32.6	**13.0**	12.5	**2.510**	.964
15	53.4	**20.5**	19.5	**2.607**	.952
20	99.6	**36.8**	34.5	**2.707**	.937
30	234	**83.6**	77.0	**2.810**	.921

Central Limit Theorem (R-estimation)

- Think of $\mathbf{u} = n^{1/2}(\phi - \phi_0)$ as an element of $\mathbb{R}^p$
- Define

$$S_n(\mathbf{u}) = \sum_{t=1}^{n-p}\left(\varphi\left(\frac{R_t(\phi_0 + n^{-1/2}\mathbf{u})}{n-p+1}\right)z_t(\phi_0 + n^{-1/2}\mathbf{u})\right) - \sum_{t=1}^{n-p}\left(\varphi\left(\frac{R_t(\phi_0\,)}{n-p+1}\right)z_t(\phi_0)\right),$$

where $R_t(\phi)$ is the rank of $z_t(\phi)$ among $z_1(\phi), \ldots, z_{n-p}(\phi)$.

- Then $S_n(\mathbf{u}) \to S(\mathbf{u})$ in distribution on $C(\mathbb{R}^p)$, where

$$S(\mathbf{u}) = |\phi_{0r}|^{-1}(\sigma^2\tilde{L} - \tilde{K})\mathbf{u}'\sigma^{-2}\Gamma_p\mathbf{u} + \mathbf{u}'\mathbf{N}, \ \ \mathbf{N} \sim N(\mathbf{0}, 2(\sigma^2\tilde{J} - \tilde{K}^2)\,|\phi_{0r}|^{-2}\,\sigma^{-2}\Gamma_p),$$

- Hence,

$$\begin{aligned}\arg\min S_n(\mathbf{u}) &= n^{1/2}(\hat{\phi}_R - \phi_0)\\ &\underset{D}{\to} \arg\min S(\mathbf{u})\\ &= -\frac{|\phi_{0r}|}{2(\sigma^2\tilde{L} - \tilde{K})}\sigma^2\Gamma_p^{-1}\mathbf{N} \sim N\left(\mathbf{0}, \frac{\sigma^2\tilde{J} - \tilde{K}^2}{2(\sigma^2\tilde{L} - \tilde{K})^2\,|\phi_{0r}|^2}\sigma^2\Gamma_p^{-1}\right)\end{aligned}$$

Main ideas (R-estimation)

- Define

$$\tilde{S}_n(\mathbf{u}) = \sum_{t=1}^{n-p} \varphi(F_z(z_t)) z_t(\phi_0 + n^{-1/2}\mathbf{u}) - \sum_{t=1}^{n-p} \varphi(F_z(z_t)) z_t(\phi_0),$$

where F_z is the df of z_t.

- Using a Taylor series, we have

$$\tilde{S}_n(\mathbf{u}) \sim n^{-1/2}\mathbf{u}' \sum_{t=1}^{n-p} \left(\varphi(F_z(z_t)) \frac{\partial z_t(\phi_0)}{\partial \phi} \right) + 2^{-1} n^{-1} \mathbf{u}' \sum_{t=1}^{n-p} \left(\varphi(F_z(z_t)) \frac{\partial^2 z_t(\phi_0)}{\partial \phi \partial \phi'} \right) \mathbf{u}$$

$$\underset{D}{\rightarrow} \mathbf{u}'\mathbf{N} - \mathbf{u}' \tilde{K} \,|\phi_{0r}|^{-1} \sigma^{-2} \Gamma_p \mathbf{u}$$

- Also,

$$S_n(\mathbf{u}) - \tilde{S}_n(\mathbf{u}) = \mathbf{u}' \sigma^2 \tilde{L} \sigma^{-2} \Gamma_p \mathbf{u} / |\phi_{0r}| + o_p(1).$$

- Hence

$$S_n(\mathbf{u}) \underset{D}{\rightarrow} |\phi_{0r}|^{-1} (\sigma^2 \tilde{L} - \tilde{K}) \mathbf{u}' \sigma^{-2} \Gamma_p \mathbf{u} + \mathbf{u}'\mathbf{N}, \;\; \mathbf{N} \sim N(\mathbf{0}, 2(\sigma^2 \tilde{J} - \tilde{K}^2) |\phi_{0r}|^{-2} \sigma^{-2} \Gamma_p).$$

Order Selection:

Partial ACF From the previous result, if true model is of order r and fitted model is of order p > r, then

$$n^{1/2}\hat{\phi}_{p,LAD} \to N\left(0, \frac{\mathrm{Var}(|Z_1|)}{2(2\sigma^2 f_\sigma(0) - E|Z_1|)^2}\right)$$

where $\hat{\phi}_{p,LAD}$ is the pth element of $\hat{\phi}_{LAD}$.

Procedure:

1. Fit high order (P-th order), obtain residuals and estimate scalar,

$$\theta^2 = \frac{\mathrm{Var}(|Z_1|)}{2(2\sigma^2 f_\sigma(0) - E|Z_1|)^2}$$

by empirical moments of residuals and density estimates.

2. Fit AP models of order p=1,2, . . . , P via LAD and obtain p-th coefficient $\hat{\phi}_{p,p}$ for each.

3. Choose model order r as the smallest order beyond which the estimated coefficients are statistically insignificant.

Note: Can replace $\hat{\phi}_{p,p}$ with $\hat{\phi}_{p,\mathrm{MLE}}$ if using MLE. In this case for p > r

$$n^{1/2}\hat{\phi}_{p,MLE} \to N(0, \frac{1}{2(\sigma^2 \hat{I} - 1)}).$$

AIC: 2*p* or not 2*p*?

- An approximately unbiased estimate of the Kullback-Leiber index of fitted to true model:

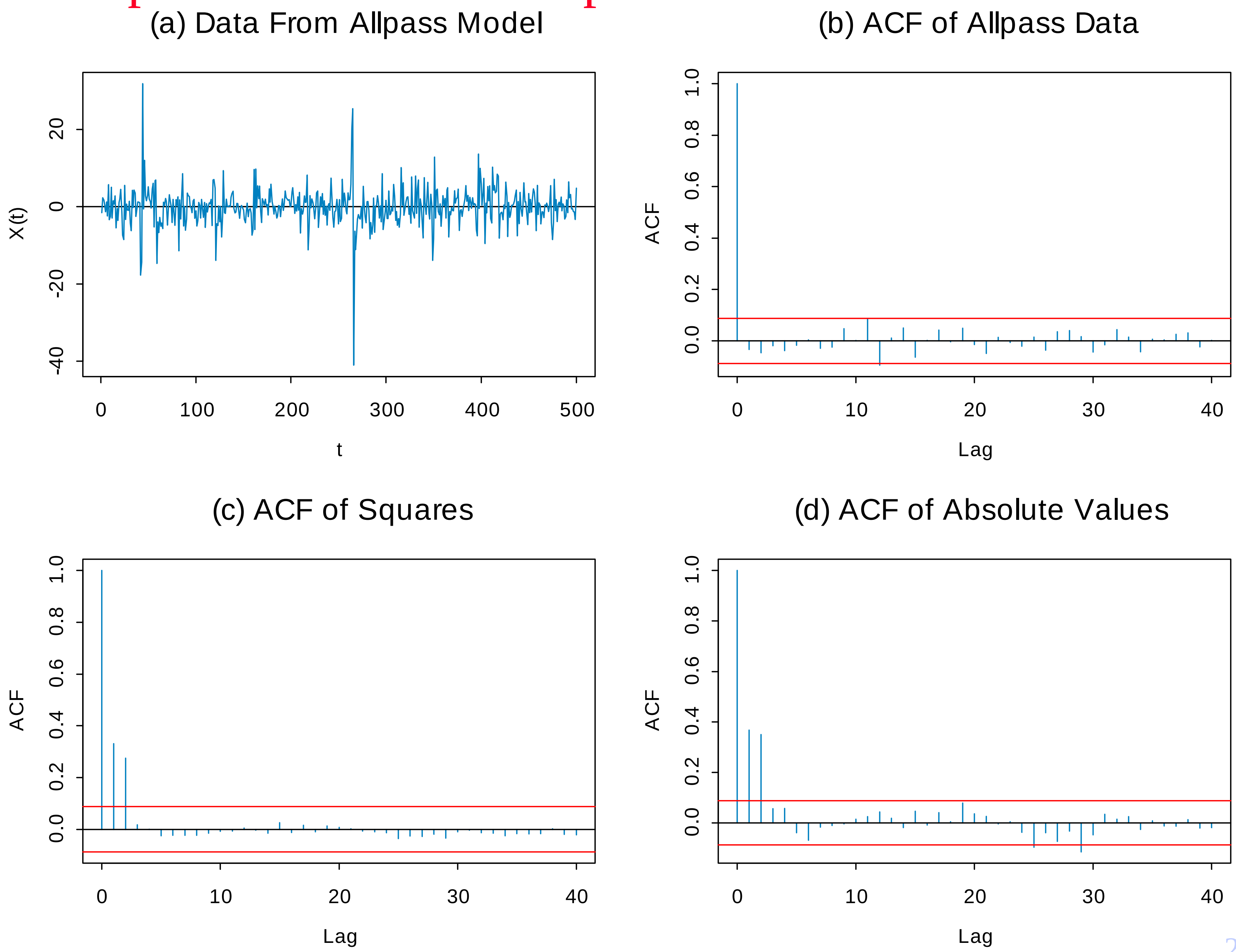
$$AIC(p) := -2L_X(\hat{\phi}, \hat{\kappa}) + \frac{\mathrm{Var}(|Z_1|)}{(2\sigma^2 f_\sigma(0) - E|Z_1|)^2}\left(\frac{2\sigma^2 f_\sigma(0)}{E|Z_1|} - 1\right)p$$

- Penalty term for Laplace case:

$$\frac{\mathrm{Var}(|Z_1|)}{(2\sigma^2 f_\sigma(0) - E|Z_1|)^2}\left(\frac{2\sigma^2 f_\sigma(0)}{E|Z_1|} - 1\right)p = p$$

- Penalty term can be estimated from the data.

Sample realization of all-pass of order 2



Simulation results:

- 1000 replicates of all-pass models
- model order parameter value

model order	parameter value
1	$\phi_1 = .5$
2	$\phi_1 = .3,\ \phi_2 = .4$

- noise distribution is t with 3 d.f.
- sample sizes n=500, 5000
- estimation method is LAD

To guard against being trapped in local minima, we adopted the following strategy.

- 250 random starting values were chosen at *random*. For model of order p, k-th starting value was computed recursively as follows:

1. Draw $\phi_{11}^{(k)}, \phi_{22}^{(k)}, \ldots, \phi_{pp}^{(k)}$ iid uniform (-1,1).
2. For j=2, …, p, compute

$$\begin{bmatrix} \phi_{j1}^{(k)} \\ \vdots \\ \phi_{j,j-1}^{(k)} \end{bmatrix} = \begin{bmatrix} \phi_{j-1,1}^{(k)} \\ \vdots \\ \phi_{j-1,j-1}^{(k)} \end{bmatrix} - \phi_{jj}^{(k)} \begin{bmatrix} \phi_{j-1,j-1}^{(k)} \\ \vdots \\ \phi_{j-1,1}^{(k)} \end{bmatrix}$$

- Select top 10 based on minimum function evaluation.

- Run Hooke and Jeeves with each of the 10 starting values and choose best optimized value.

	Asymptotic		Empirical			
N	mean	**std dev**	mean	**std dev**	%coverage	rel eff*
500	ϕ_1=.5	**.0332**	.4979	**.0397**	94.2	11.8
5000	ϕ_1=.5	**.0105**	.4998	**.0109**	95.4	9.3

	Asymptotic		Empirical		
N	mean	**std dev**	mean	**std dev**	%coverage
500	ϕ_1=.3	**.0351**	.2990	**.0456**	92.5
	ϕ_2=.4	**.0351**	.3965	**.0447**	92.1
5000	ϕ_1=.3	**.0111**	.3003	**.0118**	95.5
	ϕ_2=.4	**.0111**	.3990	**.0117**	94.7

*Efficiency relative to maximum absolute residual kurtosis:

$$\left| \frac{1}{n-p} \sum_{t=1}^{n-p} \left(\frac{z_t(\phi)}{v_2^{1/2}} \right)^4 - 3 \right|, \quad v_2 = \frac{1}{n-p} \sum_{t=1}^{n-p} (z_t(\phi) - \overline{z}(\phi))^2$$

MLE Simulations Results using t-distr(3.0)

	Asymptotic		Empirical		
N	mean	**std dev**	mean	**std dev**	%coverage
500	ϕ_1=.5	**.0274**	.4971	**.0315**	93.0
	ν=3.0	**.4480**	3.112	**.5008**	95.8
5000	ϕ_1=.5	**.0087**	.4997	**.0091**	93.4
	ν=3.0	**.1417**	3.008	**.1533**	94.0

	Asymptotic		Empirical		
N	mean	**std dev**	mean	**std dev**	%coverage
500	ϕ_1=.3	**.0290**	.2993	**.0345**	90.6
	ϕ_2=.4	**.0290**	.3964	**.0350**	90.1
	ν=3.0	**.4480**	3.079	**.4722**	94.8
5000	ϕ_1=.3	**.0092**	.2999	**.0095**	94.0
	ϕ_2=.4	**.0092**	.3999	**.0094**	94.6
	ν=3.0	**.1417**	3.008	**.1458**	95.2

R-Estimator: Minimize the objective fcn

$$S(\phi) = \sum_{t=1}^{n-p} \left(\frac{t}{n-p+1} - \frac{1}{2} \right) z_{(t)}(\phi)$$

where $\{z_{(t)}(\phi)\}$ are the ordered $\{z_t(\phi)\}$.

		Empirical		Empirical LAD	
N		mean	**std dev**	mean	**std dev**
500	ϕ_1=.5	.4978	**.0315**	.4979	**.0397**
5000	ϕ_1=.5	.4997	**.0094**	.4998	**.0109**
500	ϕ_1=.3	.2988	**.0374**	.2990	**.0456**
	ϕ_2=.4	.3957	**.0360**	.3965	**.0447**
5000	ϕ_1=.3	.3007	**.0101**	.3003	**.0118**
	ϕ_2=.4	.3993	**.0104**	.3990	**.0117**

Noninvertible MA models with heavy tailed noise

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q},$$

a. $\{Z_t\} \sim \text{IID}(\alpha)$ with Pareto tails

b. $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$

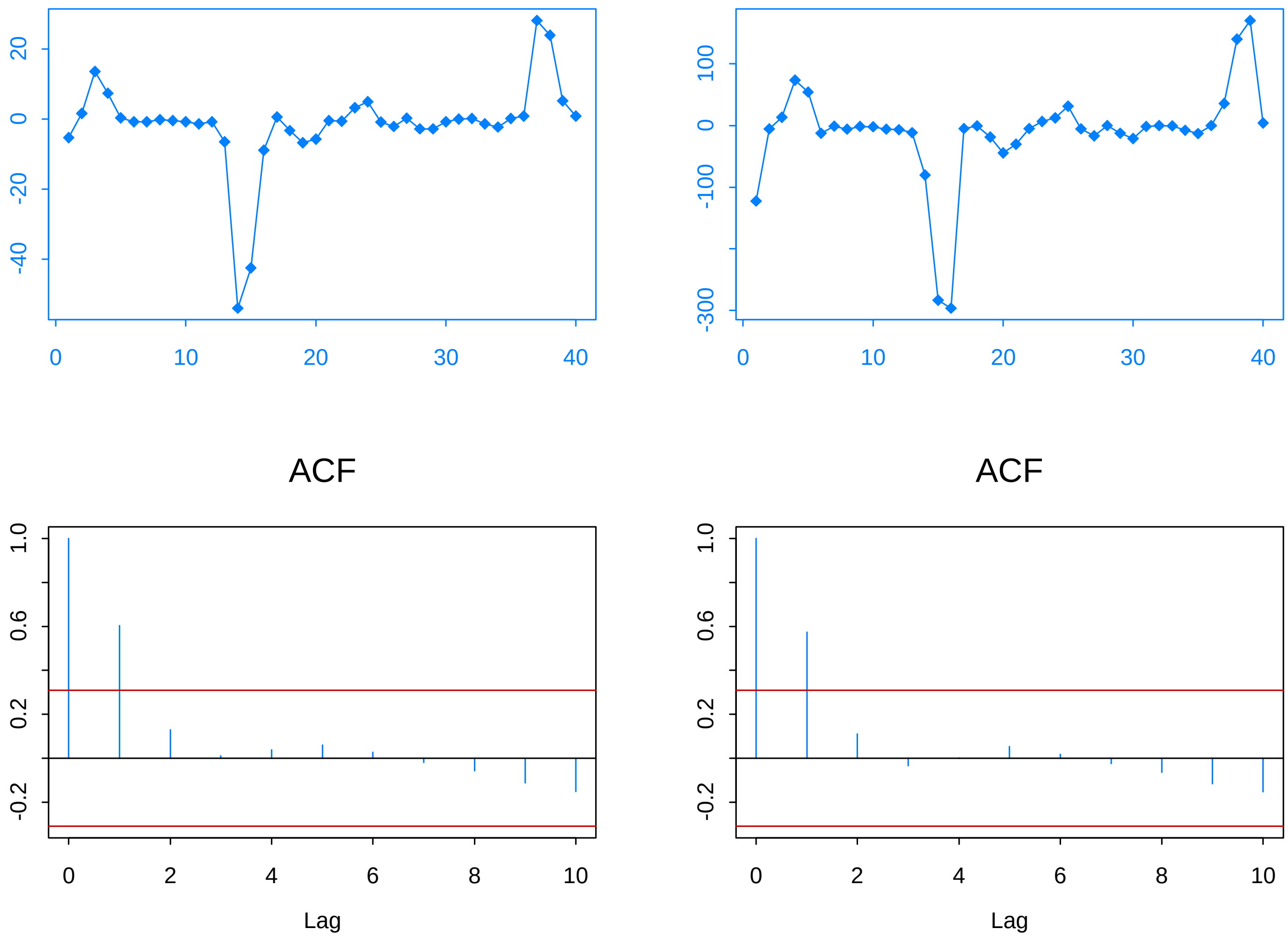
No zeros inside the unit circle $\Rightarrow$ invertible

Some zero(s) inside the unit circle $\Rightarrow$ noninvertible

Realizations of an invertible and noninvertible MA(2) processes

Model: $X_t = \theta_*(B)\, Z_t$, $\{Z_t\} \sim \text{IID}(\alpha = 1)$, where

$\theta_i(B) = (1 + 1/2B)(1 + 1/3B)$ and $\theta_{ni}(B) = (1 + 2B)(1 + 3B)$



Application of all-pass to noninvertible MA model fitting

Suppose $\{X_t\}$ follows the noninvertible MA model

$$X_t = \theta_i(B)\,\theta_{ni}(B)\,Z_t\,, \quad \{Z_t\} \sim \text{IID}.$$

Step 1: Let $\{U_t\}$ be the residuals obtained by fitting a purely invertible MA model, i.e.,

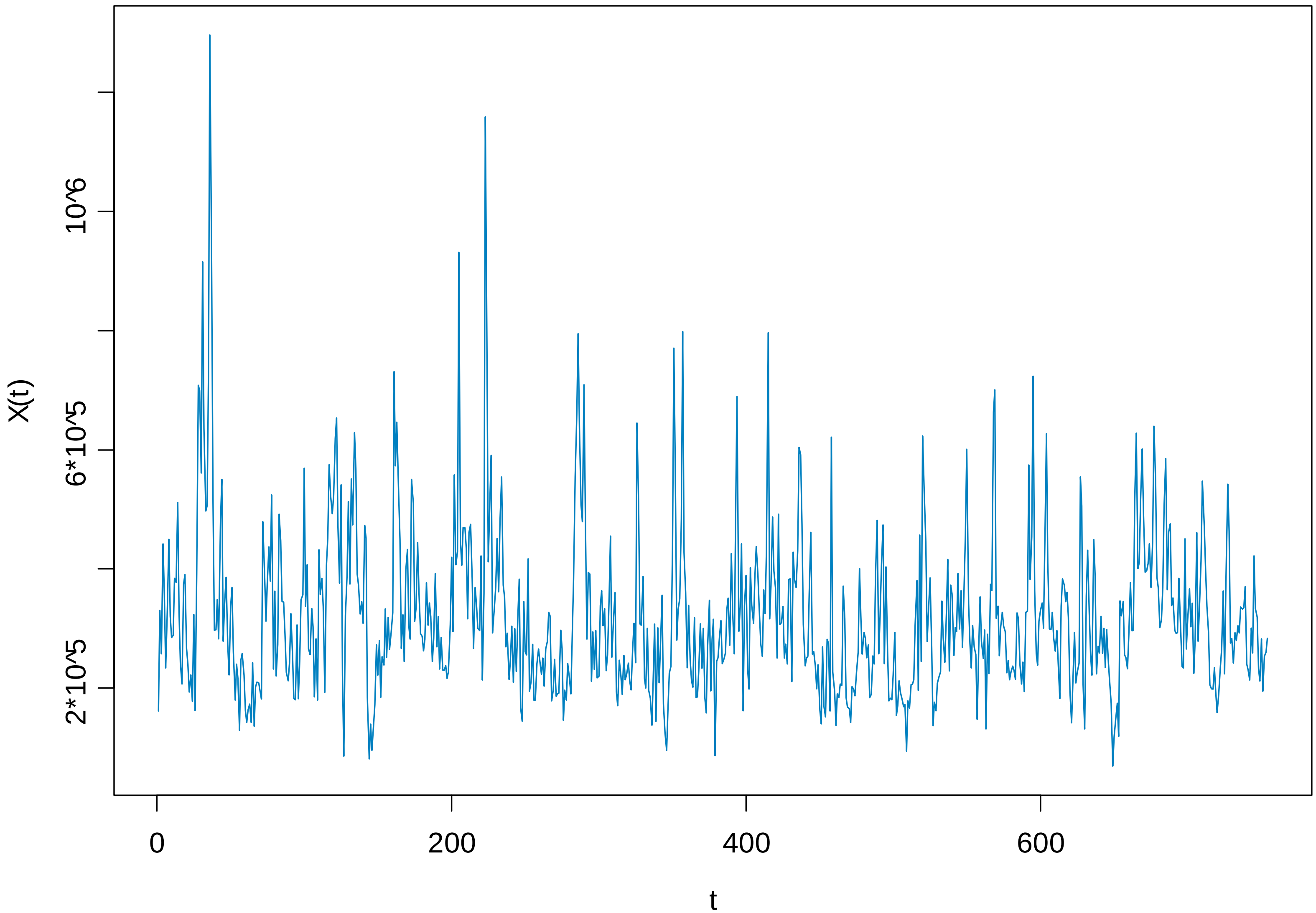
$$\begin{aligned} X_t &= \hat{\theta}(B)U_t \\ &\approx \theta_i(B)\tilde{\theta}_{ni}(B)U_t, \quad (\tilde{\theta}_{ni} \text{ is the invertible version of } \theta_{ni}). \end{aligned}$$

So $$U_t \approx \frac{\theta_{ni}(B)}{\tilde{\theta}_{ni}(B)} Z_t$$

Step 2: Fit a purely causal AP model to $\{U_t\}$

$$\tilde{\theta}_{ni}(B)U_t = \theta_{ni}(B)Z_t.$$

Volumes of Microsoft (MSFT) stock traded over 755 transaction days (6/3/96 to 5/28/99)



Analysis of MSFT:

Step 1: Log(volume) follows MA(4).

$X_t = (1+.513B+.277B^2+.270B^3+.202B^4)\, U_t$ (invertible MA(4))

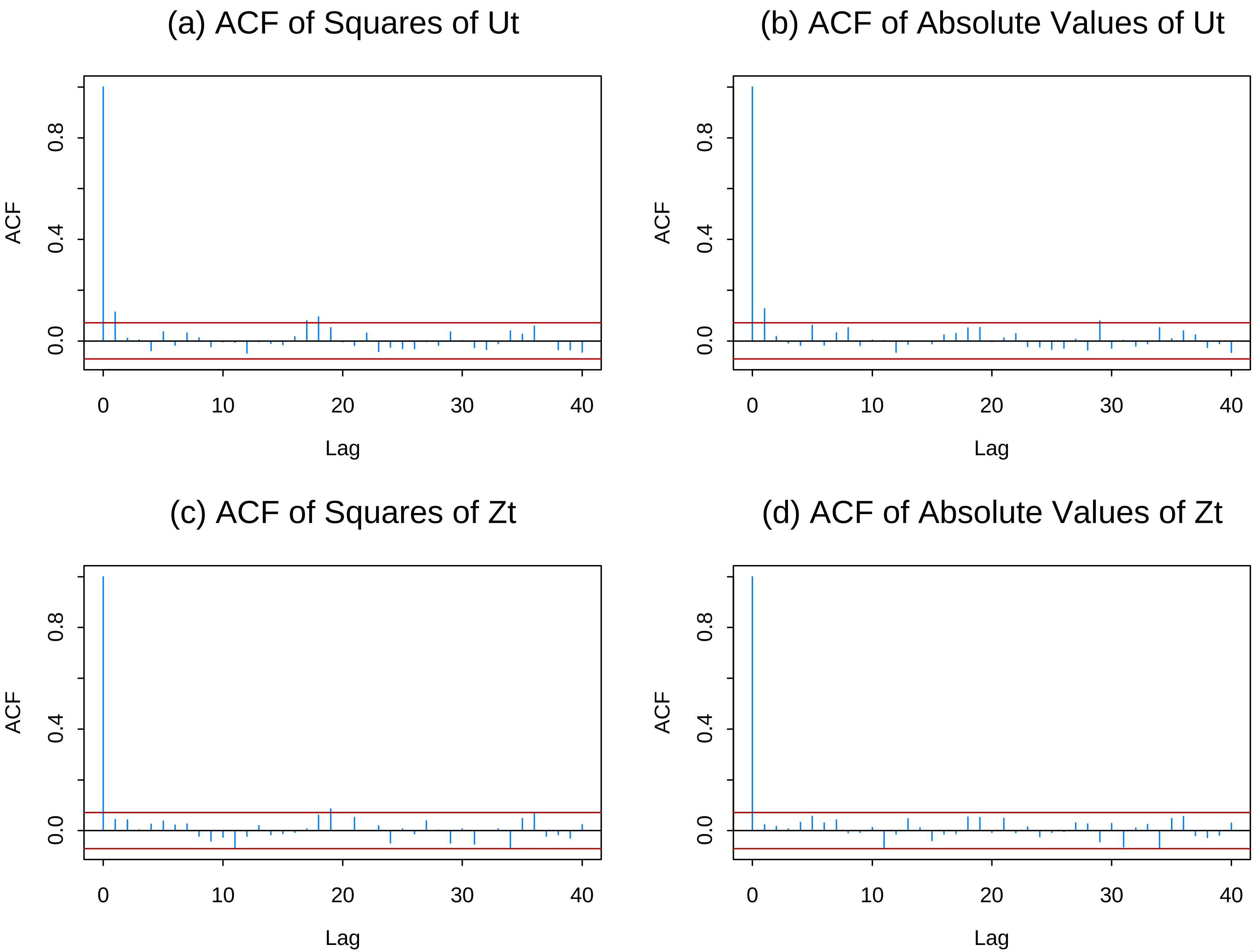
Step 2: All-pass model of order 4 fitted to $\{U_t\}$ using MLE (t-dist):

$$(1-.628B+-.229B^2+.131B^3-.202B^4)U_t$$
$$=(1-.649B+1.135B^2+3.116B^3-4.960B^4)Z_t. \; (\hat{\nu}=6.26)$$

(Model using R-estimation is nearly the same.)

Conclude that $\{X_t\}$ follows a noninvertible MA(4) which after refitting has the form:

$X_t = (1+1.34B+1.374B^2+2.54B^3+4.96B^4)\, Z_t$, $\{Z_t\}$~IID t(6.3)


(a) ACF of Squares of Ut
(b) ACF of Absolute Values of Ut
(c) ACF of Squares of Zt
(d) ACF of Absolute Values of Zt
ACF
0.0
0.4
0.8
Lag
0
10
20
30
40

Summary: Microsoft Trading Volume

☞ Two-step fit of noninvertible MA(4):

- invertible MA(4): residuals not iid
- causal AP(4); residuals iid

☞ Direct fit of purely noninvertible MA(4):

$$(1+1.34B+1.374B^2+2.54B^3+4.96B^4)$$

☞ For MCHP, invertible MA(4) fits.

Summary

☞ All-pass models and their properties

- linear time series with "nonlinear" behavior

☞ Estimation

- likelihood approximation
- MLE, LAD, R-estimation
- order selection

☞ Emprirical results

- simulation study

☞ Noninvertible moving average processes

- two-step estimation procedure using all-pass
- noninvertible MA(4) for Microsoft trading volume

Further Work

☞ Least absolute deviations

- further simulations
- order selection
- heavy-tailed case
- other smooth objective functions (e.g., min dispersion)

☞ Maximum likelihood

- Gaussian mixtures
- simulation studies
- applications

☞ Noninvertible moving average modeling

- initial estimates from two-step all-pass procedure
- adaptive procedures