MURI Research Update

Richard A. Davis

Jingjing Zou, postdoc 6/2015-present
Phyllis Wan, RA 6/2015-present

Columbia University

October 16, 2015

Summary of Research from Last year

1. Applications of Tests for Independence (with Phyllis and Thomas Mikosch)
2. Independent Component Analysis (with Jingjing, John, Sid and Alparsan)
3. Random Matrix Theory with Heavy Tails (with Mikosch and students)

1. Distance Covariance: A measure of dependence

Introduction

- Random vectors $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$,

$$X \perp Y \iff \phi_{X,Y} = \phi_X \phi_Y.$$

where ϕ denotes the characteristic function.

1. Distance Covariance: A measure of dependence

Introduction

- Random vectors $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$,

$$X \perp Y \iff \phi_{X,Y} = \phi_X \phi_Y.$$

  where ϕ denotes the characteristic function.

- Define distance covariance w.r.t. weight function $w(s,t)$

$$\mathcal{V}^2(X, Y; w) = \int_{\mathbb{R}^{p+q}} |\phi_{X,Y}(t,s) - \phi_X(t)\phi_Y(s)|^2 w(t,s) dtds$$

1. Distance Covariance: A measure of dependence

Introduction

- Random vectors $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$,

$$X \perp Y \iff \phi_{X,Y} = \phi_X \phi_Y.$$

  where ϕ denotes the characteristic function.

- Define distance covariance w.r.t. weight function $w(s,t)$

$$\mathcal{V}^2(X,Y;w) = \int_{\mathbb{R}^{p+q}} |\phi_{X,Y}(t,s) - \phi_X(t)\phi_Y(s)|^2 w(t,s)dtds$$

- Empirical version based on data (time series?) $(X_1, Y_1), \ldots, (X_n, Y_n)$

$$\hat{\mathcal{V}}_n^2(X,Y;w) = \int_{\mathbb{R}^{p+q}} |\hat{\phi}_{X,Y}(t,s) - \hat{\phi}_X(t)\hat{\phi}_Y(s)|^2 w(t,s)dtds$$

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s, t)$

- For special choices of $w(s, t)$, the empirical distance covariance can be expressed as

$$\begin{aligned}\frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)h_2(Y_k, Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r, Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k, X_l)h_2(Y_k, Y_r) \qquad (1)\end{aligned}$$

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s,t)$

- For special choices of $w(s,t)$, the empirical distance covariance can be expressed as

$$\begin{aligned}\frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)h_2(Y_k,Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r,Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k,X_l)h_2(Y_k,Y_r) \qquad (1)\end{aligned}$$

- Székely et al.: $w(t,s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s, t)$

- For special choices of $w(s, t)$, the empirical distance covariance can be expressed as

$$\begin{aligned} \frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)h_2(Y_k, Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r, Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k, X_l)h_2(Y_k, Y_r) \qquad (1) \end{aligned}$$

- Székely et al.: $w(t, s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.
  - $h_1(x, x') = |x - x'|_p^{\alpha}$, $h_2(y, y') = |y - y'|_q^{\alpha}$

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s, t)$

- For special choices of $w(s, t)$, the empirical distance covariance can be expressed as

$$\begin{aligned} \frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)h_2(Y_k, Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r, Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k, X_l)h_2(Y_k, Y_r) \qquad (1) \end{aligned}$$

- Székely et al.: $w(t, s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.
  - $h_1(x, x') = |x - x'|_p^\alpha$, $h_2(y, y') = |y - y'|_q^\alpha$
  - Requires $E|X|_p^\alpha + E|Y|_q^\alpha + E|X|_p^\alpha|Y|_q^\alpha < \infty$

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s,t)$

- For special choices of $w(s,t)$, the empirical distance covariance can be expressed as

$$\begin{aligned} \frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)h_2(Y_k,Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r,Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k,X_l)h_2(Y_k,Y_r) \qquad (1) \end{aligned}$$

- Székely et al.: $w(t,s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.
  - $h_1(x,x') = |x-x'|_p^\alpha$, $h_2(y,y') = |y-y'|_q^\alpha$
  - Requires $E|X|_p^\alpha + E|Y|_q^\alpha + E|X|_p^\alpha|Y|_q^\alpha < \infty$
- Take $w(s,t) = f_S(s)f_T(t)$ to be the product of two probability densities symmetric about the origin

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s,t)$

- For special choices of $w(s,t)$, the empirical distance covariance can be expressed as

$$\begin{aligned}\frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)h_2(Y_k,Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r,Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k,X_l)h_2(Y_k,Y_r) \qquad (1)\end{aligned}$$

- Székely et al.: $w(t,s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.
  - $h_1(x,x') = |x - x'|_p^{\alpha}$, $h_2(y,y') = |y - y'|_q^{\alpha}$
  - Requires $E|X|_p^{\alpha} + E|Y|_q^{\alpha} + E|X|_p^{\alpha}|Y|_q^{\alpha} < \infty$
- Take $w(s,t) = f_S(s)f_T(t)$ to be the product of two probability densities symmetric about the origin
  - $h_1(x,x') = \phi_S(x - x')$, $h_2(y,y') = \phi_T(y - y')$.

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s,t)$

- For special choices of $w(s,t)$, the empirical distance covariance can be expressed as

$$\begin{aligned}\frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)h_2(Y_k,Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k,X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r,Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k,X_l)h_2(Y_k,Y_r) \qquad (1)\end{aligned}$$

- Székely et al.: $w(t,s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.
  - $h_1(x,x') = |x-x'|_p^\alpha$, $h_2(y,y') = |y-y'|_q^\alpha$
  - Requires $E|X|_p^\alpha + E|Y|_q^\alpha + E|X|_p^\alpha |Y|_q^\alpha < \infty$
- Take $w(s,t) = f_S(s)f_T(t)$ to be the product of two probability densities symmetric about the origin
  - $h_1(x,x') = \phi_S(x-x')$, $h_2(y,y') = \phi_T(y-y')$.
  - No moment constraints

Distance Covariance: A measure of dependence

Choice of Weight Function $w(s, t)$

- For special choices of $w(s, t)$, the empirical distance covariance can be expressed as

$$\begin{aligned}\frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)h_2(Y_k, Y_l) \quad &+ \quad \frac{1}{n^2}\sum_{k,l} h_1(X_k, X_l)\frac{1}{n^2}\sum_{r,s} h_2(Y_r, Y_s) \\ &- \quad \frac{2}{n^3}\sum_{k,l,r} h_1(X_k, X_l)h_2(Y_k, Y_r) \qquad (1)\end{aligned}$$

- Székely et al.: $w(t, s) = (c|t|_p^{\alpha+p}|s|_q^{\alpha+q})^{-1}$, 2007, for $0 < \alpha < 2$.
  - $h_1(x, x') = |x - x'|_p^\alpha$, $h_2(y, y') = |y - y'|_q^\alpha$
  - Requires $E|X|_p^\alpha + E|Y|_q^\alpha + E|X|_p^\alpha|Y|_q^\alpha < \infty$
- Take $w(s, t) = f_S(s)f_T(t)$ to be the product of two probability densities symmetric about the origin
  - $h_1(x, x') = \phi_S(x - x')$, $h_2(y, y') = \phi_T(y - y')$.
  - No moment constraints
  - E.g., symmetric stable ($h_1(x_1, x_2) = \exp\{-c|x_1 - x_2|^\alpha\}$)

Distance Covariance: A measure of dependence

Consistency

- If $\{(X_t, Y_t)\}$ is ergodic, then

$$\hat{\mathcal{V}}_n^2(X, Y) \rightarrow_{a.s.} \mathcal{V}^2(X, Y).$$

Distance Covariance: A measure of dependence

Consistency

- If $\{(X_t, Y_t)\}$ is ergodic, then

$$\hat{\mathcal{V}}_n^2(X, Y) \to_{a.s.} \mathcal{V}^2(X, Y).$$

Limiting theory

- If $\{(X_t, Y_t)\}$ is α-mixing, and $\{X_t\} \perp \{Y_t\}$, then

$$n\hat{\mathcal{V}}_n^2(X, Y) \xrightarrow{d} \int |Q_{X,Y}(s,t)|^2 w(s,t) ds dt,$$

  where Q is a zero mean Gaussian process.
  Could take $Y_t = X_{t+h}$ to get auto-distance covariance for $\{X_t\}$.

Distance Covariance: A measure of dependence

Consistency

- If $\{(X_t, Y_t)\}$ is ergodic, then

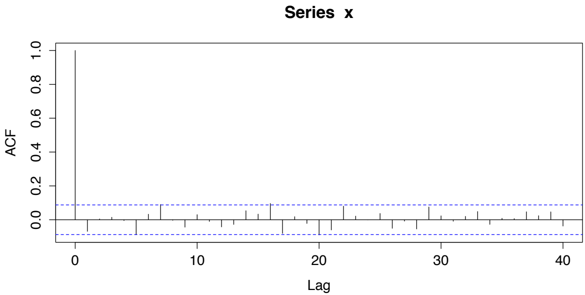
$$\hat{\mathcal{V}}_n^2(X, Y) \rightarrow_{a.s.} \mathcal{V}^2(X, Y).$$

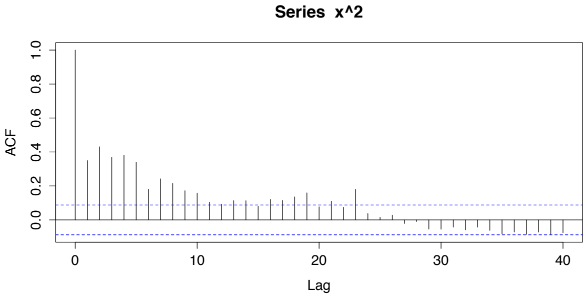
Limiting theory

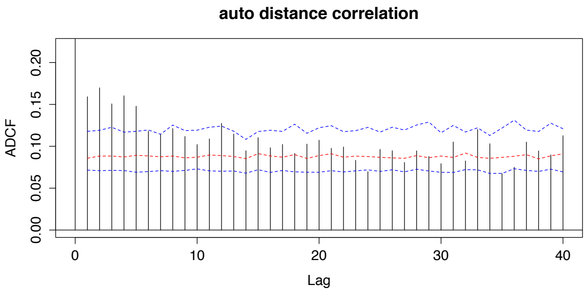
- If $\{(X_t, Y_t)\}$ is α-mixing, and $\{X_t\} \perp \{Y_t\}$, then

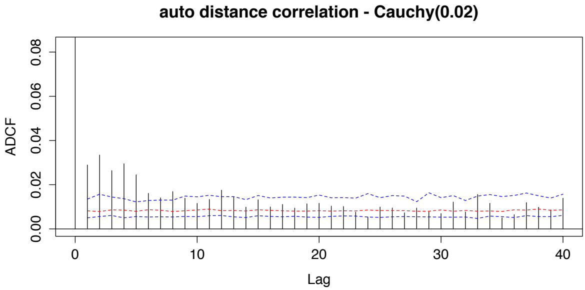
$$n\hat{\mathcal{V}}_n^2(X, Y) \stackrel{d}{\rightarrow} \int |Q_{X,Y}(s,t)|^2 w(s,t) ds dt,$$

  where Q is a zero mean Gaussian process.
  Could take $Y_t = X_{t+h}$ to get auto-distance covariance for $\{X_t\}$.
- For independence testing, we can use stationary bootstrap to approximate the distribution of the limit.


Series x
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40


Series x^2
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40


auto distance correlation
ADCF
0.00
0.05
0.10
0.15
0.20
0
10
20
30
40
Lag


auto distance correlation - Cauchy(0.02)
ADCF
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
Lag

Threshold Selection for Multivariate Heavy-Tailed Data

Multivariate Regular Variation

- Set-up: X_1,X_2,... iid $RV(\alpha)$, $R_i = ||X_i||$, $\Theta = X_i/||X_i||$.
- Recall: X_i is RV if and only if

$$P(R_i > x) = L(x)x^{-\alpha},$$

with $L(x)$ slowly varying, and:

$$P(\Theta \in \cdot | R > r) \to P_\eta(\Theta \in \cdot),\ r \to \infty.$$

Threshold Selection for Multivariate Heavy-Tailed Data

Multivariate Regular Variation

- Set-up: X_1,X_2,... iid $RV(\alpha)$, $R_i = ||X_i||$, $\Theta = X_i/||X_i||$.
- Recall: X_i is RV if and only if

$$P(R_i > x) = L(x)x^{-\alpha},$$

with $L(x)$ slowly varying, and:

$$P(\Theta \in \cdot | R > r) \to P_\eta(\Theta \in \cdot),\ r \to \infty.$$

Likelihood Estimation

- Example: Suppose $P_\eta(\Theta \in \cdot)$ has pdf $f(\theta|\eta)$. Estimate MLE from

$$L(\eta) = \sum_{i=1}^{n} \log f(\theta_i|\eta) \mathbf{1}_{[R_i > r]}.$$

Threshold Selection for Multivariate Heavy-Tailed Data

Multivariate Regular Variation

- Set-up: X_1,X_2,... iid $RV(\alpha)$, $R_i = ||X_i||$, $\Theta = X_i/||X_i||$.
- Recall: X_i is RV if and only if

$$P(R_i > x) = L(x)x^{-\alpha},$$

with $L(x)$ slowly varying, and:

$$P(\Theta \in \cdot|R > r) \to P_\eta(\Theta \in \cdot),\ r \to \infty.$$

Likelihood Estimation

- Example: Suppose $P_\eta(\Theta \in \cdot)$ has pdf $f(\theta|\eta)$. Estimate MLE from

$$L(\eta) = \sum_{i=1}^{n} \log f(\theta_i|\eta)\mathbf{1}_{[R_i > r]}.$$

- Choice of Threshold: r (large?) s.t. $(\Theta_i\mathbf{1}_{[R_i>r]}, R_i\mathbf{1}_{[R_i>r]})$ are approximately independent.

Threshold Selection for Multivariate Heavy-Tailed Data

Strategy

- Use distance correlation to approximate independence of Θ_i and R_i when R_i is large.

Threshold Selection for Multivariate Heavy-Tailed Data

Strategy

- Use distance correlation to approximate independence of Θ_i and R_i when R_i is large.
- For each r, test the independence of $(\Theta \mathbf{1}_{[R>r]}, R\mathbf{1}_{[R>r]})$ from the data until the result becomes significant.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Here R is independent of Θ iff $R > r_{0.9}$.

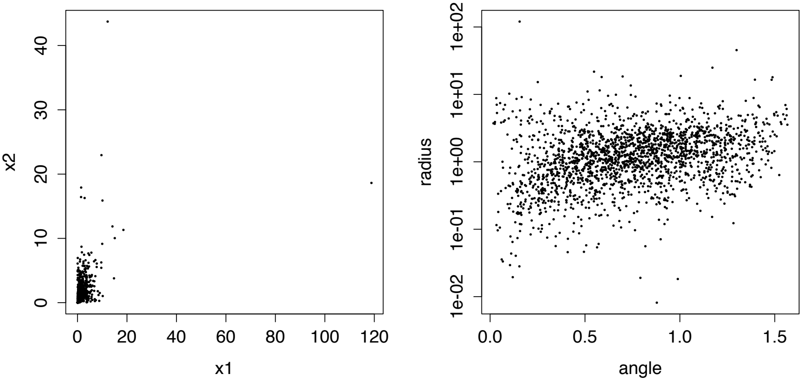


Figure: Simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Here R is independent of Θ iff $R > r_{0.9}$.

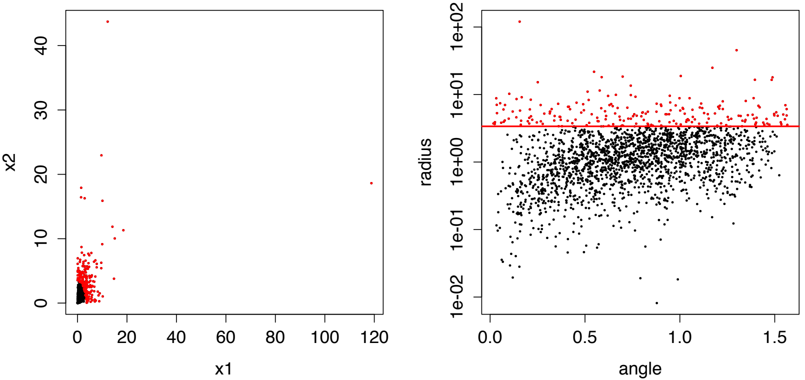


Figure: Simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

P-value path: Given each truncation level, find the p-value of independence test between $\log R_i$'s and Θ_i's after truncation.

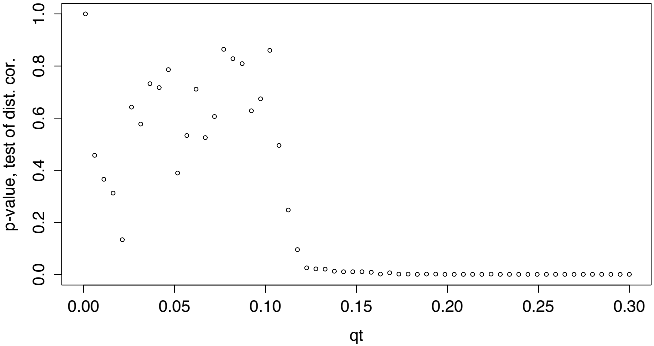


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

P-value path: Given each truncation level, find the p-value of independence test between $\log R_i$'s and Θ_i's after truncation.

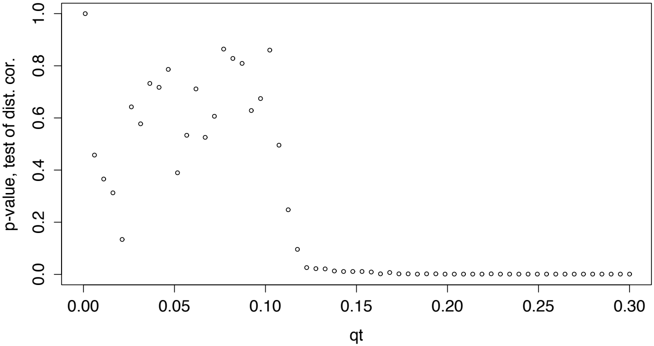


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Take an independent subsample of size 600 and plot its p-value path.

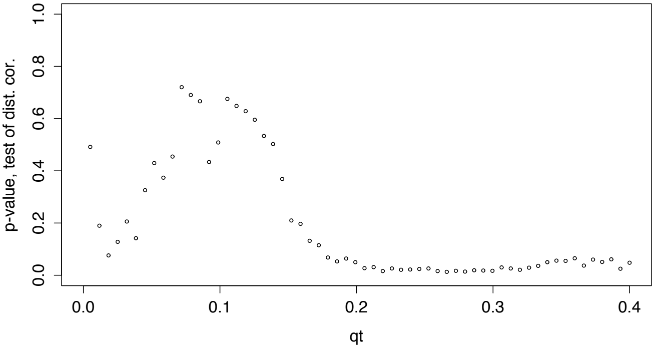


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.) Take another independent subsample and plot its p-value path.

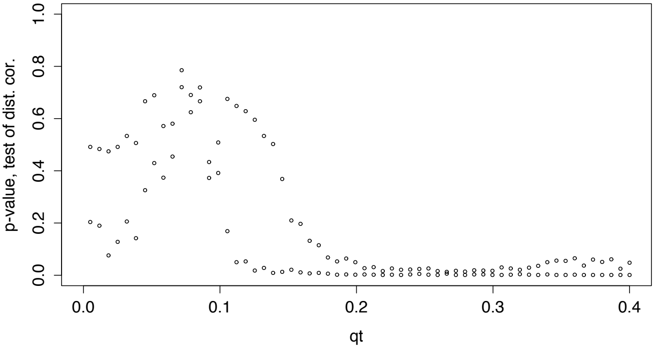


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Take another independent subsample and plot its p-value path.

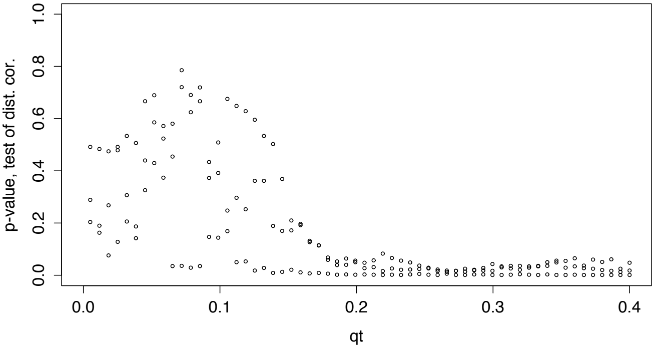


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Take 60 independent subsamples and plot their p-value paths.

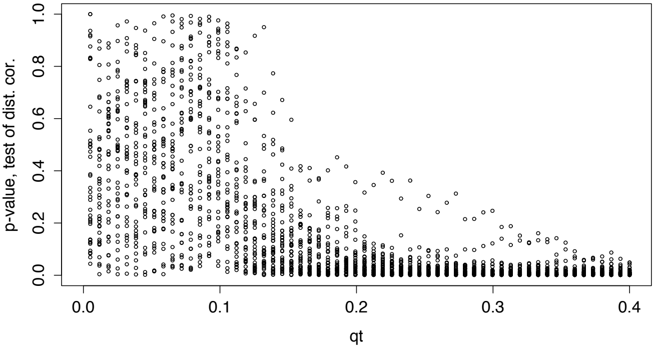


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1 for 60 independent subsamples.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Plot the mean p-value path.

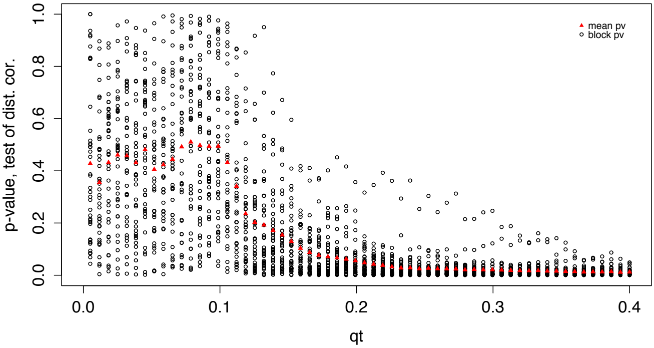


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1 for 60 independent subsamples.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

If independent, the mean p-value should to center around 0.5.

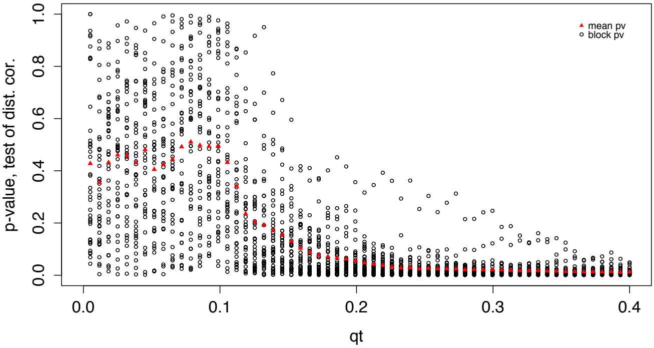


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1 for 60 independent subsamples.

Threshold Selection for Multivariate Heavy-Tailed Data

A simulated example (cont.)

Fit a piece-wise linear spline. See also Sen & Sen (2014).

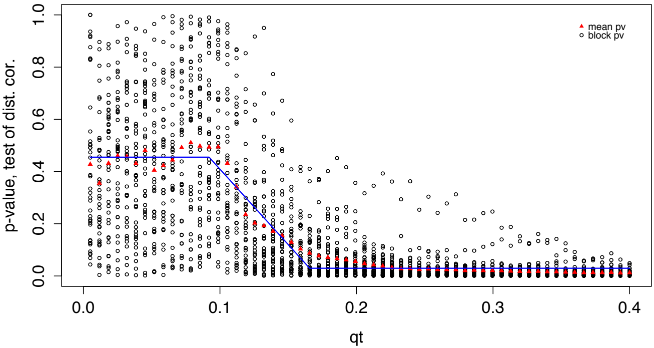


Figure: P-value of test of independence of R and Θ vs. truncation level for simulated data at independence level 0.1 for 60 independent subsamples.

To do:

- Optimal choice of weight function $w(s, t)$.
- Automatic threshold selection for multivariate heavy-tailed data (perhaps an R package?).
- Other ways to select thresholds: e.g., based on Kolmogorov-Smirnov test (Clauset et al., 2009).

2. ICA (joint with Jingjing, Nolan, Resnick, Alparslan)

The model

$$\underset{d\times 1}{X} = \underset{d\times d}{A} \cdot \underset{d\times 1}{S}$$

- Response $X = (x_1, \cdots, x_d)^T$
- Independent components $S = (S_1, \cdots, S_d)^T$
- Full rank $d \times d$ transformation matrix A
- The goal is to recover the unmixing matrix $W = A^{-1}$ and $S = W \cdot X$

A Strategy: Maximum Likelihood Estimator with Log-Concave Densities

- Find
$$f = |\det \hat{W}| \prod_{j=1}^{d} f_j(\hat{w}_j^T x)$$
that optimizes the log-likelihood
$$\int_{\mathbb{R}^d} \log(f) dP_n$$
- Use nonparametric density estimates
- Estimate f in space of d-dimensional log-concave densities
- Log-concave: exponential of piece-wise linear densities, normal, Laplace
- Not log-concave: t, stable, Pareto

Estimation Procedure

- Start from an arbitrary initial value of $\hat{W}$
- Step 1: Find log-concave density estimation $\hat{f}_j$ of $\hat{w}_j^T X$
- Step 2: With $\hat{f}_j$, update $\hat{W}$ to maximize the log-likelihood

$$\log|\det \hat{W}| + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{d}\log \hat{f}_j(\hat{w}_j^T x_i)$$

- Iterate steps 1 and 2, until convergence of the log-likelihood

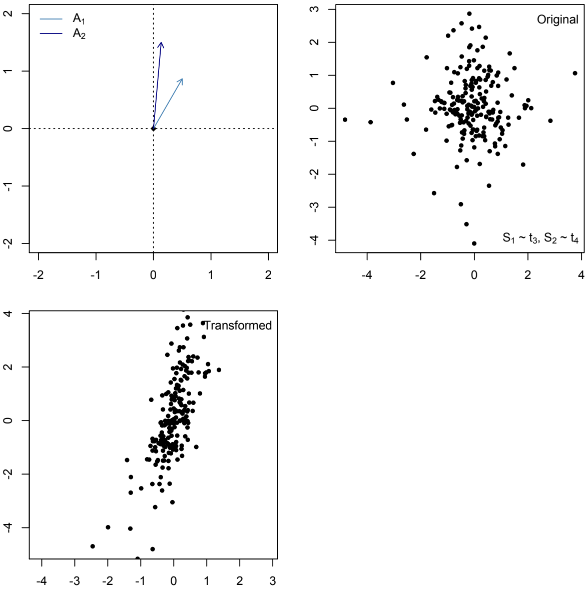
Pre-Whitening

- Assume each component of S has finite variance (can relax, c.f. Chen and Bickel (2005))
- Let $\Sigma_X = \mathrm{cov}(X)$ and $Z = \Sigma_X^{-1/2} X$
- $S = O \cdot Z$, where $O = W \cdot \Sigma_X^{-1/2}$ is an orthogonal matrix
- Number of unknown parameters reduced from d^2 to $d(d-1)/2$
- In practice, estimate Σ_X with sample covariance matrix $\hat{\Sigma}_X$

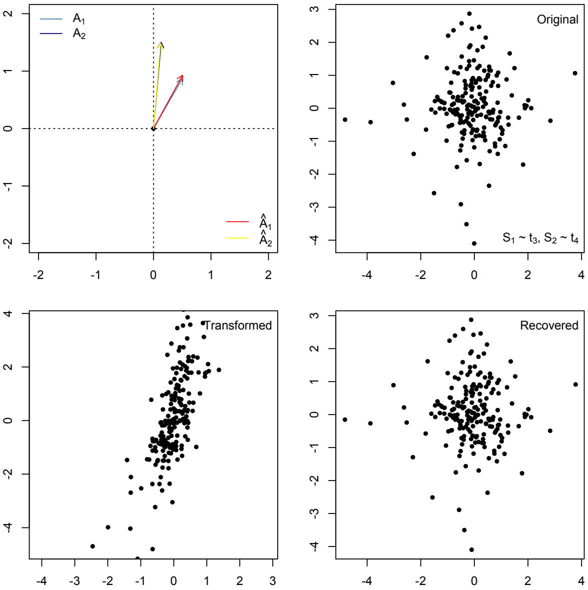
Consistency of the Maximum Likelihood Estimator

- If $\int \|x\| dP < \infty$, $\hat{W} \xrightarrow{\text{a.s.}} W$ (Samworth and Yuan (2012))
- If $\int \|x\| dP = \infty$, $\int_{\mathbb{R}^d} \log(f)\, dP = -\infty$ for all $f \in \mathcal{F}_d$
- If $\Gamma_S^{-1} \hat{\Sigma}_S^{1/2} \xrightarrow{P} I$, consistency holds for pre-whitened estimator (Chen and Bickel (2005))
  - Γ_S: sample variance matrix of S
  - $\hat{\Sigma}_S$: sample covariance matrix of S
  - Condition holds for all non-degenerate independent distributions

$S_1 \sim t_3,\ S_2 \sim t_4$

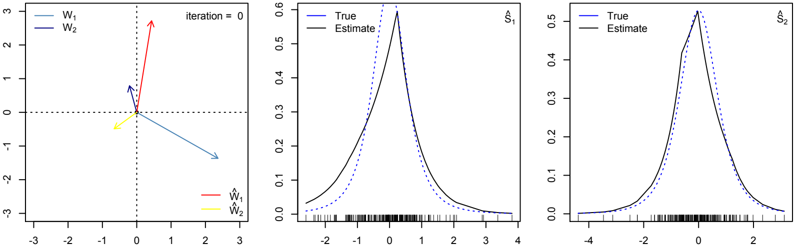


$S_1 \sim t_3$, $S_2 \sim t_4$



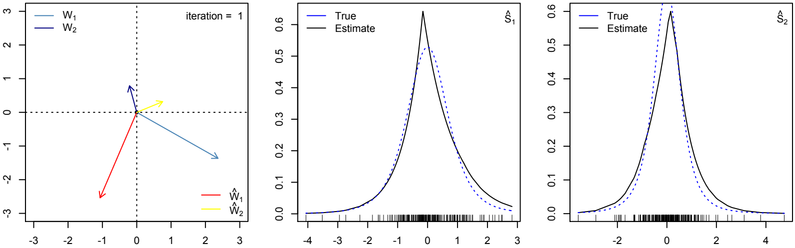
Convergence of Estimation Algorithm

$S_1 \sim t_3,\ S_2 \sim t_4$



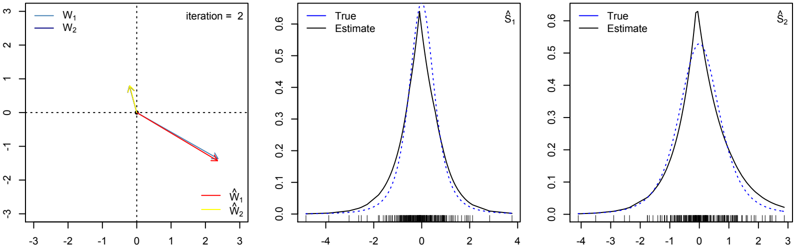
Convergence of Estimation Algorithm

$S_1 \sim t_3, \ S_2 \sim t_4$

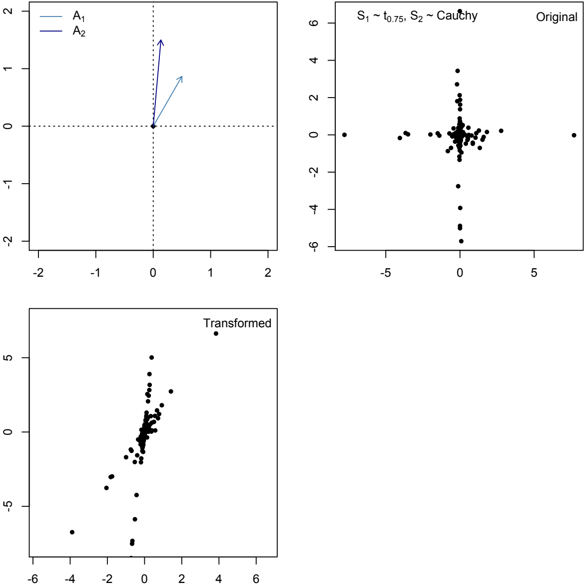


Convergence of Estimation Algorithm

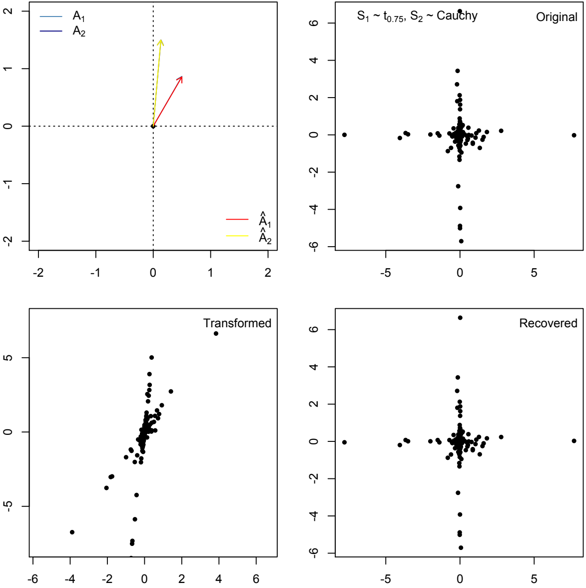
$S_1 \sim t_3, \; S_2 \sim t_4$



$S_1 \sim t_{0.75}$, $S_2 \sim$ Cauchy

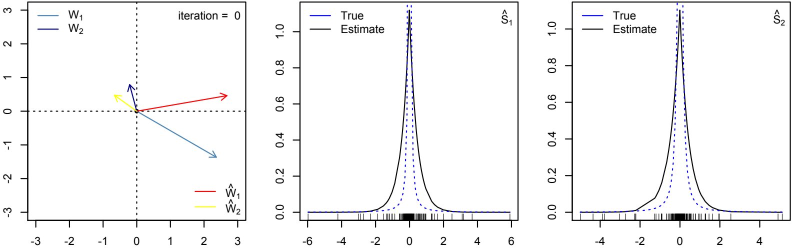


$S_1 \sim t_{0.75}$, $S_2 \sim$ Cauchy



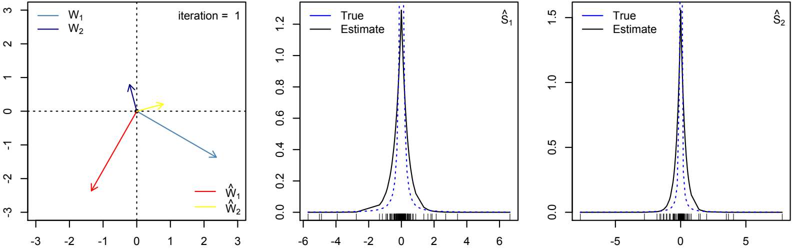
Convergence of Estimation Algorithm

$S_1 \sim t_{0.75}$, $S_2 \sim$ Cauchy



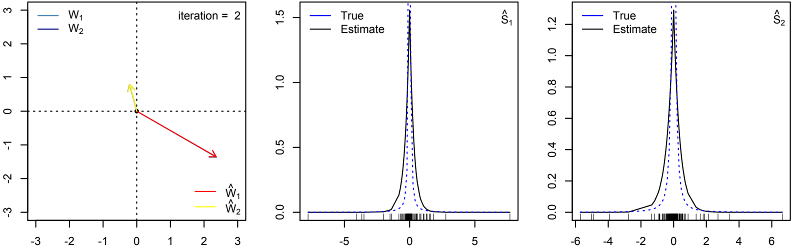
Convergence of Estimation Algorithm

$S_1 \sim t_{0.75}$, $S_2 \sim$ Cauchy



Convergence of Estimation Algorithm

$S_1 \sim t_{0.75}$, $S_2 \sim$ Cauchy



M-estimator Based on Nonparametric Mutual Information

- Aim to relax the moment condition $\int \|x\| dP < \infty$ for heavy-tailed distributions
- For fixed $a > 0$, find $\hat{W}$ that minimizes

$$\int_{\mathbb{R}^d} \hat{f} dP_n^{\hat{W},a} - \int_{\mathbb{R}^d} \tilde{f} dP_n^{\hat{W},a}$$

- $\hat{f}$: log-concave estimation of $\hat{W}X \cdot \mathbb{1}_{\{\hat{W}X \in [-a,a]^d\}}$
- $\tilde{f}$: log-concave estimation of the marginals of $\hat{W}X \cdot \mathbb{1}_{\{\hat{W}X \in [-a,a]^d\}}$

Overcomplete ICA

$$\underset{d\times 1}{X} = \underset{d\times m}{A} \cdot \underset{m\times 1}{S}$$

- $d < m$
- Challenge: A not invertible
- For fixed A, infinitely many S such that $X = AS$ because of $m - d$ missing components
- Strategy: for fixed A and f, estimate missing S with maximum likelihood estimators

Estimation Procedure Overcomplete ICA

- Start from an arbitrary initial value of $\hat{A}$ and S such that $X = \hat{A}S$
- Step 1: Find log-concave estimation $\hat{f}_j$ of the distribution of S_j
- Step 2: With $\hat{f}_j$, update $\hat{A}$ to maximize the log-likelihood
  - With an extra optimization step to estimate missing S
- Iterate steps 1 and 2, until convergence of the log-likelihood

Plans for next year

- Complete project on testing independence in a time series setting.
- Complete project on selection of threshold for multivariate RV data and its offshoots. (Partly joint with Cornell group.)
- Expand development of statistical tools for multivariate heavy tails: parameter estimation of parameters in the spectral distribution with emphasis on network models; changepoints in heavy-tails, etc. (Joint with Cornell team.)
- Complete theory and develop algorithms for the ICA model in the undercomplete case.
- Develop methodology to handle the overcomplete case in ICA (with Nolan)
- Expand the random matrix work (with p and n going to infinity) to the case of nonlinear models. Rank transforms of the rows may be of particular interest.