Multivariate Heavy Tailed Phenomena: Modeling, Diagnostics and Applications

Sidney Resnick
School of Operations Research and Information Engineering
Rhodes Hall, Cornell University
Ithaca NY 14853 USA

http://people.orie.cornell.edu/∼sid
sir1@cornell.edu sidresnick@gmail.com

MURI Review: Cornell

September 30, 2013

1. Hidden Regular Variation.

(Joint with Filip Lindskog, Joyjit Roy.)

Goal:

- Find a framework in which multiple multivariate heavy tail properties exist simultaneously; improve risk estimates.
- Give significant examples illustrating this framework.

Given a *risk vector*

$$\boldsymbol{X} = (X_1, \dots, X_d), \qquad d \leq \infty,$$

estimate the probability of a *risk region* $\mathcal{R}$

$$P[\boldsymbol{X} \in \mathcal{R}]$$

where $\mathcal{R}$ is beyond the range of observed data. Solution based on asymptotics: Assumption of multivariate heavy tails:

- Choose a state space; for example
  - $\mathbb{E} = [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$, or
  - $\mathbb{E} = (\mathbf{0}, \boldsymbol{\infty})$.

- Choose a scaling function $b(t) \to \infty$ appropriate for $\mathbb{E}$ such that that in $M_+(\mathbb{E})$,

$$nP\Big[\frac{\boldsymbol{X}}{b(n)} \in \cdot\Big] \to \nu(\cdot). \qquad \text{(DOA)}$$

Issues:

- What is "$\to$"?
- What is $\mathbb{E}$? $M_+(\mathbb{E})$?
  - Traditionally used one point uncompactification and vague convergence.
  - Has major drawbacks.
  - We now have a flexible framework in CSMS.
  - $\mathbb{E}$ and the scaling function $b(\cdot)$ are always related.
- Specify $\nu(\cdot)$ and estimate.

Quasi-solution to estimation problem:

$$P[\boldsymbol{X} \in \mathcal{R}] \approx \frac{1}{n}\hat{\nu}(\mathcal{R}/\hat{b}(n)).$$

But

- $\nu(\cdot)$ may concentrate on a small region of the state space.
- There may be simultaneous regular variation properties coexisting under different choices of state space $\mathbb{E}$ and scaling functions $b(\cdot)$.

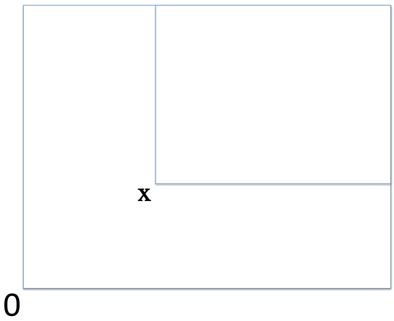
Example:

$d = 2$ and

$\mathcal{R} = (\mathbf{x}, \boldsymbol{\infty}] = (x_1, \infty] \times (x_2, \infty]$

and

$P[\boldsymbol{X} \in \mathcal{R}] = P[X_1 > x_1, X_2 > x_2].$



Risk contagion: Can two or more components of the risk vector $\boldsymbol{X}$ be simultaneously large?

Ambiguity: Should we do the approximation assuming $\mathbb{E} = [0, \infty)^2 \setminus \{\mathbf{0}\}$ or assuming $\mathbb{E} = (0, \infty)^2$.

2. Asymptotic independence

- If $\mathbb{E} = [0, \infty)^d \setminus \{\mathbf{0}\}$, asymptotic independence means

$$\nu(\{\mathbf{x} : x_i > y_i,\ x_j > y_j), \}) = 0, \quad y_i > 0,\, y_j > 0. \quad \text{(AsyIndep)}$$

  for all $1 \leq i < j \leq d$ and thus such an *asymptotic* model has no risk contagion since we estimate

$$P[\text{ two or more components of } \boldsymbol{X} \text{ are large simultaneously }] \approx 0.$$

- Can we improve on this asymptotic method?

2.1. (AsyIndep) not uncommon:

- $\boldsymbol{X}$ has independent components.
- Gaussian copula model: Heavy tailed marginals but Gaussian dependence with

$$\mathrm{corr}(X_i, X_j) = \rho(i, j) < 1.$$

- Let $U \sim U(0, 1)$ and

$$\boldsymbol{X} = \Big(\frac{1}{U}, \frac{1}{1-U}\Big).$$

2.2. Standard Example.

For $d = 2$: If $\boldsymbol{X} = (X_1, X_2)$ and $X_1 \perp\!\!\!\perp X_2$, X_1, X_2 iid with

$$P[X_i > y] = y^{-1}, \quad y > 1.$$

- Then for $x_1 > 0$, $x_2 > 0$, as $n \to \infty$

$$nP[X_i > nx_i] \to x_i^{-1}, \qquad i = 1, 2,$$
$$nP[X_1 > nx_1, X_2 > nx_2] \to 0,$$

  so $\boldsymbol{X}$ is regularly varying on $\mathbb{E} = [0, \infty)^2 \setminus \{\mathbf{0}\}$ with index 1 and limit measure concentrating on the axes.

- Also for $x_1 > 0, x_2 > 0$,

$$\sqrt{n}P[X_1 > \sqrt{n}x_1] \cdot \sqrt{n}P[X_2 > \sqrt{n}x_2] =$$
$$nP[X_1 > \sqrt{n}x_1, X_2 > \sqrt{n}x_2] \to \frac{1}{x_1 x_2},$$

  so $\boldsymbol{X}$ is regularly varying on $\mathbb{E}_0 = (0, \infty)^2 \setminus \{\mathbf{0}\}$ with index 2 and limit measure giving positive mass to $(\mathbf{x}, \boldsymbol{\infty}]$.

Conclude for this example:

- $\boldsymbol{X}$ is regularly varying on $\mathbb{E} = [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$ with index 1 (scale by n) and limit measure concentrating on lines through $\{\mathbf{0}\}$, and giving zero mass to $(\mathbf{0}, \boldsymbol{\infty})$.
- $\boldsymbol{X}$ is regularly varying on $\mathbb{E}_0 = (\mathbf{0}, \boldsymbol{\infty})$ with index 2 (scale by $\sqrt{n}$) and the limit measure gives positive mass to $(\mathbf{0}, \boldsymbol{\infty})$.

Some progress (in low dimensions):

- Detection.
- Estimation.
- Non-parametric technique using rank methods detects hidden structure.
- Framework applicable to $\mathbb{R}^\infty$, $C[0,1]$, $D[0,1]$.
- Examples where an infinite number of regular variation properties coexist
  - iid with regularly varying marginals.
  - Lévy processes with regularly varying Lévy measure.
- Framework includes classical theory as well as the conditional extreme value model (condition on one component of a vector being large).

3. Framework: Regular variation on cones in CSMS (with Lindskog & Roy).

Context: Consider random element $\boldsymbol{X}$ of a complete separable metric space $\mathbb{S}$ with an origin and scaling operation: Examples of $\mathbb{S}$:

- $\mathbb{R}_+$; $\boldsymbol{X}$ is a random variable,
- $\mathbb{R}_+^d$; $\boldsymbol{X}$ is a random vector,
- $\mathbb{R}_+^\infty$; $\boldsymbol{X}$ is a random sequence,
- $D[0,1]$=càdlàg space; $\boldsymbol{X}$ is a càdlàg process such as a Lévy process.

Suppose $\mathbb{F}_1 \subset \mathbb{S}$ closed (cone) containing $\mathbf{0}$ and define the TABOF space

$$\mathbb{S}_{\mathbb{F}_1} = \mathbb{S} \setminus \mathbb{F}_1.$$

→ The random element $\boldsymbol{X} \in \mathbb{S}$ has a distribution with a regularly varying tail on $\mathbb{S}_{\mathbb{F}_1}$ if $\exists\, b(t) \uparrow \infty$ and measure $\nu \not\equiv 0$ on $\mathbb{S}_{\mathbb{F}_1}$ such that

$$tP\Big[\frac{\boldsymbol{X}}{b(t)} \in \cdot\Big] \to \nu(\cdot), \qquad \text{in } M^*(\mathbb{S}_{\mathbb{F}_1}).$$

[Must define topology on $M^*(\mathbb{S}_{\mathbb{F}_1})$, the measures on $\mathbb{S}_{F_1}$ that are finite on sets at positive distance from F_1; fairly routine.]

Let $\mathbb{F}_2$ be another closed (cone) containing $\mathbf{0}$ and set

$$\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2} = \mathbb{S} \setminus (\mathbb{F}_1 \cup \mathbb{F}_1).$$

→ The random $\boldsymbol{X}$ has a distribution with hidden regular variation on $\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}$ if there is regular variation on $\mathbb{S}_{\mathbb{F}_1}$ AND if $\exists\, b_1(t) \uparrow \infty$ and a measure $\nu_1(\cdot) \not\equiv 0$ on $\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}$ such that

$$tP\Big[\frac{\boldsymbol{X}}{b_1(t)} \in \cdot\Big] \to \nu_1(\cdot), \qquad \text{in } M^*(\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}),$$

AND

$$b(t)/b_1(t) \to \infty$$

(which makes the behavior on $\mathbb{S}_{\mathbb{F}_1 \cup \mathbb{F}_2}$ hidden).

3.0.1. Example: iid on $\mathbb{R}_+^\infty$.

$\mathbb{S} = \mathbb{R}_+^\infty$, $\mathbb{F} = \mathbb{F}^{(j)}$, $j \geq 0$, where

$$\begin{aligned}\mathbb{F}^{(j)} =& \{\mathbf{x} := (x_1, x_2, \dots) \in \mathbb{R}_+^\infty : \sum_{j=1}^\infty \epsilon_{x_j}(0,\infty) \leq j\} \\ =& \{\mathbf{x} : \text{ at most } j \text{ components} > 0\}.\end{aligned}$$

So

$$\begin{aligned}\mathbb{F}^{(0)} =& \{\mathbf{0}_\infty\} \\ \mathbb{F}^{(1)} =& \text{axes in } \mathbb{R}_+^\infty \text{ through } \mathbf{0}, \text{ including } \mathbf{0} \\ =& \bigcup_{j=1}^\infty \{0\}^{j-1} \times (0,\infty) \times \{0\}^\infty \cup \{\mathbf{0}_\infty\}, \\ & \vdots\end{aligned}$$

Leads to a sequence of spaces:

$$\underbrace{\mathbb{R}_+^\infty \setminus \mathbb{F}^{(0)}}_{\text{remove } \mathbf{0}_\infty} \supset \underbrace{\mathbb{R}_+^\infty \setminus \mathbb{F}^{(1)}}_{\text{remove axes}} \supset \underbrace{\mathbb{R}_+^\infty \setminus \mathbb{F}^{(2)}}_{\text{remove 2-dim faces}} \supset \dots \quad \dots.$$

Suppose,

- $\mathbb{S} = \mathbb{R}_+^\infty$;
- $\boldsymbol{X} = (X_1, X_2 \dots)$ has iid components with each X_i having a regularly varying tail with scaling function $b(t)$. Means:

$$tP[X_j > b(t)x] \to \nu_\alpha(x, \infty) = x^{-\alpha}, \quad t \to \infty, \alpha > 0,$$

and $b(t)$ satisfies,

$$b(t) = \Big(\frac{1}{P[X_j > \cdot]}\Big)^{-1}(t), \quad P[X_j > b(t)] \sim \frac{1}{t}.$$

Then as $t \to \infty$, for $j \geq 1$

$$tP[\boldsymbol{X}/b(t^{1/j}) \in \cdot] \to \mu^{(j)} \text{ in } \mathbb{M}(\mathbb{R}_+^\infty \setminus \mathbb{F}^{(j-1)})$$

and $\mu^{(j)}$ concentrates on $\mathbb{F}^{(j)} \setminus \mathbb{F}^{(j-1)}$, the sequences with *exactly* j positive components.

Recent extensions:

1. J. Roy extends this to study HRV for moving averages of the form

$$X_n = \sum_{j=0}^{\infty} \psi_j Z_{n-j}, \qquad n \in \mathbb{Z}$$

   where $\{Z_n\}$ is a doubly infinite positive iid sequence with regularly varying marginals and $\psi_j \geq 0$.

   As a random element of the CSMS $\prod_{i=-\infty}^{\infty} \mathbb{R}$ of double sided sequences,

   - $\boldsymbol{X} = \{X_n\}$ is regularly varying and
   - an infinite number of regular variation properties coexist.

2. Alternative method of counteracting tendency toward asymptotic independence by gradually increasing dependence at the correct rate. Related to

   - Hüsler–Reiss distributions and
   - CEV model.

4. Challenges.

- Practical?
  - Limitations of asymptotic methods: rates of convergence?
  - Instead of estimating a risk probability as 0, estimate is a very small number.
- Need for more formal inference for estimation including confidence statements.
- General HRV technique requires knowing the support of the limit measure. Estimate support?
- High dimension problems? How to sift through different possible subcones? There could be a sequence of cones with regular variation on each. How to teach a computer to find the cones?
- How to go from standard to more realistic non-standard case where components not scaled the same; still some inference problems.

5. Don't know! Coming year.

- Core issues (even in 2 dimensions):
  - What does is mean for

$$P[X = i, Y = j] =: f_{ij}$$

    to be multivariate regularly varying (have a power law)?
  - Embedding problem: Assuming we know what it means for f_{ij} to be regularly varying, when does there exist a regularly varying function $U(x, y)$ of continuous variables such that

$$f_{ij} = U(i, j)?$$

    (Always true in one dimension.)
  - When does this imply the measure

$$P[(X, Y) \in \cdot\,]$$

    is regularly varying?
    Note: The integral of a regularly varying function $u(s, t)$

$$U(x, y) = \int_0^x \int_0^y u(s, t)dsdt,$$

    is not necessarily regularly varying. (Always true in one dimension.)

- Tauberian theorems in higher dimensions are known in restricted circumstances. Needed for network models where limiting frequencies are given in terms of generating functions.

- What is the origin of heavy tails? For reasonable models, presumably, as $n \to \infty$,

$$\frac{N_n(i,j)}{n} \to f_{i,j}$$
$$\frac{\mathrm{CT}_n}{n} \to \mathrm{CT}_\infty =: \sum_{(i,j)} f_{i,j} \epsilon_{(i,j)}.$$

  where

  $N_n(i,j) = \#\{\text{nodes at time } n \text{ with in-degree} = i, \text{out-degree} = j\}.$

  - Generalize sexy martingale arguments used for undirected graphs to prove the limits exist for directed graphs.
  - Find broad assumptions under which $f_{i,j}$ exhibit power law behavior. (Nobody really believes networks evolve like, eg, Krapivsky.)
  - Use a martingale central limit theorem (?) to understand asymptotic normality of $\frac{N_n(i,j)}{n}$; gateway to inference?

- Continue to think about MLE or other forms of estimation as applied to network data.
- Understand the role of censoring in data such as *slashdot*.

Contents