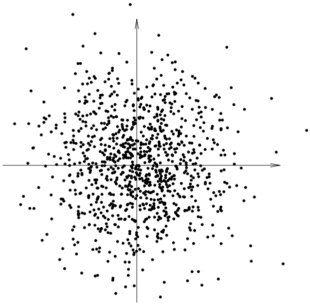
Multivariate Extremes and Geometry

Guus Balkema Paul Embrechts* Natalia Nolde

*Department of Mathematics
and
RiskLab
ETH Zurich, Switzerland

Introduction

- **Key tools**:
  - sample clouds
  - densities

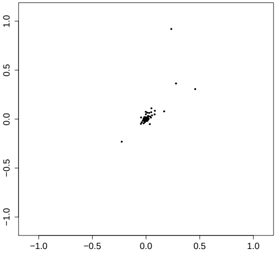


- Limiting shape of sample clouds is at the heart of many results of the thesis
  - Global description of multivariate extremes
  - The shape of the sample clouds and of the level sets of densities is crucial in determining various probabilistic properties of light-tailed distributions

Sample clouds: Notation

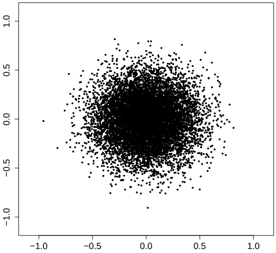
- $\mathbf{X}_1, \mathbf{X}_2, \ldots$ i.i.d. random vectors on $\mathbb{R}^d$
- $N_n = \{\mathbf{X}_1/a_n, \ldots, \mathbf{X}_n/a_n\}$: n-point sample cloud with scaling constants $a_n > 0$, $a_n \to \infty$ as $n \to \infty$
- $N_n(A) = \sum_{i=1}^n \mathbf{1}_A(\mathbf{X}_i/a_n)$ for any Borel set $A \subset \mathbb{R}^d$
  $\uparrow$
  number of points of N_n contained in set A

Examples of sample clouds ($n = 10,000$)

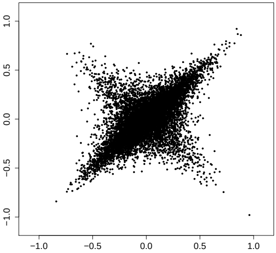


elliptic Cauchy with dispersion matrix

$$\Sigma = \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix}$$



standard normal



meta-Cauchy with normal marginals

Sample clouds: Convergence onto a set

Definition

The sample clouds N_n converge onto a compact set E in $\mathbb{R}^d$ if

- $\mathbb{P}\{N_n(U^c) > 0\} \to 0$ *for open sets U containing E, and*
- $\mathbb{P}\{N_n(\mathbf{p} + \epsilon B) > m\} \to 1$, $m \geq 1$, $\epsilon > 0$, $\mathbf{p} \in E$

  where B denotes the unit Euclidean ball

The set E is called a limit set

Part I

Meta densities and the shape of their sample clouds

References:

A.A. Balkema, P. Embrechts and N. Nolde (2010). Meta densities and the shape of their sample clouds. *J. Multivariate Analysis* 101: 1738-1754.

A.A. Balkema, P. Embrechts and N. Nolde (2010). Sensitivity of the asymptotic behaviour of meta distributions. *Submitted*.

Why meta distributions?

Recent popularity in applications of multivariate probability theory, especially in finance

- Asymptotic independence of coordinatewise maxima as a shortcoming of the multivariate Gaussian model
- Go beyond normality by introducing stronger tail dependence while preserving normal marginals

Why meta distributions?

Recent popularity in applications of multivariate probability theory, especially in finance

- Asymptotic independence of coordinatewise maxima as a shortcoming of the multivariate Gaussian model
- Go beyond normality by introducing stronger tail dependence while preserving normal marginals

- A typical example:
  - Start with a multivariate t distribution (tail dependence and heavy tails)
  - Transform each coordinate so that the new distribution has normal marginals (light tails)
  - The new distribution is called a meta distribution with normal marginals based on the original t distribution (tail dependence and light tails)

Definition (Meta distribution)

- Random vector $\mathbf{Z}$ in $\mathbb{R}^d$ with df F and continuous marginals F_i, $i = 1, \dots, d$
- $G_1, \dots, G_d$: continuous df's on $\mathbb{R}$, strictly increasing on $\mathrm{I_i} = \{0 < \mathrm{G_i} < 1\}$
- Define transformation:

$$K(x_1, \dots, x_d) = \big(K_1(x_1), \dots, K_d(x_d)\big), \quad K_i(s) = F_i^{-1}\big(G_i(s)\big), \quad i = 1, \dots, d$$

Definition (Meta distribution)

- Random vector $\mathbf{Z}$ in $\mathbb{R}^d$ with df F and continuous marginals F_i, $i = 1, \ldots, d$
- $G_1, \ldots, G_d$: continuous df's on $\mathbb{R}$, strictly increasing on $\mathrm{I_i} = \{0 < \mathrm{G_i} < 1\}$
- Define transformation:

$$K(x_1, \ldots, x_d) = \big(K_1(x_1), \ldots, K_d(x_d)\big), \quad K_i(s) = F_i^{-1}\big(G_i(s)\big), \quad i = 1, \ldots, d$$

- The df $G = F \circ K$ is the meta distribution (with marginals G_i) based on original df F

- The coordinatewise map $K = K_1 \otimes \cdots \otimes K_d$ which maps $\mathbf{x} = (x_1, \ldots, x_d) \in \mathrm{I} = \mathrm{I_1} \times \cdots \times \mathrm{I_d}$ into the vector $\mathbf{z} = (K_1(x_1), \ldots, K_d(x_d))$ is called the meta transformation

Definition (Standard set-up)

In the standard set-up, the meta density g is based on the original density f and has marginals which are all equal to g_0, where

- $f \sim f_*(n_D(\mathbf{z}))$ with
  - $f_* \in RV_{-(\lambda+d)}$ for some $\lambda > 0$, continuous and decreasing on $[0, \infty)$
  - n_D: the gauge function of the set D
  - D: bounded, open, star-shaped set with a continuous boundary and containing the origin
- g_0 is continuous, positive, symmetric, and asymptotic to a von Mises function $e^{-\psi}$ with $\psi \in RV_\theta$ for $\theta > 0$

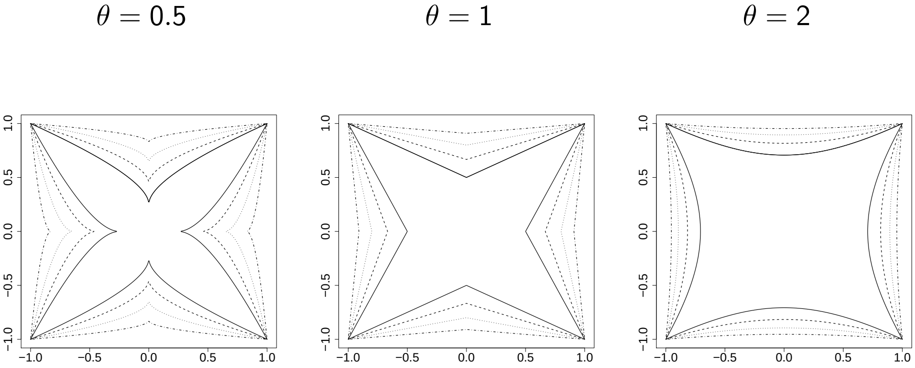
Theorem

- *Let f, g and $g_0 \sim e^{-\psi}$ satisfy the assumptions of the standard set-up*
- *Let $\psi(s_n) \sim \log n$*

For the sequence of independent observations $\mathbf{X}_n$ from the meta density g, the sample clouds $N_n = \{\mathbf{X}_1/s_n, \dots, \mathbf{X}_n/s_n\}$ converge onto the limit set $E_{\lambda,\theta}$ given by

$$E_{\lambda,\theta} := \{\mathbf{u} \in \mathbb{R}^d \mid |u_1|^\theta + \cdots + |u_d|^\theta + \lambda \geq (\lambda + d)\|\mathbf{u}\|_\infty^\theta\}$$

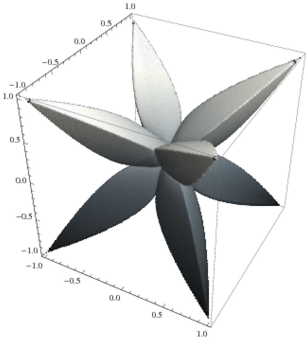
Examples of limit sets $E_{\lambda,\theta}$ in $\mathbb{R}^2$



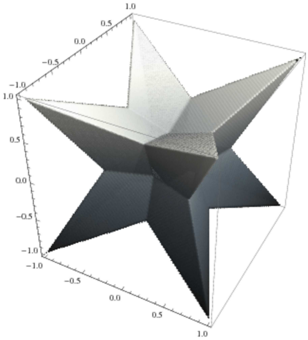
Legend: $\lambda = 1$ (solid), $\lambda = 2$ (dashed), $\lambda = 4$ (dotted), $\lambda = 10$ (dotdash)

Examples of limit sets $E_{\lambda,\theta}$ in $\mathbb{R}^3$ ($\lambda = 1$)

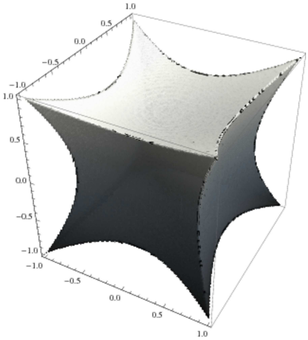
$\theta = 0.5$



$\theta = 1$



$\theta = 2$



- *How stable is the shape of the limit set?*
  - How much the original distribution can be altered without affecting its marginal tail behaviour and the shape of the limit set?
  - How sensitive is the shape of the limit set to perturbations of the original distribution?

- *How stable is the shape of the limit set?*
  - How much the original distribution can be altered without affecting its marginal tail behaviour and the shape of the limit set?
  - How sensitive is the shape of the limit set to perturbations of the original distribution?

- We investigate above questions using two procedures:
  - block partitions
  - mixtures

Assumption:

marginal dfs of F are all equal to F_0 with positive continuous symmetric density

$\Rightarrow$ meta transformation K has equal components:

$$K : \mathbf{x} \mapsto \mathbf{z} = (K_0(x_1), \ldots, K_0(x_d))$$

$$K_0 = F_0^{-1} \circ G_0 \qquad K_0(-t) = -K_0(t)$$

Theorem 1

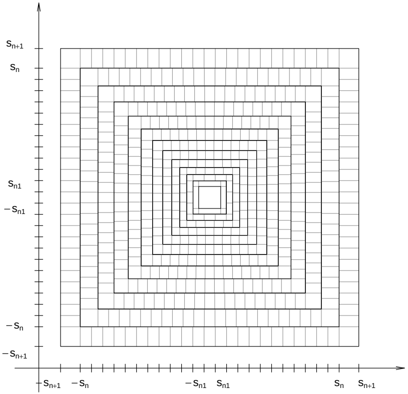
- Assume the standard set-up
- Let $\tilde{\rho}$ be an excess measure on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ with marginal densities $\lambda/|t|^{\lambda+1}$, $\lambda > 0$

One may choose $\tilde{F}$ such that

- sample clouds converge to Poisson point process $\tilde{N}$ with mean measure $\tilde{\rho}$
- marginals are tail asymptotic to those of F
- meta dfs $\tilde{G} = \tilde{F} \circ K$ and G have the same asymptotics:
  - the sample clouds converge onto $E_{\lambda,\theta}$ with the same scaling
  - the marginals are tail asymptotic

Block partitions - Construction

- Partitions of $\mathbb{R}^d$ into bounded Borel sets B_n, where B_n are coordinate blocks
- Start with an increasing sequence of cubes: $s_n C = [-s_n, s_n]^d$ with $0 < s_1 < s_2 < \cdots, \; s_n \to \infty$
- Subdivide ring $R_n = s_{n+1}C \setminus s_n C$ into blocks by a symmetric partition of interval $[-s_{n+1}, s_{n+1}]$ with division points $\pm s_{nj}$, $j = 1, \ldots, m_n$, with



$$-s_{n+1} < -s_n < \cdots < -s_{n1} < s_{n1} < \cdots < s_{nm_n} = s_n < s_{nm_n+1} = s_{n+1}$$

Regular block partitions

- If blocks are relatively small, the asymptotic behaviour of a distribution is not affected if the distribution is replaced by one which gives (asymptotically) the same mass to each block

  - **regular block partitions**: a block partition is regular if and only if

$$s_{n+1} \sim s_n \text{ and } \Delta_n / s_n \to 0$$

where $\Delta_n = \max\{s_{n1}, s_{n2} - s_{n1}, \ldots, s_{nm_n} - s_{nm_n - 1}\}$

Proof of Theorem 1 - sketch of construction

Notation:

(A_n): block partition in **x**-space with division points s_n

(B_n): block partition in **z**-space with division points $t_n = K_0(s_n)$

- Suppose (A_n) is regular ($\Rightarrow \quad s_{n+1} \sim s_n$)
  It is possible that $t_n << t_{n+1}$ ($\Rightarrow \quad (B_n)$ is non-regular)

- Choose $s_{nm_n-1} = s_{n-1}$; define sets:
  - $U = \bigcup_n [t_{n-1}, t_{n+1}]^d$
  - Note: $U/t \to (0, \infty)^d$ for $t \to \infty$ if $t_n << t_{n+1}$
  - Let U_δ be the image of U in orthant Q_δ for $\delta \in \{-1, 1\}^d$

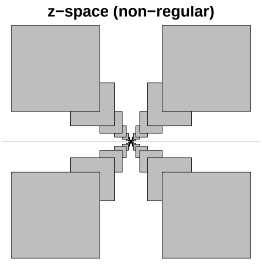
Sketch of construction (cont'd)

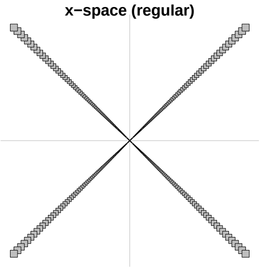
- Choose density $\hat{f}$ such that sample clouds converge to $\tilde{N}$ with mean measure $\tilde{\rho}$
- Let $\tilde{f}$ agree with $\hat{f}$ on sets U_δ and with f elsewhere $\Rightarrow$ $\tilde{f}$ and $\hat{f}$ differ only on asymptotically negligible set
- Alter $\tilde{f}$ on a bounded set to make it a probability density

$\Rightarrow$ sample clouds from $\tilde{F}$ converge to $\tilde{N}$

- In corr. partition (A_n) on **x**-space, the measure is changed only on "tiny" blocks $[s_{n-1}, s_{n+1}]^d$ (with $s_{n-1} \sim s_{n+1}$) and their reflections

$\Rightarrow$ sample clouds from $\tilde{G} = \tilde{F} \circ K$ converge onto $E_{\lambda,\theta}$





Theorem 2

- Assume the standard set-up
- $A \subset [-1,1]^d$: a star-shaped closed set with continuous boundary and containing the origin as interior point
- $E_{00} = \{r\delta \mid 0 \le r \le 1,\ \delta \in \{-1,1\}^d\}$: the diagonal cross

There exists a df $\tilde{F}$ such that

- $\tilde{F}$ and F have the same asymptotics:
  - the sample clouds converge to the same point process
  - the marginals are tail asymptotic
- the sample clouds from the meta distribution $\tilde{G}$ converge onto the set $E = A \cup E_{00}$

Idea behind a proof of Theorem 2

- Replace dF by a probability measure $d\tilde{F}$ which agrees outside a bounded set with $d(F + F^o)$, where F^o has lighter marginals than F:

$$F_j^o(-t) << F_0(-t), \quad 1 - F_j^o(t) << 1 - F_0(t), \quad t \to \infty \; (j = 1, \ldots, d)$$

$\Rightarrow$ $\tilde{F}$ and F have the same asymptotics

- dfs $\tilde{G}$ and G may have different asymptotics:
  - the scaling constants $a_n^o \sim a_n$ even though G^o has lighter tails than G
  - suppose sample clouds from G^o converge onto a compact set E^o

then: sample clouds from $\tilde{G}$ converge onto $E_{\lambda,\theta} \cup E^o$

Some concluding remarks

- Distributions with the same asymptotics in one space (**x**- or **z**-space) may give rise to different asymptotic behaviour in the other space
  - Sensitivity of the limit shape may be radical due to even slight perturbations of the original distribution, *perturbations which do not affect the asymptotics of the coordinatewise extremes or the marginals*
- It is possible to manipulate the shape of the limit set without affecting the distribution of coordinatewise extremes

Part II

Asymptotic independence for unimodal densities

References:

A.A. Balkema and N. Nolde (2010). Asymptotic independence for unimodal densities. *Advances in Applied Probability*: 42(2), in press.

N. Lysenko, P. Roy and R. Waeber (2009). Multivariate extremes of generalized skew-normal distributions. *Statist. Probab. Let.* 79:525-533.

Early work goes back to 1960's...

Masaaki Sibuya

For bivariate data, asymptotic independence means:

large values in one coordinate are unlikely to be accompanied by large values in the other coordinate

Definition *(Sibuya's condition)*

Let X and Y have continuous dfs F_1 and F_2, resp.

Define a function P on $[0,1]^2$ via

$$P(F_1(x), F_2(y)) = \mathbb{P}\{X > x, Y > y\}$$

X and Y are asymptotically independent if and only if

$$P(1-s, 1-s) = o(s) \qquad s \to 0^+ \qquad \text{[Sibuya (1960)]}$$

Motivation (1/2): Relevance for applications

- The concept of asymptotic independence has been used in many applications
  - e.g. modelling of environmental, financial, and traffic network data
- Model risk: not being aware of the properties of a model with respect to extremal dependence may potentially lead to risk underestimation
  - e.g. the Gaussian copula model and the recent financial crisis

Motivation (2/2): Densities vs. distribution functions (dfs)

- "Simple" to check asymptotic independence using dfs as the standard criteria are given in terms of dfs
- However, explicit expressions for dfs are not always available
- Typically we are given a density in analytic form
- For a light-tailed density, the shape of sample clouds may give a good geometric image of the asymptotic shape of its level sets

Motivation (2/2): Densities vs. distribution functions (dfs)

- "Simple" to check asymptotic independence using dfs as the standard criteria are given in terms of dfs
- However, explicit expressions for dfs are not always available
- Typically we are given a density in analytic form
- For a light-tailed density, the shape of sample clouds may give a good geometric image of the asymptotic shape of its level sets

Objective:

Give sufficient conditions for asymptotic independence in terms of the (asymptotic) shape of the level sets of the density or of suitably scaled sample clouds

The class $\mathcal{A}(D)$

- $\mathcal{D}_d$: the class of all bounded open star-shaped sets $D \subset \mathbb{R}^d$ for which
  - the cone $D_\infty = \bigcup_{n>0} nD$ is convex
  - the gauge function n_D is continuous on D_∞

- A positive density g on $\mathbb{R}^d$ belongs to the class $\mathcal{A}(D)$ if
  - $D \in \mathcal{D}_d$
  - there exist sequences $c_n > 0$ and $r_n \to \infty$ with $c_{n+1}/c_n \to 0$ and $r_{n+1} \sim r_n$ such that for any $\epsilon > 0$ eventually

$$\boxed{e^{-\epsilon} r_n D \subset \{g > c_n\} \subset e^{\epsilon} r_n D} \qquad (\star)$$

  - Write $\{g > c_n\}/r_n \to D$

- A continuous positive function $\tilde{g}$ is shape equivalent to g if its level sets satisfy $(\star)$

The class $\mathcal{A}(D)$: Shape-equivalent densities

- How large is the class $\mathcal{A}(D)$?

Proposition:

Suppose $f_* = e^{-\varphi}$, where φ is continuous, strictly increasing and varies regularly at ∞ with exponent $\theta > 0$

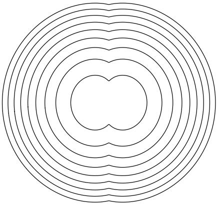
Then $g = e^{-\gamma} \in \mathcal{A}(D)$ is shape equivalent to $f = f_*(n_D)$ if and only if $\gamma(\mathbf{x}_n) \sim \varphi(n_D(\mathbf{x}_n))$ whenever $\|\mathbf{x}_n\| \to \infty$

The class $\mathcal{A}(D)$: Shape-equivalent densities (cont'd)

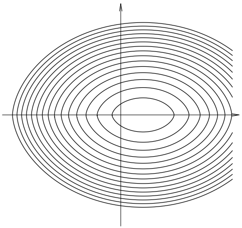
- The behaviour of the quotient of two shape equivalent functions may be quite erratic

Examples: $\tilde{f} = qf$ is shape equivalent to standard normal density f on $\mathbb{R}^2$ for the following choices of q:

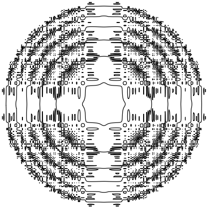
$q = 1 + |x|$



$q = \exp(x - |y|^{3/2})$



$q = \exp\{\sin(\pi e^{x^2}) \sin(\pi e^{y^2})\}$



Criteria for asymptotic independence - I

Definition

The set $D \in \mathcal{D}_2$ is blunt if the coordinatewise supremum of the points in D does not lie in the closure of D

Criteria for asymptotic independence - I

Definition

The set $D \in \mathcal{D}_2$ is blunt if the coordinatewise supremum of the points in D does not lie in the closure of D

Theorem 1

- *Suppose $\mathbf{X} = (X_1, \ldots, X_d)$ has density $f \in \mathcal{A}(D)$*

If the bivariate projection D_{12} of the set D on x_1, x_2-plane is blunt, then X_1 and X_2 are asymptotically independent

Criteria for asymptotic independence - II

Theorem 2

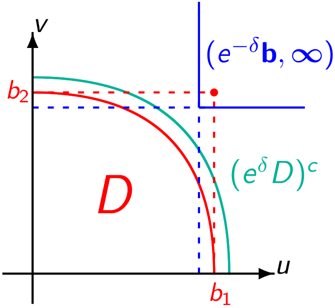
- $\mathbf{X}_1, \mathbf{X}_2, \ldots$ *i.i.d. observations from a continuous df F on* $\mathbb{R}^2$
- *Suppose there exist scaling constants* s_n *such that the sample clouds* $N_n = \{\mathbf{X}_1/s_n, \ldots \mathbf{X}_n/s_n\}$ *converge onto the closure of* $D \in \mathcal{D}_2$

If D is blunt then the two coordinates of the vector $\mathbf{X}_1$ *are asymptotically independent*

Proof of Theorem 2 - outline

- Let $\mathbf{b} = (b_1, b_2) := \sup D$

D blunt $\Rightarrow$ $\exists\delta > 0$: $(e^{-\delta}\mathbf{b}, \infty) \cap e^{\delta} D = \emptyset$



- Let $n\pi_n$ denote the mean measure of N_n

$$n\pi_n(e^{-\delta}\mathbf{b}, \infty) \le n\pi_n(e^{\delta}D)^c \to 0$$
$$n\pi_n(D \cap \{u > e^{-\delta} b_1\}) \to \infty$$
$$n\pi_n(D \cap \{v > e^{-\delta} b_2\}) \to \infty$$

since N_n converges onto D

- Let $\mathbf{W}_n = (U_n, V_n)$ denote the componentwise maximum of N_n

$$\mathbb{P}\{\mathbf{W}_n \in N_n\} \le \mathbb{P}\{N_n((e^{-\delta}\mathbf{b}, \infty)) > 0\} + \mathbb{P}\{N_n(\{u > e^{-\delta} b_1\}) = 0\} + \mathbb{P}\{N_n(\{v > e^{-\delta} b_2\}) = 0\} \to 0$$

[criterion due to Gnedin (1994)]

Example

- A *d*-dimensional skew-normal distribution with positive-definite scale matrix $\Omega \in \mathbb{R}^{d \times d}$ and shape parameter $\boldsymbol{\alpha} \in \mathbb{R}^d$ has a density of the form
$$g(\mathbf{x}) = 2\phi_d(\mathbf{x}; \mathbf{0}, \Omega)\Phi(\boldsymbol{\alpha}^T \mathbf{x}) \qquad \mathbf{x} \in \mathbb{R}^d$$
where $\phi_d(\,\cdot\,; \boldsymbol{\mu}, \Sigma)$ and $\Phi(\cdot)$ denote the d-dimensional normal density with mean $\boldsymbol{\mu}$ and covariance matrix Σ, and the univariate standard normal df, resp.

- It can be shown that $g \in \mathcal{A}(D)$ with $D = \{\mathbf{u} \in \mathbb{R}^d \mid n_D(\mathbf{u}) < 1\}$, where
$$n_D(\mathbf{u}) = \begin{cases} \mathbf{u}^T \Omega^{-1} \mathbf{u}, & \boldsymbol{\alpha}^T \mathbf{u} \geq 0, \\ \mathbf{u}^T \Omega^{-1} \mathbf{u} + (\boldsymbol{\alpha}^T \mathbf{u})^2, & \boldsymbol{\alpha}^T \mathbf{u} < 0 \end{cases}$$

- D has blunt bivariate projections (in fact, convex with C^1 boundary) and hence the skew-normal distribution is asymptotically independent
  - cf. *[Lysenko, Roy, Waeber (2009)]*

Future work:

- A more delicate question is the relation between the shape and asymptotic independence when the bivariate projections are not blunt
- Criteria for asymptotic dependence for light-tailed densities

Additional references

Balkema, G., and Embrechts, P. (2007). *High Risk Scenarios and Extremes. A geometric approach.* European Mathematical Society.

Gnedin, A.V. (1994). On the best choice problem with dependent criteria. *J. Appl. Probab.* 31:221-234.

Sibuya, M. (1960). Bivariate extreme statistics. *Ann. Inst. Stat. Math.* 11:195-210.

Thank you!