Directional histograms
Measuring independence for stable distributions

John Nolan

American University
Washington, DC, USA

MURI Workshop, NYC
6 March 2015

1. Directional histograms

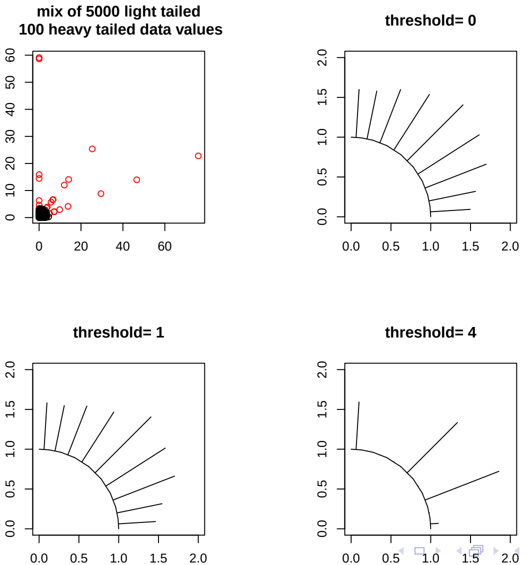
2. Independence measure η_p for bivariate stable r. vectors

3. Sample measure $\widehat{\eta}_p$

Outline

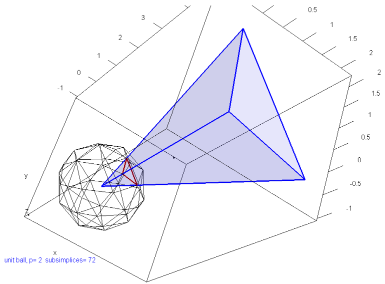
1. Directional histograms
2. Independence measure η_p for bivariate stable r. vectors
3. Sample measure $\widehat{\eta}_p$

Directional histogram $d = 2$ - count how many in each "direction"

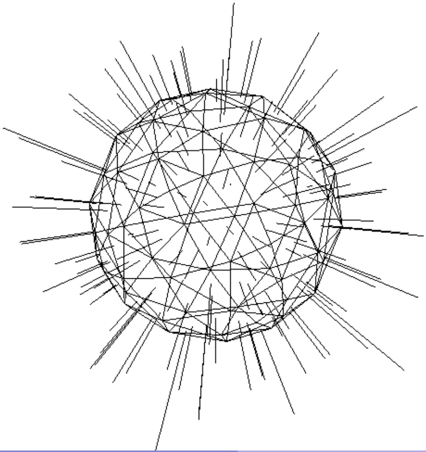


Generalize to $d \geq 3$?

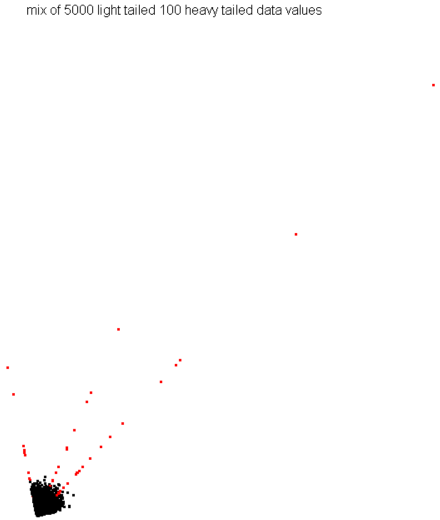
- triangulate sphere
- each simplex on sphere determines a cone
- loop through data points, seeing which cone each falls in
- If $d = 3$, plot
- Variations:
  - threshold based on distance from center
  - use ℓ_p ball
  - restrict to positive orthant



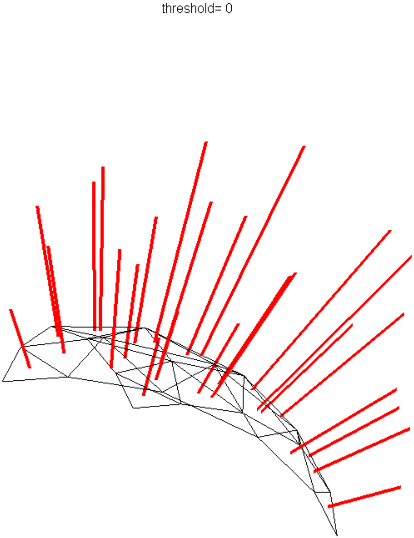
Directional histogram $d = 3$



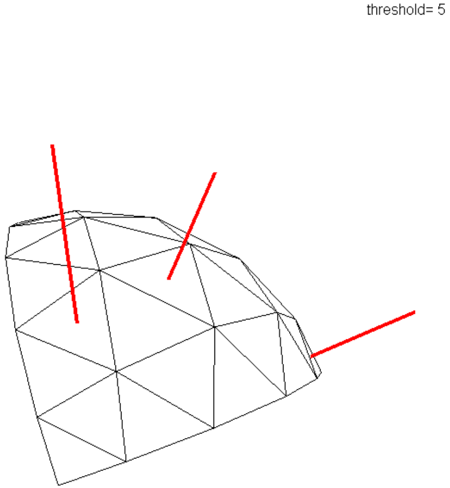
Directional dependence (simulated data)



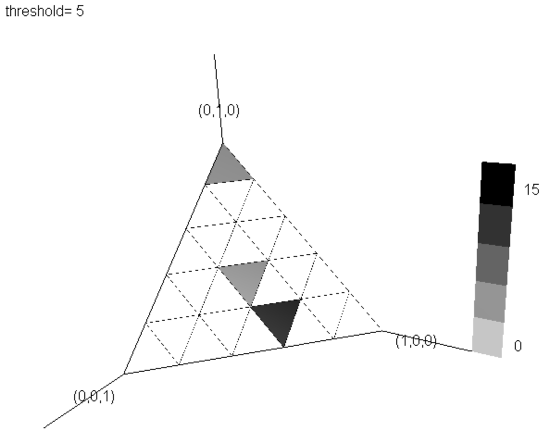
All data



Thresholding by distance from origin



Thresholding by distance from origin (alternate view)



Directional histogram $d > 3$

Subdivision routines return a list of simplices in some order. For any d, can compute the directional histogram counts.

Then plot the a standard histogram using index of simplex.

Directional histogram $d > 3$

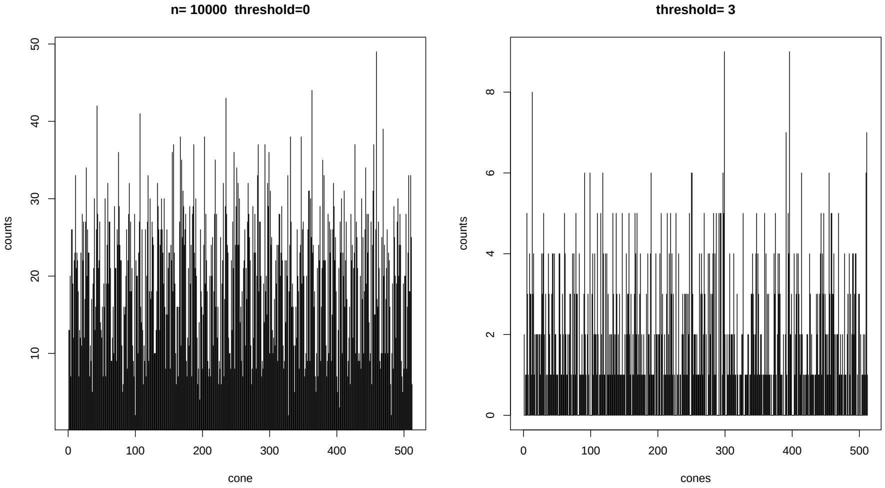
Subdivision routines return a list of simplices in some order. For any d, can compute the directional histogram counts.

Then plot the a standard histogram using index of simplex.

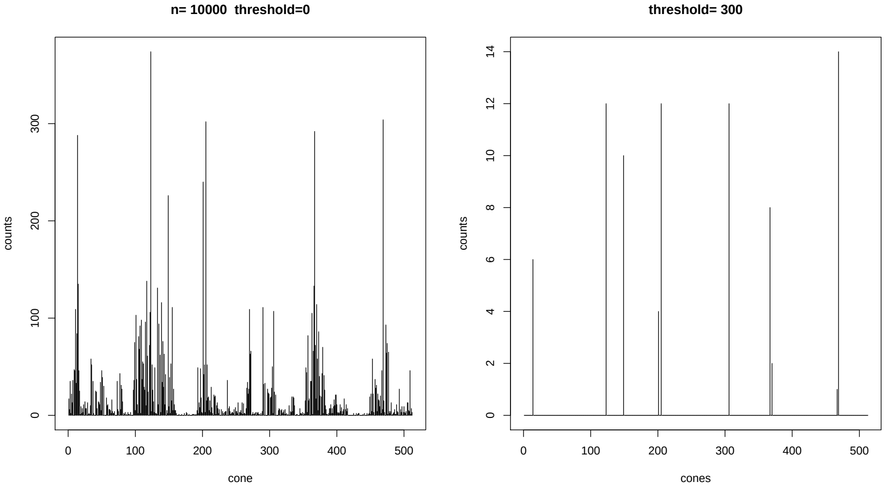
Lose geometry, but can show concentration in different directions. Thresholding may reveal a few directions where extremes lie.

Can use to select model to use on a given data set, e.g. isotropic when histogram is roughly uniform, discrete angular measure when just a few directions present after thresholding.

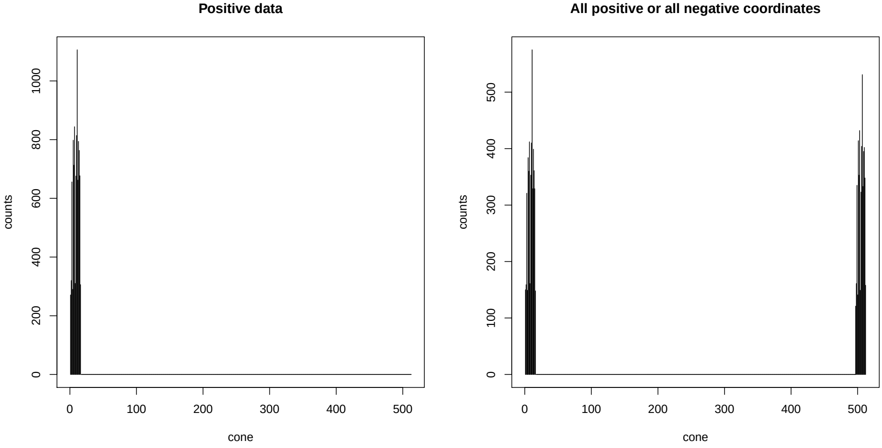
$d = 5$, with 512 cones/directions - isotropic



$d = 5$, with 512 cones/directions - $m = 7$ point masses



$d = 5$, with 512 cones/directions - concentration in sectors



Outline

Spectral measure characterization

We will say $\mathbf{X} \sim \mathbf{S}(\alpha, \Lambda, \delta; j)$, $j = 0, 1$ if its joint characteristic function is given by

$$\phi(\mathbf{u}) = E \exp(i\langle \mathbf{u}, \mathbf{X}\rangle) = \exp\left(-\int_{\mathbb{S}} \omega\left(\langle \mathbf{u}, \mathbf{s}\rangle | \alpha; j\right) \Lambda(d\mathbf{s}) + i\langle \mathbf{u}, \delta\rangle\right),$$

where

$$\omega(t|\alpha; j) = \begin{cases} |t|^{\alpha}[1 + i\,\mathrm{sign}\,(t) \tan \frac{\pi\alpha}{2}(|t|^{1-\alpha} - 1)] & \alpha \neq 1, j = 0 \\ |t|^{\alpha}[1 - i\,\mathrm{sign}\,(t) \tan \frac{\pi\alpha}{2}] & \alpha \neq 1, j = 1 \\ |t|[1 + i\,\mathrm{sign}\,(t) \frac{2}{\pi} \log |t|] & \alpha = 1, j = 0, 1. \end{cases}$$

The 1-parameterization is more commonly used, but discontinuous in α. 0-parameterization is a continuous parameterization.

Projection parameterization

Every one dimensional projection $\langle \mathbf{u}, \mathbf{X} \rangle = u_1 X_1 + u_2 X_2 + \cdots + u_d X_d$ has a univariate stable distribution, with a constant index of stability α and skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$ that depend on the direction $\mathbf{u}$.

Projection parameterization

Every one dimensional projection $\langle \mathbf{u}, \mathbf{X} \rangle = u_1 X_1 + u_2 X_2 + \cdots + u_d X_d$ has a univariate stable distribution, with a constant index of stability α and skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$ that depend on the direction $\mathbf{u}$.

We will call the functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ the projection parameter functions. They determine the joint distribution via the Cramér-Wold device, so we can parameterize $\mathbf{X}$ by these projection parameter functions: $\mathbf{X} \sim \mathbf{S}\left(\alpha, \beta(\cdot), \gamma(\cdot), \delta(\cdot); j\right)$, $j = 0$ or $j = 1$.

Projection parameterization

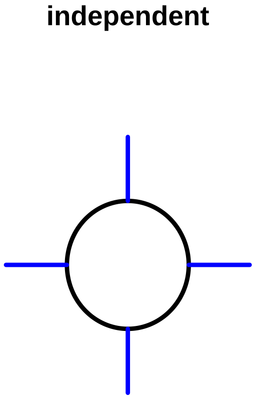
Every one dimensional projection $\langle \mathbf{u}, \mathbf{X} \rangle = u_1 X_1 + u_2 X_2 + \cdots + u_d X_d$ has a univariate stable distribution, with a constant index of stability α and skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$ that depend on the direction $\mathbf{u}$.

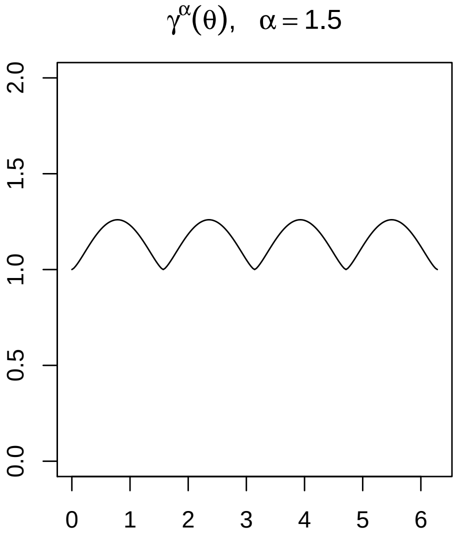
We will call the functions $\beta(\cdot)$, $\gamma(\cdot)$ and $\delta(\cdot)$ the projection parameter functions. They determine the joint distribution via the Cramér-Wold device, so we can parameterize $\mathbf{X}$ by these projection parameter functions: $\mathbf{X} \sim \mathbf{S}(\alpha, \beta(\cdot), \gamma(\cdot), \delta(\cdot); j)$, $j = 0$ or $j = 1$.

In what follows, we will always assume that $\mathbf{X}$ has normalized components: $\gamma(1,0) = \gamma(0,1) = 1$.

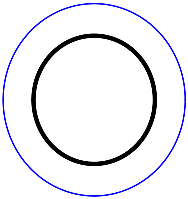
Will sometimes use polar notation: $\gamma(\theta) := \gamma(\cos\theta, \sin\theta)$.

$\Lambda(\cdot)$ and $\gamma(\cdot)$

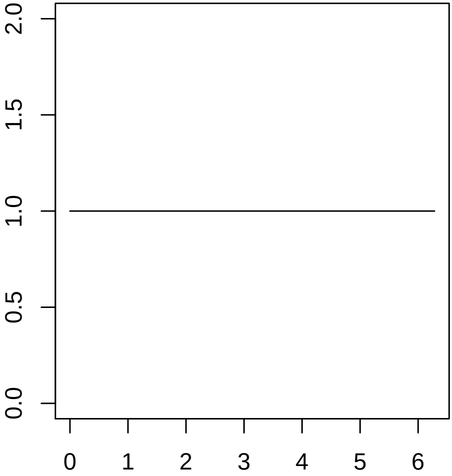


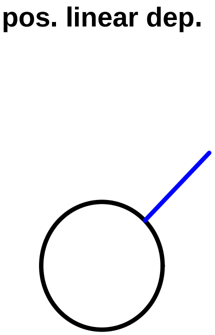


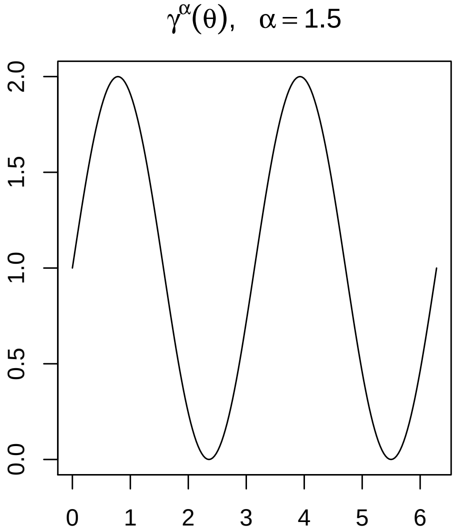
isotropic



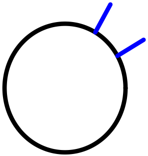
$\gamma^{\alpha}(\theta), \quad \alpha = 1.5$



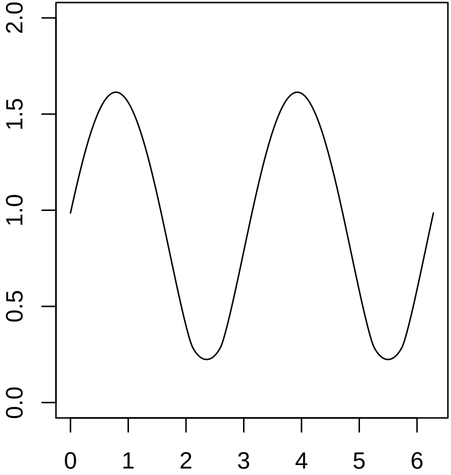

pos. linear dep.


$\gamma^{\alpha}(\theta), \quad \alpha = 1.5$
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6

pos. associated



$\gamma^{\alpha}(\theta), \quad \alpha = 1.5$



Set $\gamma_\perp(\mathbf{u}) = (|u_1|^\alpha + |u_2|^\alpha)^{1/\alpha}$ (independence), $p \in [1, \infty]$

$$\eta_p = \eta_p(X_1, X_2) = \|\gamma^\alpha(u_1, u_2) - \gamma_\perp^\alpha(u_1, u_2)\|_{L^p(\mathbb{S}, d\mathbf{u})}. \tag{1}$$

Here $d\mathbf{u}$ is (unnormalized) surface area on $\mathbb{S}$.

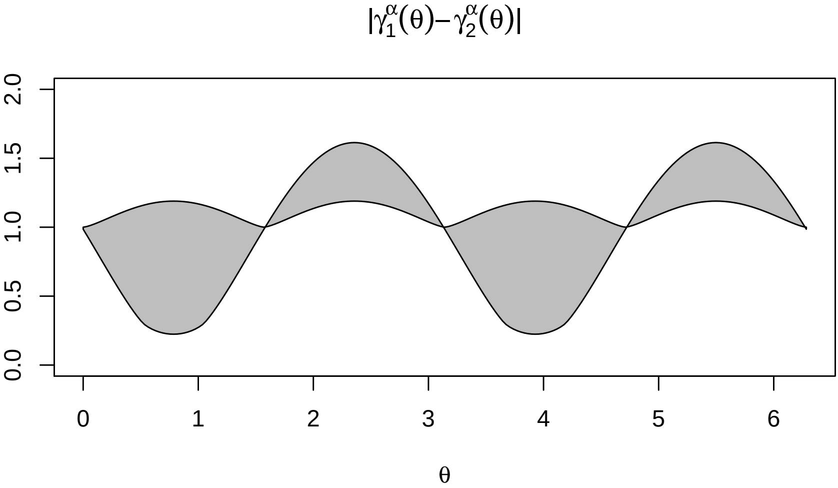
Set $\gamma_\perp(\mathbf{u}) = (|u_1|^\alpha + |u_2|^\alpha)^{1/\alpha}$ (independence), $p \in [1, \infty]$

$$\eta_p = \eta_p(X_1, X_2) = \|\gamma^\alpha(u_1, u_2) - \gamma_\perp^\alpha(u_1, u_2)\|_{L^p(\mathbb{S}, d\mathbf{u})}. \qquad (1)$$

Here $d\mathbf{u}$ is (unnormalized) surface area on $\mathbb{S}$.

$\mathbf{X}$ has independent components if and only if $\eta_p = 0$ for some (every) $p \in [1, \infty]$.

η_p measures how far the scale function of $\mathbf{X}$ is from the scale function of a stable r. vector with independent components: when $\mathbf{X}$ is symmetric, earlier work shows $\sup_{\mathbf{x} \in \mathbb{R}^2} |f(\mathbf{x}) - f_\perp(\mathbf{x})| \le k_\alpha \|\gamma(\cdot) - \gamma_\perp(\cdot)\|$.



$|\gamma_1^\alpha(\theta) - \gamma_2^\alpha(\theta)|$
0.0
0.5
1.0
1.5
2.0
0
1
2
3
4
5
6
θ

Properties of η_p

- The p-norm in (1) is evaluated as an integral over the unit circle $\mathbb{S}$, not all of $\mathbb{R}^2$. In polar coordinates,

$$\eta_p = \left(2 \int_0^{\pi} |\gamma^{\alpha}(\cos\theta, \sin\theta) - \gamma_{\perp}^{\alpha}(\cos\theta, \sin\theta)|^p \, d\theta \right)^{1/p}, \qquad (2)$$

  where the interval of integration has been reduced by using the fact that $\gamma(\cdot)$ is π-periodic
- α can be any value in $(0, 2)$ and $\mathbf{X}$ can have symmetric or non-symmetric components, and it can be centered or shifted.
- η_p is symmetric: $\eta_p(X_1, X_2) = \eta_p(X_2, X_1)$.

- $\eta_p \geq 0$ by definition, not measuring positive/negative dependence, just distance from independence. Don't think there is a general way of assigning a sign, e.g. rotate the indep. components case by $\pi/4$ and the resulting distribution bunches around both the lines $y = x$ and $y = -x$ for large values of $|\mathbf{X}|$.
- The definition makes sense in the Gaussian case: when $\alpha = 2$, the scale function for a bivariate Gaussian distribution with correlation ρ is $\gamma(\mathbf{u})^2 = 1 + 2\rho u_1 u_2$ and $\gamma_\perp = 1$. Then $\eta_p^p = |2\rho|^p \int_{\mathbb{S}} |u_1 u_2|^p d\mathbf{u}$, so $\eta_p = k_p|\rho|$. In elliptically contoured/sub-Gaussian case, can get an integral expression that can be evaluated numerically.
- Multivariate stable $\mathbf{X} = (X_1, \ldots, X_d)$ has mutually independent components if and only if all pairs are independent, so the components of $\mathbf{X}$ are mutually independent if and only if $\eta_p(X_i, X_j) = 0$ for all $i > j$.

Covariation and co-difference in terms of $\gamma(\cdot)$

For $\alpha > 1$, the covariation is

$$[X_1, X_2]_\alpha = \int_{\mathbb{S}} s_1 s_2^{<\alpha-1>} \Lambda(d\mathbf{s}) = \frac{1}{\alpha} \left. \frac{\partial \gamma^\alpha(u_1, u_2)}{\partial u_1} \right|_{(u_1=0, u_2=1)} .$$

Thus the covariation depends only on the behavior of $\gamma(\cdot, \cdot)$ near the point $(1, 0)$. If X_1 and X_2 are independent, then $[X_1, X_2]_\alpha = 0$; but the converse is false.

Covariation and co-difference in terms of $\gamma(\cdot)$

For $\alpha > 1$, the covariation is

$$[X_1, X_2]_\alpha = \int_{\mathbb{S}} s_1 s_2^{<\alpha-1>} \Lambda(d\mathbf{s}) = \frac{1}{\alpha} \left. \frac{\partial \gamma^\alpha(u_1, u_2)}{\partial u_1} \right|_{(u_1=0, u_2=1)} .$$

Thus the covariation depends only on the behavior of $\gamma(\cdot,\cdot)$ near the point $(1,0)$. If X_1 and X_2 are independent, then $[X_1, X_2]_\alpha = 0$; but the converse is false.

The co-difference is defined for symmetric α-stable vectors, and can be written as

$$\tau = \tau(X_1, X_2) = \gamma^\alpha(1,0) + \gamma^\alpha(0,1) - \gamma^\alpha(1,-1),$$

and is defined for any $\alpha \in (0,2)$. If X_1 and X_2 are independent, then $\tau = 0$. When $\alpha < 1$ and $\tau = 0$, then indep. If $\alpha > 1$, need both $\tau(X_1, X_2) = 0$ and $\tau(X_2, X_1) = 0$ to guarantee indep.

Outline

Use max. likelihood estimation of the marginals and get $\widehat{\alpha}$, normalize each component. For angles $0 \leq \theta_1 < \theta_2 < \cdots < \theta_m \leq \pi$, define $\widehat{\gamma}_j = \widehat{\gamma}(\cos\theta_j, \sin\theta_j) =$ ML estimate of the scale of the projected data set $\langle \mathbf{Y}_i, (\cos\theta_j, \sin\theta_j)\rangle$, $i = 1, \ldots, n$

Use max. likelihood estimation of the marginals and get $\widehat{\alpha}$, normalize each component. For angles $0 \leq \theta_1 < \theta_2 < \cdots < \theta_m \leq \pi$, define $\widehat{\gamma}_j = \widehat{\gamma}(\cos\theta_j, \sin\theta_j) =$ ML estimate of the scale of the projected data set $\langle \mathbf{Y}_i, (\cos\theta_j, \sin\theta_j)\rangle$, $i = 1, \ldots, n$

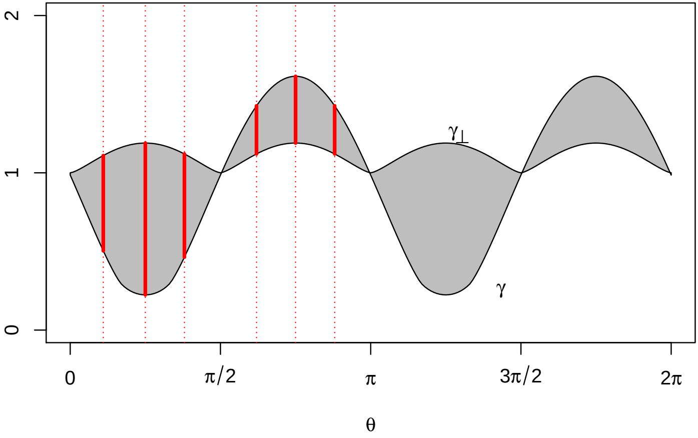
Define

$$\widehat{\eta}_2 = \left(\sum_{j=1}^{m} \left(\widehat{\gamma}_j^{\widehat{\alpha}} - \gamma_{\perp,j}^{\widehat{\alpha}} \right)^2 \right)^{1/2} .$$

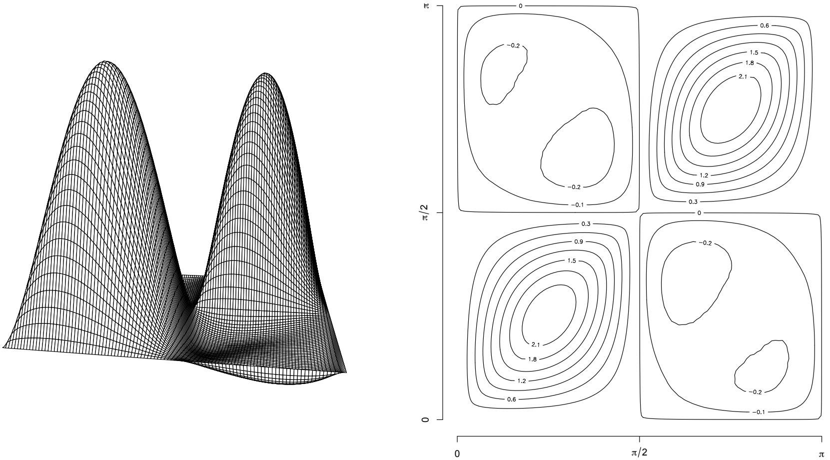
Get critical values by simulation, depends on α and grid.

Suggest uniform grid with m points in first and second quadrant that avoid 0, $\pi/2$, π

Uniform grid with $m = 3$ in each quadrant



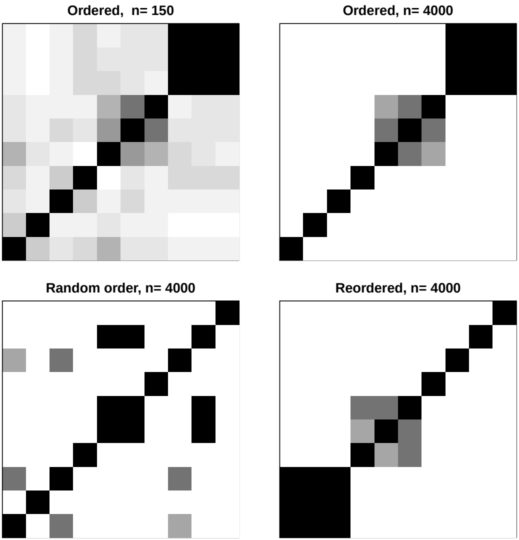
Covariance of $\widehat{\gamma}(\theta_1)$ and $\widehat{\gamma}(\theta_2)$



Power calculation via simulation, $\alpha = 1.5$, 5 grid points per quadrant, 1000 simulations

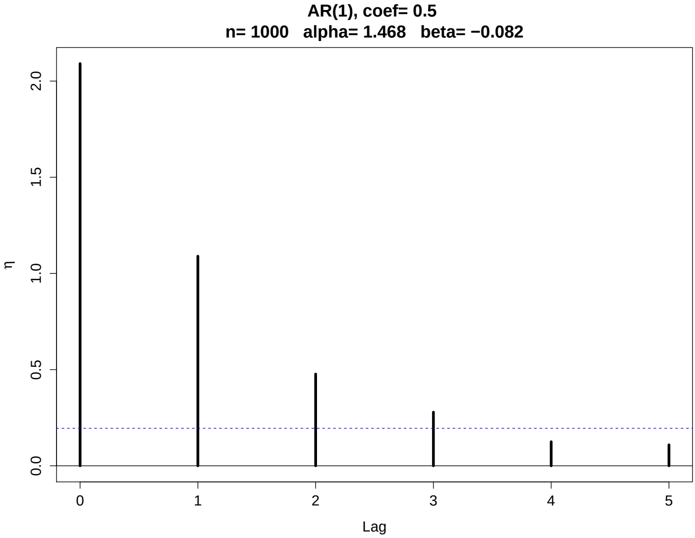
n	isotropic	indep. $\circlearrowleft \pi/4$	indep. $\circlearrowleft \pi/8$	indep. $\circlearrowleft \pi/16$	exact linear dep.
25	0.191	0.322	0.243	0.213	1
50	0.223	0.624	0.381	0.183	1
100	0.344	0.918	0.644	0.214	1
200	0.636	0.998	0.937	0.440	1
300	0.874	1	0.997	0.627	1
400	0.960	1	1	0.791	1
500	0.989	1	1	0.893	1
600	0.999	1	1	0.959	1
700	1	1	1	0.980	1
800	1	1	1	0.985	1
900	1	1	1	0.998	1
1000	1	1	1	0.997	1

Multivariate: compute $\widehat{\eta}_{i,j}$ between all pairs (X_i, X_j)

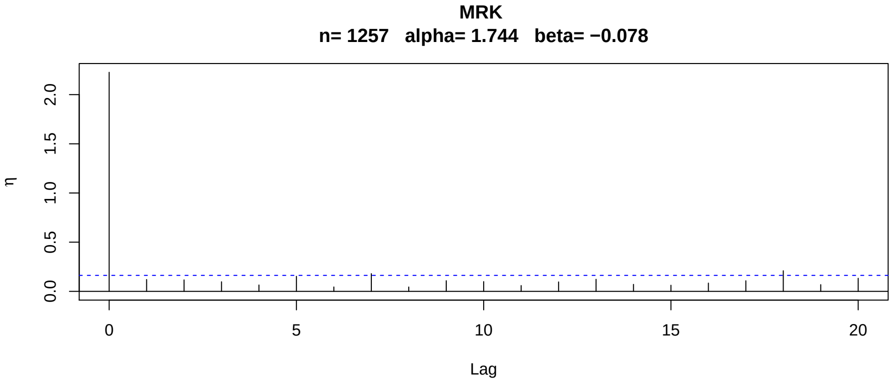


Time series - plot $\eta(X_t, X_{t+h})$

Simulated data with stable innovations:

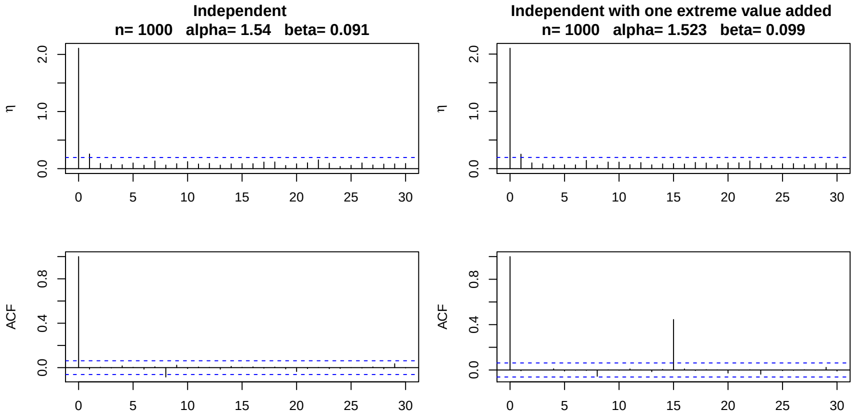


Time series - returns of Merck stock for 2010-2014



Robustness of acf vs η plot

Simulated time series with independent stable terms. In this simulation, the η and acf plots look similar (left). Changing one point by replacing a point 15 time periods away from max with 0.8*max shows η plot unchanged, but acf shows strong dependence (right).



η for **X** in the domain of attraction of stable

The calculation of η only requires an estimate of the tail index α and scale in directions $\theta_1, \ldots, \theta_m$. Can use any tail estimator of the univariate data sets obtained by projecting the data in different directions. The following examples used a simple tail estimator - regression on the tail probabilities.

η for **X** in the domain of attraction of stable

The calculation of η only requires an estimate of the tail index α and scale in directions $\theta_1, \ldots, \theta_m$. Can use any tail estimator of the univariate data sets obtained by projecting the data in different directions. The following examples used a simple tail estimator - regression on the tail probabilities.

Simulated using symmetrized Paretos: $X = Y_1 - Y_2$ where each term is indep. Pareto($\alpha = 1.5$).

- Fix n=sample size.
- Find critical value by simulation. Bootstrap indep. components (X_1, X_2), compute $\widehat{\eta}$ and tabulate. Repeat $M = 10000$ times and find a critical value c_p based on $(1-p)$ quantile of tabulated values.
- Simulate different data sets: isotropic $(\cos U, \sin U)X$ where $U \sim$Uniform$(0, 2\pi)$; rotations of independent case $R(\theta)(X_1, X_2)$ for $\theta = pi/4, pi/8, pi/16$; exact linear dependence $\epsilon(X, X)$ where $\epsilon = \pm 1$ w/ prob. 1/2.
- Vary n and tabulate power

Power calculations in DOA case

sample size n	isotropic	independent $\circlearrowleft \pi/4$	independent $\circlearrowleft \pi/8$	independent $\circlearrowleft \pi/16$	exact linear dependence
100	0.253	0.057	0.049	0.058	0.161
200	0.708	0.025	0.040	0.049	0.342
300	0.844	0.010	0.013	0.023	0.481
400	0.940	0.011	0.020	0.022	0.995
500	0.956	0.011	0.007	0.018	1
600	0.986	0.024	0.013	0.028	1
700	0.988	0.023	0.003	0.009	1
800	0.995	0.258	0.012	0.019	1
900	0.998	0.284	0.013	0.011	1
1000	0.993	0.498	0.006	0.009	1
2000	1	0.996	0.376	0.008	1
3000	1	1	0.876	0.003	1
4000	1	1	0.989	0.003	1
5000	1	1	1	0.004	1

Require larger sample to detect dependence; depends on choosing cutoff correctly and estimators of α and scale.