New developments for a geometric approach to multivariate extremal inference

Ryan Campbell

with Jennifer L. Wadsworth

Lancaster University

UQÀM

12 September 2024

Multivariate Extreme Value Theory (MEVT)

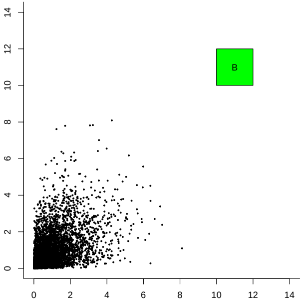
- Interested in $\Pr(\boldsymbol{X} \in B)$ for $\boldsymbol{X} \in \mathbb{R}^d$.

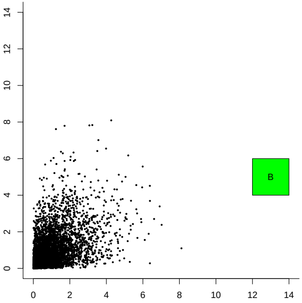
Multivariate Extreme Value Theory (MEVT)

- Interested in $\Pr(\boldsymbol{X} \in B)$ for $\boldsymbol{X} \in \mathbb{R}^d$.
- B not in range of data

Multivariate Extreme Value Theory (MEVT)

- Interested in $\Pr(\boldsymbol{X} \in B)$ for $\boldsymbol{X} \in \mathbb{R}^d$.
- B not in range of data



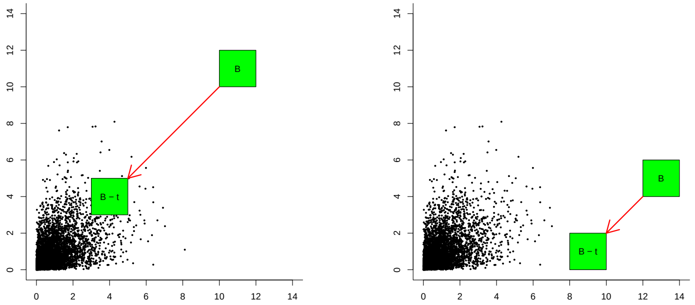


MEVT: Multivariate Regular Variation

- Suppose $\boldsymbol{X}_E = (X_{E,1}, \ldots, X_{E,d})^\top$, $X_{E,j} \sim \text{Exp}(1)$.
- Perform *extrapolation*: $\Pr(\boldsymbol{X}_E \in B) \approx e^{-t}\Pr(\boldsymbol{X}_E \in B - t)$

MEVT: Multivariate Regular Variation

- Suppose $\boldsymbol{X}_E = (X_{E,1}, \ldots, X_{E,d})^\top$, $X_{E,j} \sim \text{Exp}(1)$.
- Perform *extrapolation*: $\Pr(\boldsymbol{X}_E \in B) \approx e^{-t}\Pr(\boldsymbol{X}_E \in B - t)$
- de Haan (1970), has drawbacks...

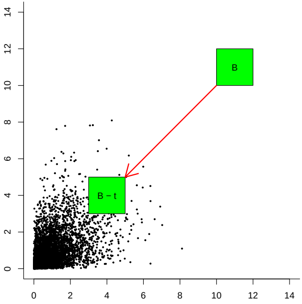


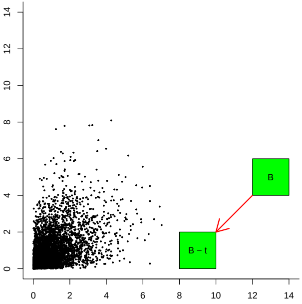
MEVT: Hidden Regular Variation

- Introduce η to correct tail decay.
- Perform *extrapolation*: $\Pr(\boldsymbol{X}_E \in B) \approx e^{-t/\eta} \Pr(\boldsymbol{X}_E \in B - t)$

MEVT: Hidden Regular Variation

- Introduce η to correct tail decay.
- Perform *extrapolation*: $\Pr(\boldsymbol{X}_E \in B) \approx e^{-t/\eta} \Pr(\boldsymbol{X}_E \in B - t)$
- Ledford and Tawn (1997), also has drawbacks...





MEVT: Conditional Extremes

- Given $\boldsymbol{X} = (X, Y)$, model $Y|X > x$, x large
- Heffernan and Tawn (2004)

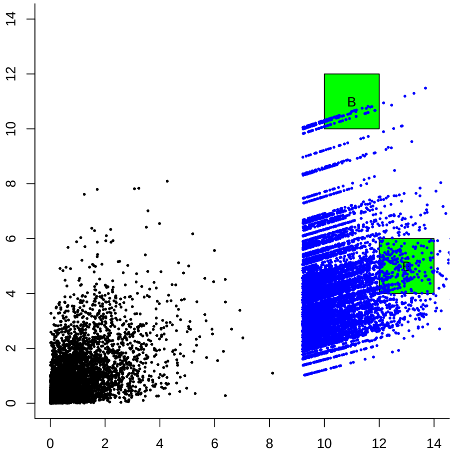
MEVT: Conditional Extremes

- Given $\boldsymbol{X} = (X, Y)$, model $Y|X > x$, x large
- Heffernan and Tawn (2004)
- Relies on false working assumption

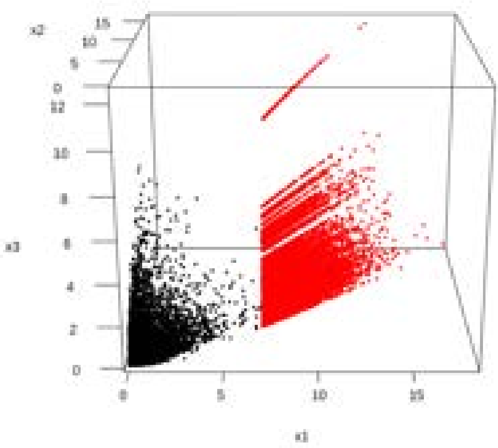
$$\frac{Y - \alpha x}{x^{\beta}}\bigg| X > x \sim \mathcal{N}(0, 1)$$

- Complicated inference for $d > 2$.
- Doesn't capture complex dependence structures.
- Requires prior knowledge of dependence structure.

MEVT: Conditional Extremes



MEVT: Conditional Extremes



Limit sets and gauge functions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} f_{\boldsymbol{X}_E}$, d-dimensional, Exponential(1) margins.

Limit sets and gauge functions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} f_{\boldsymbol{X}_E}$, d-dimensional, Exponential(1) margins.
- If not in Exponential(1) margins, can standardize using

$$\begin{aligned} X_{E,j} &= F_E^{-1}\left(F_{X_j}(X_j)\right) \\ &= -\log\left(1 - F_{X_j}(X_j)\right) \end{aligned}$$

Limit sets and gauge functions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} f_{\boldsymbol{X}_E}$, d-dimensional, Exponential(1) margins.
- If not in Exponential(1) margins, can standardize using

$$\begin{aligned} X_{E,j} &= F_E^{-1}\left(F_{X_j}(X_j)\right) \\ &= -\log\left(1 - F_{X_j}(X_j)\right) \end{aligned}$$

- Scaled sample clouds $\left\{\frac{\boldsymbol{X}_1}{\log n}, \ldots, \frac{\boldsymbol{X}_n}{\log n}\right\}$ converge onto a **limit set**,

$$G := \left\{\boldsymbol{x} \in \mathbb{R}^d \,\middle|\, g(\boldsymbol{x}) \leq 1\right\}$$

as $n \to \infty$ (Balkema and Nolde, 2010).

Limit sets and gauge functions

- The **gauge function**, $g : \mathbb{R}^d \longrightarrow \mathbb{R}$, is 1-homogeneous and is obtained through:

$$g(\boldsymbol{x}) = \lim_{t\to\infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t}$$

Limit sets and gauge functions

- The **gauge function**, $g : \mathbb{R}^d \longrightarrow \mathbb{R}$, is 1-homogeneous and is obtained through:

$$g(\boldsymbol{x}) = \lim_{t \to \infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t}$$

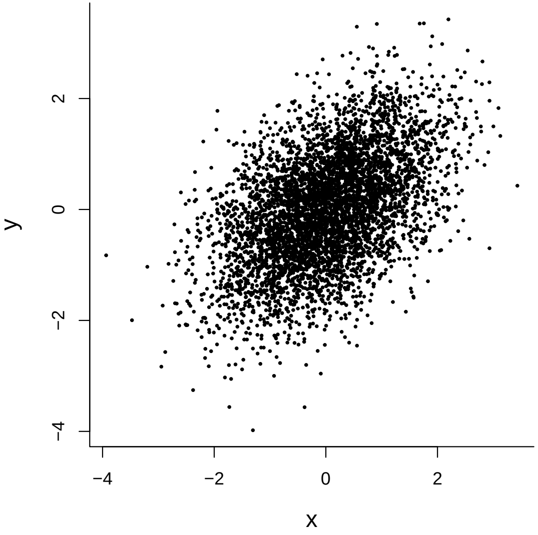
- Nolde (2014); Nolde and Wadsworth (2022) use g to describe extremal dependence properties.

Example: Gaussian

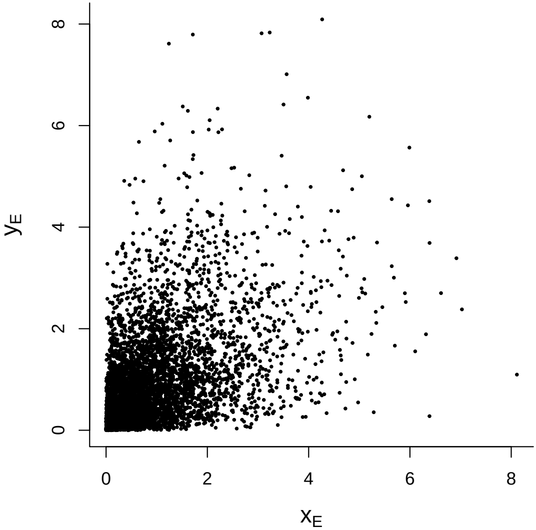
$$f_{\text{Gauss.}}(\boldsymbol{z}) = \frac{1}{\sqrt{2\pi\,|\boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{z}\right\}$$

$$\begin{aligned} g(\boldsymbol{x}) \quad &= \lim_{t\to\infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t} \\ &= \left(\boldsymbol{x}^{1/2}\right)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{x}^{1/2} \end{aligned}$$

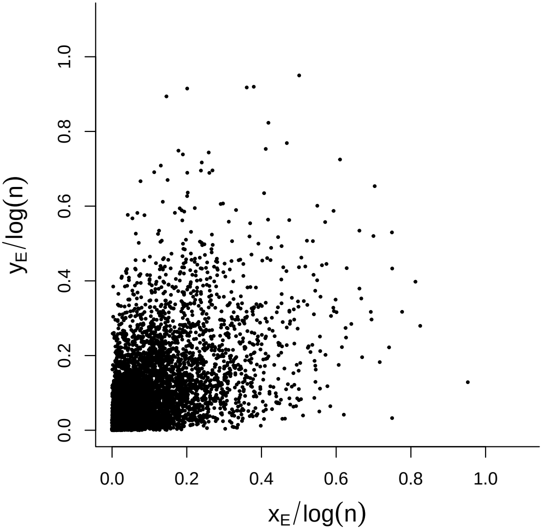
$d = 2$ Gaussian, $\rho = 0.5$



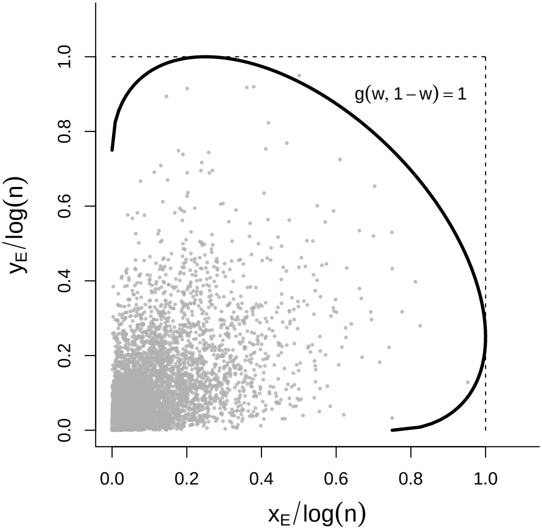
$d = 2$ Gaussian, $\rho = 0.5$



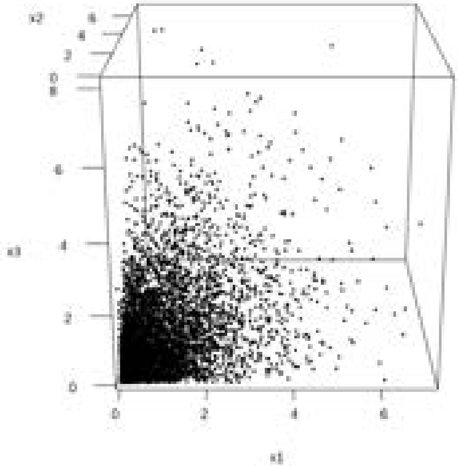
$d = 2$ Gaussian, $\rho = 0.5$



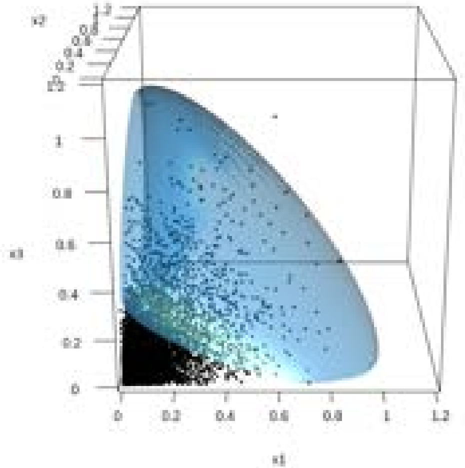
$d = 2$ Gaussian, $\rho = 0.5$



$d = 3$ Gaussian, $\rho_{12} = 0.5$, $\rho_{13} = 0.2$, $\rho_{23} = 0.8$



$d = 3$ Gaussian, $\rho_{12} = 0.5$, $\rho_{13} = 0.2$, $\rho_{23} = 0.8$

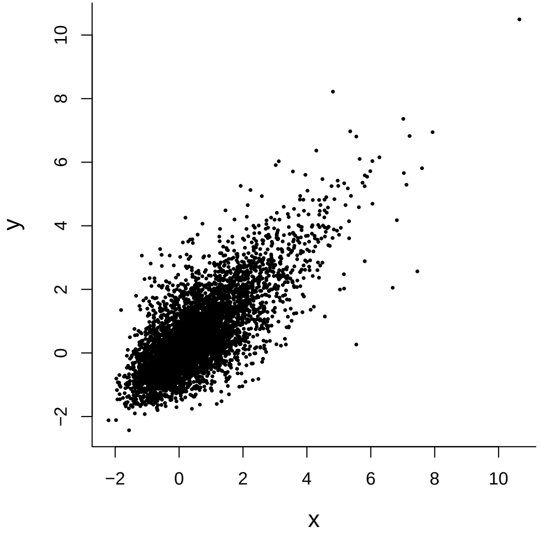


Example: Logistic

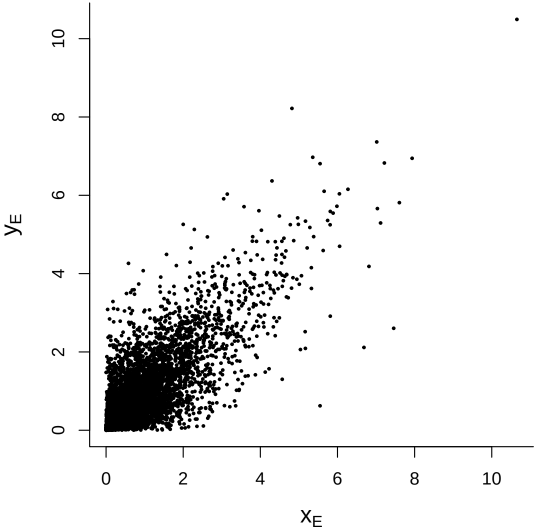
$$f_{\text{Fréchet}}(\boldsymbol{z}) = \exp\left\{ -\left(\sum_{j=1}^{d} z_j^{-1/\theta} \right)^{\theta} \right\}$$

$$\begin{aligned} g(\boldsymbol{x}) &= \lim_{t\to\infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t} \\ &= \frac{1}{\theta}\sum_{j=1}^{d} x_j + \left(1 - \frac{d}{\theta}\right) \min\{x_1, \ldots, x_d\} \end{aligned}$$

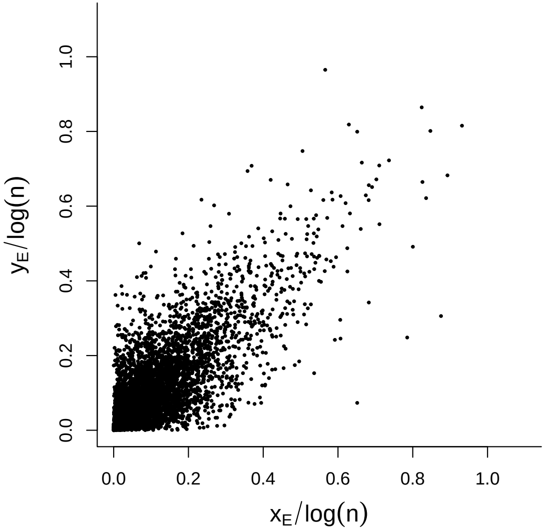
$d = 2$ Logistic, $\theta = 0.5$



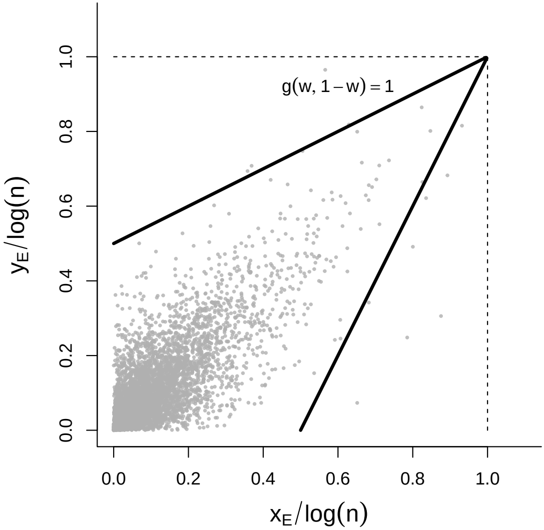
$d = 2$ Logistic, $\theta = 0.5$



$d = 2$ Logistic, $\theta = 0.5$



$d = 2$ Logistic, $\theta = 0.5$



Example: Asymmetric logistic

Let $\mathcal{D} = \{1, \ldots, d\}$ and $P(\mathcal{D})$ = power set of $\mathcal{D}$.

$$f_{\text{Gumbel}}(\boldsymbol{z}) = \exp\left\{ - \sum_{c \in P(\mathcal{D})} \gamma_c \left(\sum_{j=1}^{d} z_j^{-1/\theta_c} \right)^{\theta_c} \right\}$$

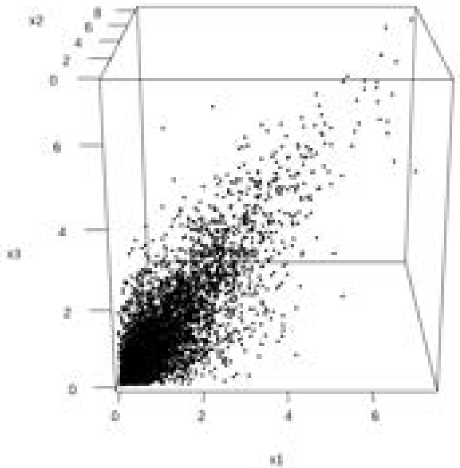
$$\begin{aligned} g(\boldsymbol{x}) \quad &= \lim_{t \to \infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t} \\ &= \text{very complicated!} \end{aligned}$$

Example: Asymmetric logistic

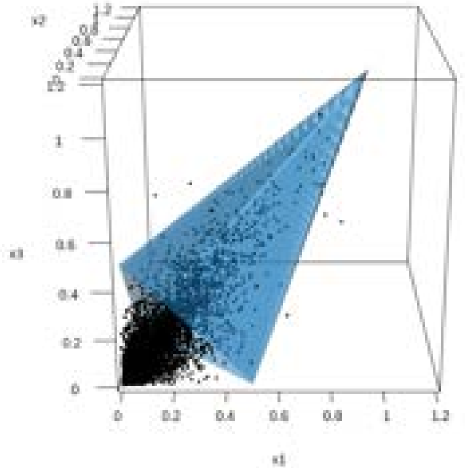
- **Example:** $\gamma_{\{1,\dots,d\}} = 1$, $\gamma_c = 0$ otherwise.

$$g(\boldsymbol{x}) = \frac{1}{\theta_{\{1,\dots,d\}}} \sum_{j=1}^{d} x_j + \left(1 - \frac{d}{\theta_{\{1,\dots,d\}}}\right) \min\{x_1, \dots, x_d\}$$

$d = 3$ Asymmetric logistic, variables $\{1, 2, 3\}$ large



$d = 3$ Asymmetric logistic, variables $\{1, 2, 3\}$ large

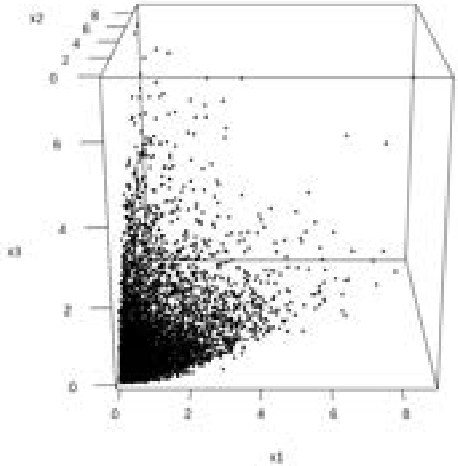


Example: Asymmetric logistic

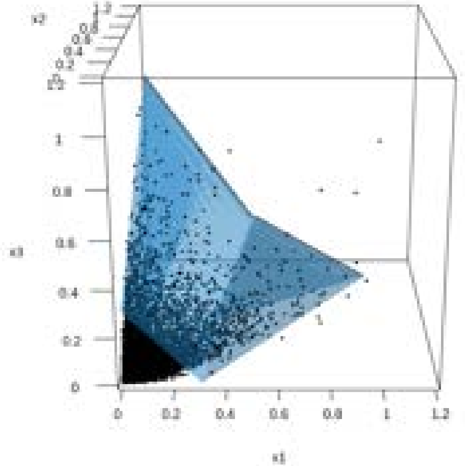
- **Example:** $\gamma_{\{1,2\}} = 1$, $\gamma_{\{2,3\}} = 1$, $\gamma_c = 0$ otherwise.

$$g(\boldsymbol{x}) = \min\left\{ \frac{x_1 + x_2}{\theta_{\{1,2\}}} + \frac{x_3}{\theta_{\{2,3\}}} + \left(1 - \frac{2}{\theta_{\{1,2\}}}\right)\min(x_1, x_2) + \left(1 - \frac{1}{\theta_{\{2,3\}}}\right)\min(x_2, x_3), \right.$$
$$\left. \frac{x_2 + x_3}{\theta_{\{2,3\}}} + \frac{x_1}{\theta_{\{1,2\}}} + \left(1 - \frac{2}{\theta_{\{2,3\}}}\right)\min(x_2, x_3) + \left(1 - \frac{1}{\theta_{\{1,2\}}}\right)\min(x_1, x_2) \right\}$$

$d = 3$ Asymmetric logistic, variables $\{1, 2\}$ and $\{2, 3\}$ large



$d = 3$ Asymmetric logistic, variables $\{1, 2\}$ and $\{2, 3\}$ large



How can we estimate g from data and use it for extremal statistical inference?

Extremal inference with gauge functions

Consider

$$\begin{aligned} R &= \|\boldsymbol{X}_E\|_1 \\ &= \sum_{j=1}^{d} \boldsymbol{X}_{E,j} \end{aligned}$$

Extremal inference with gauge functions

Consider

$$
\begin{aligned}
R &= \|\boldsymbol{X}_E\|_1 \\
&= \sum_{j=1}^{d} \boldsymbol{X}_{E,j}
\end{aligned}
$$

$$
\boldsymbol{W} = \boldsymbol{X}_E / \|\boldsymbol{X}_E\|_1
$$

Extremal inference with gauge functions

Consider

$$
\begin{aligned}
R &= \|\boldsymbol{X}_E\|_1 \\
&= \sum_{j=1}^{d} \boldsymbol{X}_{E,j}
\end{aligned}
$$

$$\boldsymbol{W} = \boldsymbol{X}_E / \|\boldsymbol{X}_E\|_1$$

$$r_q(\boldsymbol{W}) = q^{\text{th}} \text{ quantile of } R \text{ given angle } \boldsymbol{W}$$

Extremal inference with gauge functions

Consider

$$
\begin{aligned}
R &= \|\boldsymbol{X}_E\|_1 \\
&= \sum_{j=1}^{d} \boldsymbol{X}_{E,j}
\end{aligned}
$$

$$
\boldsymbol{W} = \boldsymbol{X}_E / \|\boldsymbol{X}_E\|_1
$$

$$
r_q(\boldsymbol{W}) = q^{\text{th}} \text{ quantile of } R \text{ given angle } \boldsymbol{W}
$$

$$
\text{i.e.,} \quad \Pr(R < r_q(\boldsymbol{w}) \mid \boldsymbol{W} = \boldsymbol{w}) = q
$$

Extremal inference with gauge functions

Consider

$$
\begin{aligned}
R &= \|\boldsymbol{X}_E\|_1 \\
&= \sum_{j=1}^{d} \boldsymbol{X}_{E,j}
\end{aligned}
$$

$$\boldsymbol{W} = \boldsymbol{X}_E / \|\boldsymbol{X}_E\|_1$$

$$r_q(\boldsymbol{W}) = q^{\text{th}} \text{ quantile of } R \text{ given angle } \boldsymbol{W}$$

$$\text{i.e.,} \quad \Pr(R < r_q(\boldsymbol{w}) \mid \boldsymbol{W} = \boldsymbol{w}) = q$$

- Note that $\boldsymbol{X}_E = R\boldsymbol{W}$

Extremal inference with gauge functions

$$R \in \mathbb{R}_+$$

$$\boldsymbol{W} \in \mathcal{S}_{d-1} = \left\{ \boldsymbol{x} \in \mathbb{R}^d : \sum_{j=1}^{d} |x_j| = 1 \right\}$$

Extremal inference with gauge functions

$$\lim_{t\to\infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t} = g(\boldsymbol{x})$$

Extremal inference with gauge functions

$$\lim_{t\to\infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t} = g(\boldsymbol{x})$$

$$\begin{aligned} -\log f_{\boldsymbol{X}_E}(t\boldsymbol{x}) &\sim tg(\boldsymbol{x}) \qquad ; \qquad t \to \infty \\ &= tg(\boldsymbol{x})\,[1 + o(1)] \end{aligned}$$

Extremal inference with gauge functions

$$\lim_{t\to\infty} \frac{-\log f_{\boldsymbol{X}_E}(t\boldsymbol{x})}{t} = g(\boldsymbol{x})$$

$$\begin{aligned} -\log f_{\boldsymbol{X}_E}(t\boldsymbol{x}) &\sim tg(\boldsymbol{x}) \quad ; \quad t \to \infty \\ &= tg(\boldsymbol{x})\,[1 + o(1)] \end{aligned}$$

$$f_{\boldsymbol{X}_E}(t\boldsymbol{x}) = e^{-tg(\boldsymbol{x})[1+o(1)]}$$

Extremal inference with gauge functions

$$f_{\boldsymbol{X}_E}(r\boldsymbol{w}) = e^{-rg(\boldsymbol{w})[1+o(1)]}$$

Extremal inference with gauge functions

$$f_{\boldsymbol{X}_E}(r\boldsymbol{w}) = e^{-rg(\boldsymbol{w})[1+o(1)]}$$

$$\begin{aligned} f_{R,\boldsymbol{W}}(r,\boldsymbol{w}) &= |\nabla r\boldsymbol{w}|\, f_{\boldsymbol{X}_E}(r\boldsymbol{w}) \\ &= r^{d-1} e^{-rg(\boldsymbol{w})[1+o(1)]} \end{aligned}$$

Extremal inference with gauge functions

$$
\begin{aligned}
f_{R|\boldsymbol{W}}(r \mid \boldsymbol{w}) &= f_{R,\boldsymbol{W}}(r, \boldsymbol{w}) / f_{\boldsymbol{W}}(\boldsymbol{w}) \\
&\propto r^{d-1} e^{-r g(\boldsymbol{w})[1+o(1)]}
\end{aligned}
$$

Extremal inference with gauge functions

$$
\begin{aligned}
f_{R|\boldsymbol{W}}(r \mid \boldsymbol{w}) &= f_{R,\boldsymbol{W}}(r, \boldsymbol{w}) / f_{\boldsymbol{W}}(\boldsymbol{w}) \\
&\propto r^{d-1} e^{-r g(\boldsymbol{w})[1+o(1)]}
\end{aligned}
$$

$$
R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\} \sim \text{truncGamma}(d, g(\boldsymbol{w}))
$$

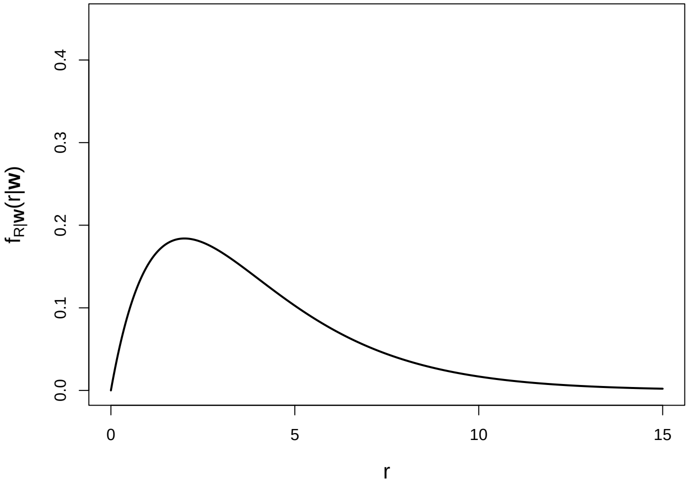
Extremal inference with gauge functions

$$f_{R|\boldsymbol{W}}(r \mid \boldsymbol{w}) = f_{R,\boldsymbol{W}}(r, \boldsymbol{w}) / f_{\boldsymbol{W}}(\boldsymbol{w})$$
$$\propto r^{d-1} e^{-rg(\boldsymbol{w})[1+o(1)]}$$

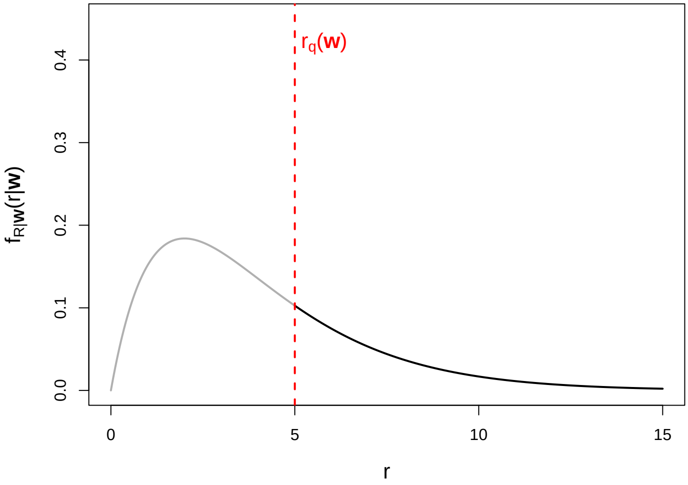
$$R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\} \sim \text{truncGamma}(d, g(\boldsymbol{w}))$$

$$f_{\text{truncGamma}}(r \mid \boldsymbol{w}) = \begin{cases} \frac{f_{\text{Gamma}}(r;\alpha, g(\boldsymbol{w};\boldsymbol{\theta}))}{\bar{F}_{\text{Gamma}}(r_q(\boldsymbol{w});\alpha, g(\boldsymbol{w};\boldsymbol{\theta}))} & ;\ r > r_q(\boldsymbol{w}) \\ 0 & ;\ r \leq r_q(\boldsymbol{w}) \end{cases}$$

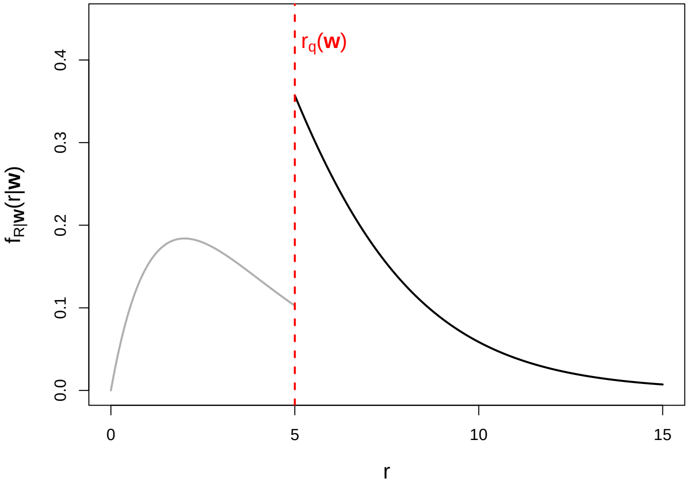
Extremal inference with gauge functions



Extremal inference with gauge functions



Extremal inference with gauge functions

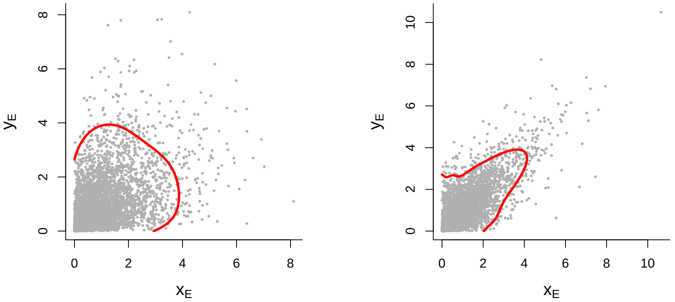


Extremal inference with gauge functions

- Interested in when $R \mid \{\boldsymbol{W} = \boldsymbol{w}\}$ is large, or $R > r_q(\boldsymbol{W})$

Extremal inference with gauge functions

- Interested in when $R \mid \{\boldsymbol{W} = \boldsymbol{w}\}$ is large, or $R > r_q(\boldsymbol{W})$



Parametric approach: model fitting

- Fit the model

$$R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\} \sim \mathrm{truncGamma}(\alpha, g(\boldsymbol{w}; \boldsymbol{\theta}))$$

by maximizing

$$L(\alpha, \boldsymbol{\theta} \mid r_{1:n}, \boldsymbol{w}_{1:n}) = \prod_{i: r_i > r_q(\boldsymbol{w}_i)} \frac{f_{\mathrm{Gamma}}\left(r_i; \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}{\bar{F}_{\mathrm{Gamma}}\left(r_q(\boldsymbol{w}_i); \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}$$

for different parametric choices of g.

Parametric approach: model fitting

- Fit the model

$$R \mid \{\boldsymbol{W} = \boldsymbol{w}, R > r_q(\boldsymbol{W})\} \sim \text{truncGamma}(\alpha, g(\boldsymbol{w}; \boldsymbol{\theta}))$$

by maximizing

$$L(\alpha, \boldsymbol{\theta} \mid r_{1:n}, \boldsymbol{w}_{1:n}) = \prod_{i: r_i > r_q(\boldsymbol{w}_i)} \frac{f_{\text{Gamma}}\left(r_i; \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}{\bar{F}_{\text{Gamma}}\left(r_q(\boldsymbol{w}_i); \alpha, g(\boldsymbol{w}_i; \boldsymbol{\theta})\right)}$$

for different parametric choices of g.

- Select best model using AIC.

Parametric approach: extrapolation

- Can estimate extremal probabilities via

$$\Pr(\boldsymbol{X} \in B) = \Pr(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})) \Pr(R > r_q(\boldsymbol{W}))$$

Parametric approach: extrapolation

- Can estimate extremal probabilities via

$$\Pr(\boldsymbol{X} \in B) = \Pr(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})) \Pr(R > r_q(\boldsymbol{W}))$$

- Compute $\Pr(R > r_q(\boldsymbol{W}))$ empirically.

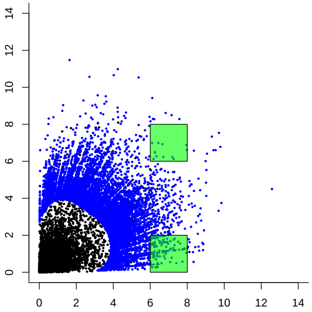
Parametric approach: extrapolation

- Can estimate extremal probabilities via

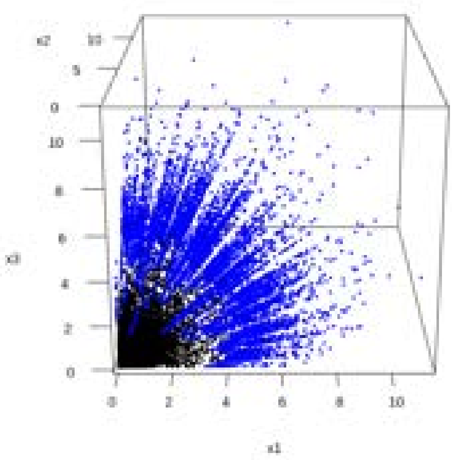
$$\Pr(\boldsymbol{X} \in B) = \Pr(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W})) \Pr(R > r_q(\boldsymbol{W}))$$

- Compute $\Pr(R > r_q(\boldsymbol{W}))$ empirically.
- Compute $\Pr(\boldsymbol{X} \in B | R > r_q(\boldsymbol{W}))$ via sampling:
  1. generate $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$
  2. generate r_i from truncGamma$(\widehat{\alpha}, g(\boldsymbol{w}_i; \widehat{\boldsymbol{\theta}}))$ for $i = 1, \ldots, N$
  3. return $\boldsymbol{x}_i = r_i \boldsymbol{w}_i$, $i = 1, \ldots, N$

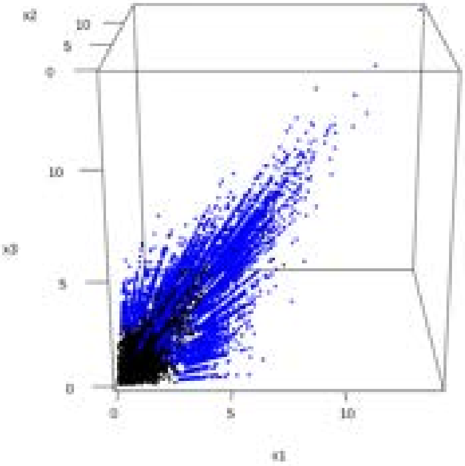
Parametric approach: extrapolation



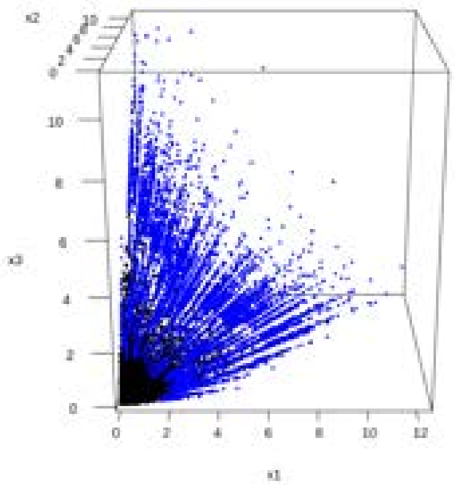
Parametric approach: extrapolation



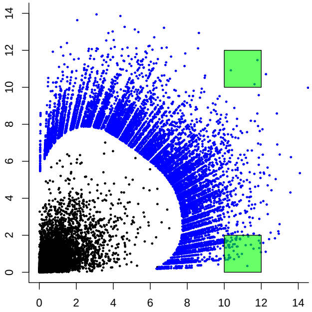
Parametric approach: extrapolation

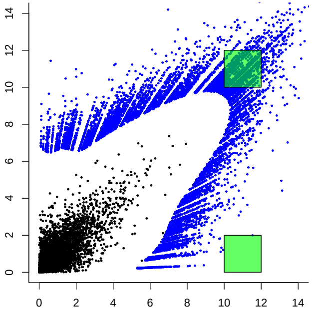


Parametric approach: extrapolation

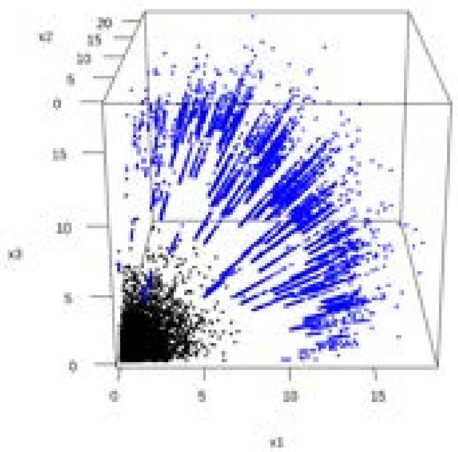


Parametric approach: further extrapolation

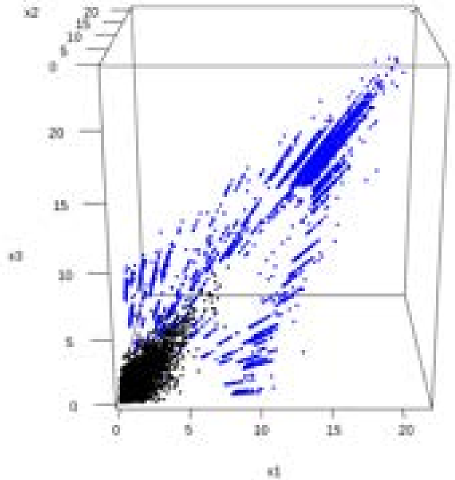




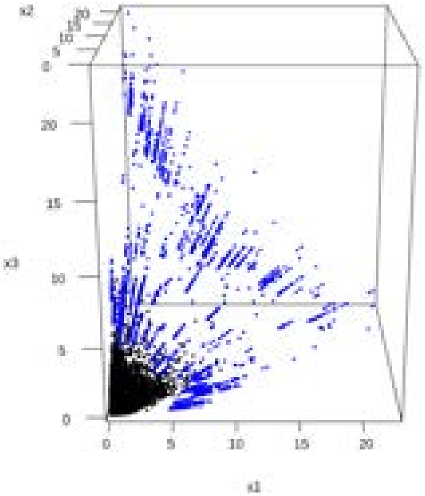
Parametric approach: further extrapolation



Parametric approach: further extrapolation



Parametric approach: further extrapolation



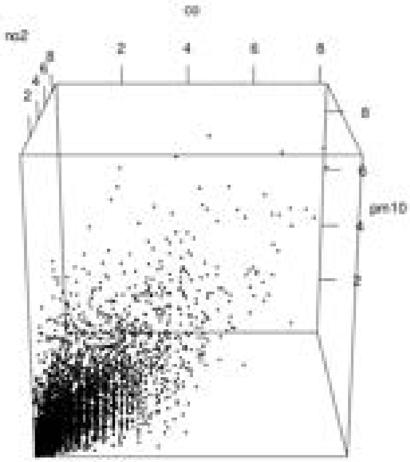
Application to air pollution measurements ($d = 3$)

- North Kensington site, London, UK
- carbon monoxide (CO, mg$/m^3$), nitrogen dioxide (NO_2, μg$/m^3$), and particles with a diameter of 10 μm or less (PM10, mg$/m^3$).
- $n = 5,584$ daily maximum measurements, October–April. 1996–2024.

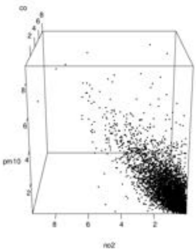
Application to air pollution measurements ($d = 3$)

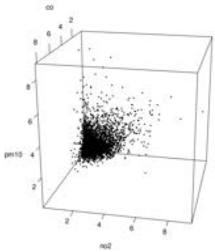
- North Kensington site, London, UK
- carbon monoxide (CO, mg/m^3), nitrogen dioxide (NO_2, μg/m^3), and particles with a diameter of 10 μm or less (PM10, mg/m^3).
- $n = 5,584$ daily maximum measurements, October–April. 1996–2024.
- Evidence (Simpson et al., 2020) that PM10 is large when CO and NO_2 are small, and that all three grow large together.

Application to air pollution measurements ($d = 3$)

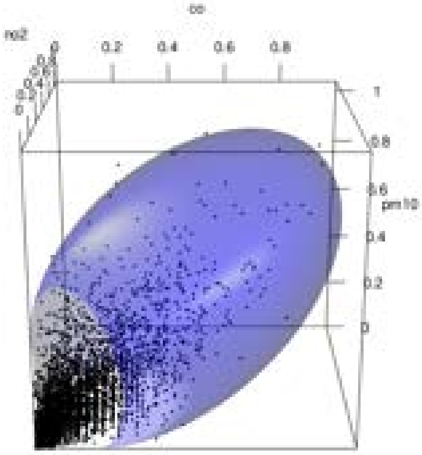


Application to air pollution measurements ($d = 3$)

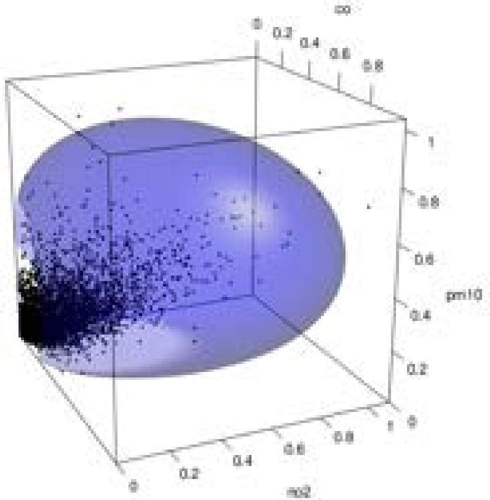




Application to air pollution measurements ($d = 3$)



Application to air pollution measurements ($d = 3$)



Application to air pollution measurements ($d = 3$)

A commonly-used measure of extremal dependence is

$$\chi_c(u) = \left(\frac{1}{1-u}\right) \Pr\left[F_E\left(X_j\right) > u \, \forall \, j \in c \subseteq \{1,2,3\}\right]$$

for u close to 1.

Application to air pollution measurements ($d = 3$)

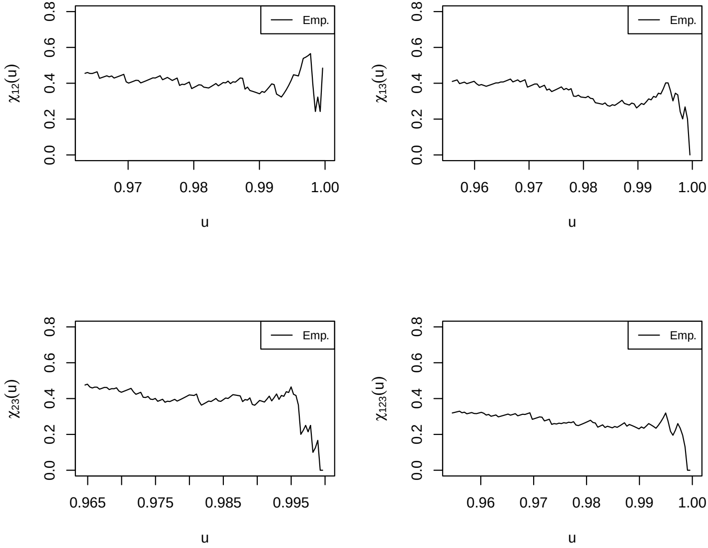
A commonly-used measure of extremal dependence is

$$\chi_c(u) = \left(\frac{1}{1-u}\right) \Pr\left[F_E\left(X_j\right) > u \,\forall\, j \in c \subseteq \{1, 2, 3\}\right]$$

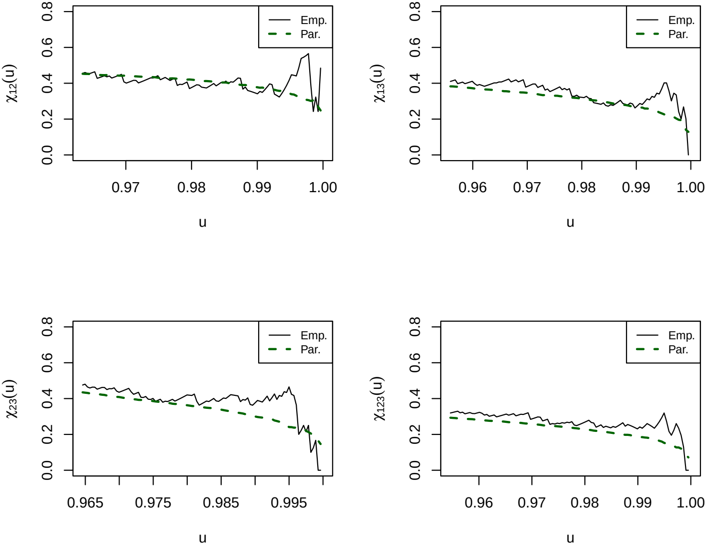
for u close to 1.

- $\chi_{12}(u) = \Pr\left(F_E(X_2) > u | F_E(X_1) > u\right)$
- $\chi_{13}(u) = \Pr\left(F_E(X_3) > u | F_E(X_1) > u\right)$
- $\chi_{23}(u) = \Pr\left(F_E(X_3) > u | F_E(X_2) > u\right)$
- $\chi_{123}(u) = \Pr\left(F_E(X_2) > u, F_E(X_3) > u | F_E(X_1) > u\right)$

Application to air pollution measurements ($d = 3$)



Application to air pollution measurements ($d = 3$)



Parametric approach

- A "rough sketch" on how to perform geometric multivariate extremal inference.
- Statistical inference for multivariate tails **for any dependence structure**.
- Comparable (and sometimes outperforms) conditional extremes.
- Simple inference.

Parametric approach

- Parametric models are too rigid for real-data examples, and mixing gauges is slow.
- Parameter estimation and sampling is slow when $d \geq 3$.
- Re-sampling exceedance angles with replacement is undesirable when $d > 3$, would like a model for $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$

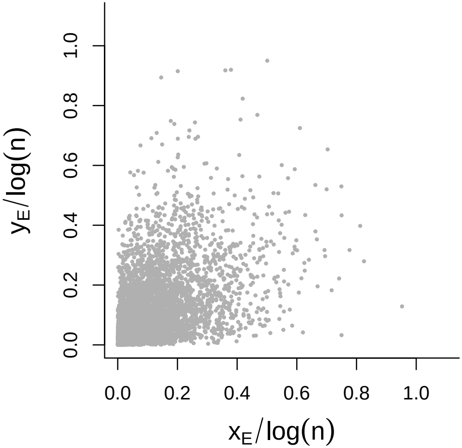
Parametric approach

- Parametric models are too rigid for real-data examples, and mixing gauges is slow.
- Parameter estimation and sampling is slow when $d \geq 3$.
- Re-sampling exceedance angles with replacement is undesirable when $d > 3$, would like a model for $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$
- Current quantile regression approaches are only suitable for $d = 2, 3$.

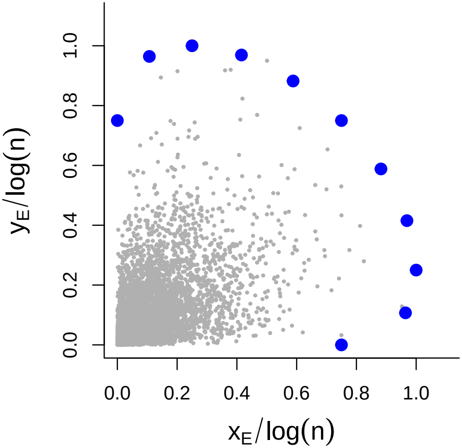
Parametric approach

- Parametric models are too rigid for real-data examples, and mixing gauges is slow.
- Parameter estimation and sampling is slow when $d \geq 3$.
- Re-sampling exceedance angles with replacement is undesirable when $d > 3$, would like a model for $\boldsymbol{W} \mid \{R > r_q(\boldsymbol{W})\}$
- Current quantile regression approaches are only suitable for $d = 2, 3$.
- **Ultimate goal is for a fast and accurate model when $d = 4, 5, 6, \ldots$**

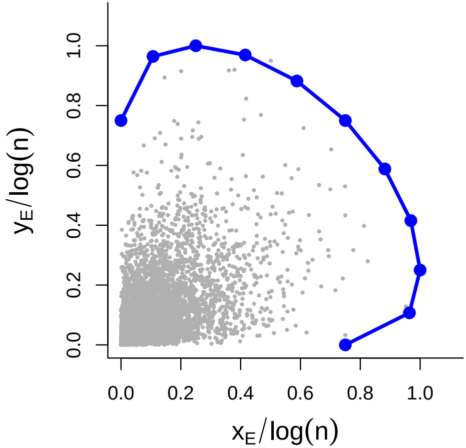
Semiparametric piecewise-linear approach



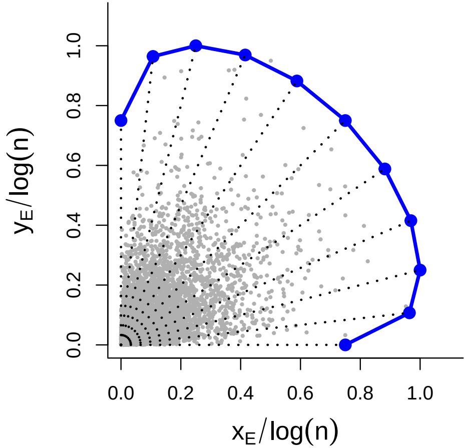
Semiparametric piecewise-linear approach



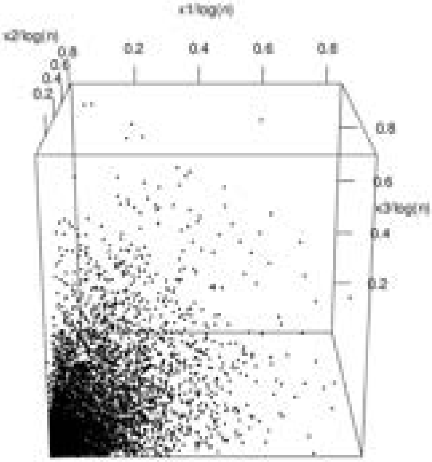
Semiparametric piecewise-linear approach



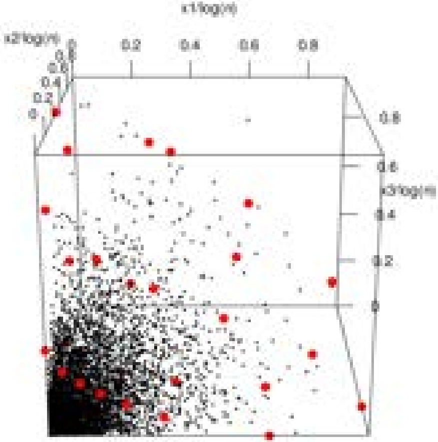
Semiparametric piecewise-linear approach



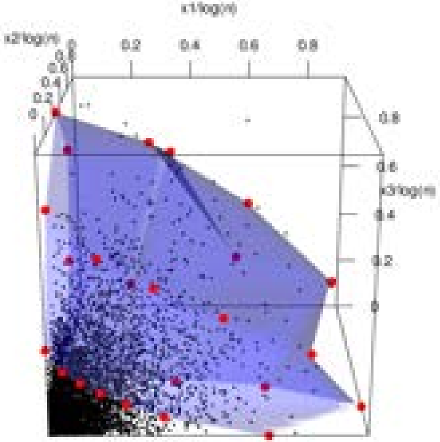
Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach

For $d = 2$...

Semiparametric piecewise-linear approach

For $d = 2$...

- Define a set of N reference angles $w^{\star 1}, \ldots, w^{\star N} \in [0, 1]$.

Semiparametric piecewise-linear approach

For $d = 2$...

- Define a set of N reference angles $w^{\star 1}, \ldots, w^{\star N} \in [0, 1]$.
- Results in $N - 1$ regions: $[w^{\star 1}, w^{\star 2}], \ldots, [w^{\star N-1}, w^{\star N}]$.

Semiparametric piecewise-linear approach

For $d = 2$...

- Define a set of N reference angles $w^{\star 1}, \ldots, w^{\star N} \in [0, 1]$.
- Results in $N - 1$ regions: $[w^{\star 1}, w^{\star 2}], \ldots, [w^{\star N-1}, w^{\star N}]$.
- At each $w^{\star k}$, define a parameter $\theta_k > 0$ such that $\theta_i (w^{\star k}, 1 - w^{\star k})^\top$ lies on the boundary of the limit set.

Semiparametric piecewise-linear approach

For $d = 2$...

- Define a set of N reference angles $w^{\star 1}, \ldots, w^{\star N} \in [0, 1]$.
- Results in $N - 1$ regions: $[w^{\star 1}, w^{\star 2}], \ldots, [w^{\star N-1}, w^{\star N}]$.
- At each $w^{\star k}$, define a parameter $\theta_k > 0$ such that $\theta_i (w^{\star k}, 1 - w^{\star k})^\top$ lies on the boundary of the limit set.
- Linear interpolation between neighbouring points points.

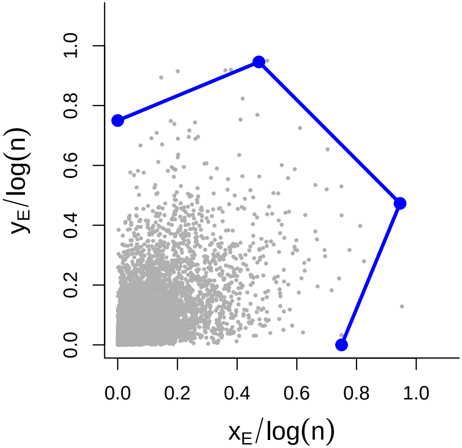
Semiparametric piecewise-linear approach

For $d = 2$...

- Define a set of N reference angles $w^{\star 1}, \dots, w^{\star N} \in [0, 1]$.
- Results in $N-1$ regions: $[w^{\star 1}, w^{\star 2}], \dots, [w^{\star N-1}, w^{\star N}]$.
- At each $w^{\star k}$, define a parameter $\theta_k > 0$ such that $\theta_i(w^{\star k}, 1 - w^{\star k})^\top$ lies on the boundary of the limit set.
- Linear interpolation between neighbouring points points.
- At a point $(x_1, x_2)^\top$, the gauge function value is given by

$$
\begin{aligned}
g(x_1, x_2; \boldsymbol{\theta}) = &\sum_{k=1}^{N-1} \mathbf{1}_{(w^{\star k}, w^{\star k+1})}\left(\frac{x_1}{x_1 + x_2}\right) \\
&\times \frac{\left[\theta_k(1 - w^{\star k}) - \theta_{k+1}(1 - w^{\star k+1})\right] x_1 - \left[\theta_k w^{\star k} - \theta_{k+1} w^{\star k+1}\right] x_2}{\left[\theta_k(1 - w^{\star k}) - \theta_{k+1}(1 - w^{\star k+1})\right] \theta_k w^{\star k} - \left[\theta_k w^{\star k} - \theta_{k+1} w^{\star k+1}\right] \theta_k (1 - w^{\star k})}
\end{aligned}
$$

Semiparametric piecewise-linear approach



Semiparametric piecewise-linear approach

In d-dimensions...

Semiparametric piecewise-linear approach

In d-dimensions...

- Define a set of N reference angles $\boldsymbol{w}^{\star 1}, \ldots, \boldsymbol{w}^{\star N} \in \mathcal{S}_{d-1}$.

Semiparametric piecewise-linear approach

In d-dimensions...

- Define a set of N reference angles $\boldsymbol{w}^{\star 1}, \ldots, \boldsymbol{w}^{\star N} \in \mathcal{S}_{d-1}$.
- Results in M regions: $\triangle^{(1)}, \triangle^{(2)}, \ldots, \triangle^{(M)}$.

Semiparametric piecewise-linear approach

In d-dimensions...

- Define a set of N reference angles $\boldsymbol{w}^{\star 1}, \ldots, \boldsymbol{w}^{\star N} \in \mathcal{S}_{d-1}$.
- Results in M regions: $\triangle^{(1)}, \triangle^{(2)}, \ldots, \triangle^{(M)}$.
- Region $\triangle^{(k)}$ has d vertices: $\theta_1^{(k)} \boldsymbol{w}^{\star (k),1}, \ldots, \theta_d^{(k)} \boldsymbol{w}^{\star (k),d}$

Semiparametric piecewise-linear approach

In d-dimensions...

- Define a set of N reference angles $\boldsymbol{w}^{\star 1}, \ldots, \boldsymbol{w}^{\star N} \in \mathcal{S}_{d-1}$.
- Results in M regions: $\triangle^{(1)}, \triangle^{(2)}, \ldots, \triangle^{(M)}$.
- Region $\triangle^{(k)}$ has d vertices: $\theta_1^{(k)} \boldsymbol{w}^{\star(k),1}, \ldots, \theta_d^{(k)} \boldsymbol{w}^{\star(k),d}$
- At a point $\boldsymbol{x}$, the gauge function value is given by

$$g(\boldsymbol{x}; \boldsymbol{\theta}) = \sum_{k=1}^{M} \mathbf{1}_{\triangle^{(k)}}\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right) \frac{\boldsymbol{n}^{(k)\top} \boldsymbol{x}}{\boldsymbol{n}^{(k)\top} \theta_1^{(k)} \boldsymbol{w}^{\star(k),1}}$$

Semiparametric piecewise-linear approach

- **Problem:** N large leads to variability in MLEs $\widehat{\boldsymbol{\theta}}$.

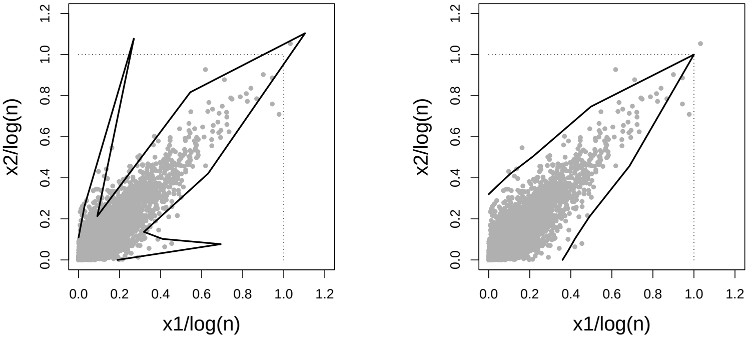
Semiparametric piecewise-linear approach

- **Problem:** N large leads to variability in MLEs $\widehat{\boldsymbol{\theta}}$.
- **Solution:** Penalise the gradients during model fitting:

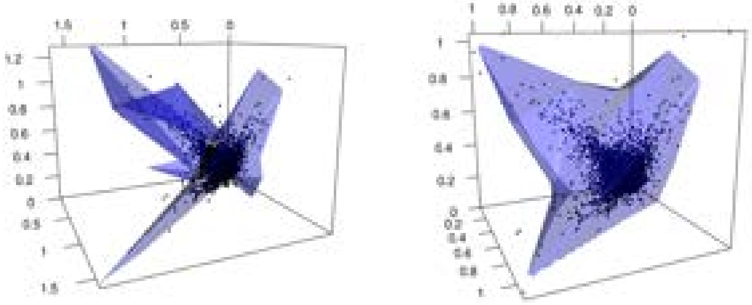
$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \mathbb{R}_+^d}{\operatorname{argmin}} - \log L(\boldsymbol{\theta} \mid r_{1:n}, \boldsymbol{w}_{1:m}) + \lambda \sum_{(i,j) \in \mathcal{I}} \left\| \nabla g_{\boldsymbol{\theta}}^{(i)} - \nabla g_{\boldsymbol{\theta}}^{(j)} \right\|_2^2$$

$\lambda \geq 0$.

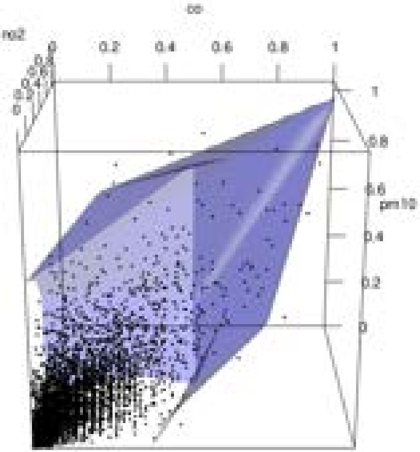
Semiparametric piecewise-linear approach



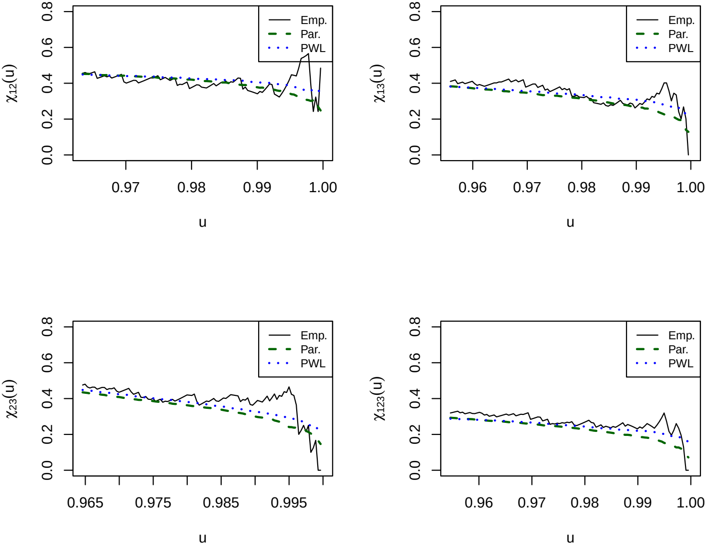
Semiparametric piecewise-linear approach



Air pollution revisited



Air pollution revisited



In conclusion...

- Lack of a unified approach for multivariate extremal inference.
- Geometric approach tackles problems with difficult dependence structures.
- Wadsworth and Campbell (2024) is first to use the geometric approach for *statistical inference*.
- The parametric approach works well in $d = 2, 3$, is okay when $d = 4$, hasn't been tested for $d \geq 5$.
- Campbell and Wadsworth (???) can scale up to $d = 5, 6, 7, \ldots$
- We also have improved modelling the angles and quantile estimation!

Bibliography

A. A. Balkema and N. Nolde. Asymptotic independence for unimodal densities. *Advances in Applied Probability*, 42(2):411–432, 2010. ISSN 00018678. URL http://www.jstor.org/stable/25683827.

L. de Haan. *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes, by L. de Haan*. Mathematical Centre tracts. 1970. URL https://books.google.ca/books?id=cJ3uswEACAAJ.

J. E. Heffernan and J. A. Tawn. A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):497–546, 2004.

A. W. Ledford and J. A. Tawn. Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59(2): 475–499, 1997.

N. Nolde. Geometric interpretation of the residual dependence coefficient. *Journal of Multivariate Analysis*, 123:85–95, 2014. ISSN 0047-259X.

N. Nolde and J. L. Wadsworth. Linking representations for multivariate extremes via a limit set. *Advances in Applied Probability*, 54(3):688–717, 2022. doi: 10.1017/apr.2021.51.

E. S. Simpson, J. L. Wadsworth, and J. A. Tawn. Determining the dependence structure of multivariate extremes. *Biometrika*, 107(3):513–532, 2020.

J. L. Wadsworth and R. Campbell. Statistical inference for multivariate extremes via a geometric approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, page qkae030, 03 2024. ISSN 1369-7412. doi: 10.1093/jrsssb/qkae030. URL https://doi.org/10.1093/jrsssb/qkae030.