Dependence Estimation in Spatial Extremes

Michaël Lalancette[1]
Joint work with Sebastian Engelke[2] and Stanislav Volgushev[1]

[1]Department of Statistical Sciences, University of Toronto

[2]Research Center for Statistics, University of Geneva

April 23, 2021

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Mathematically, EVT estimates the tail of the d.f. F of a random variable

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Mathematically, EVT estimates the tail of the d.f. F of a random variable
- Univariate methods fairly well developed

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Mathematically, EVT estimates the tail of the d.f. F of a random variable
- Univariate methods fairly well developed
- Multivariate and spatial methods: dependence modeling is challenging

Extreme value theory

- Extrapolate outside the range of data and estimate properties of the underlying distribution *beyond the observed regions*
- Mathematically, EVT estimates the tail of the d.f. F of a random variable
- Univariate methods fairly well developed
- Multivariate and spatial methods: dependence modeling is challenging
- Paper: Bivariate $\Rightarrow$ Spatial

Motivating data

- $n \approx 18\,000$ rainfall measurements at $d = 40$ locations in southern Australia (Le et al., 2018)

Motivating data

- $n \approx 18\,000$ rainfall measurements at $d = 40$ locations in southern Australia (Le et al., 2018)
- Typical problem: estimate

$$\mathbb{P}(\text{rainfall at location } j > x_j, \text{for each } j \in J),$$

for large values x_j and finite collection of locations J

Motivating data

- $n \approx 18\,000$ rainfall measurements at $d = 40$ locations in southern Australia (Le et al., 2018)
- Typical problem: estimate

$$\mathbb{P}(\text{rainfall at location } j > x_j, \text{for each } j \in J),$$

  for large values x_j and finite collection of locations J
- Both the locations J and the levels x_j could be unobserved

Motivating data

- $n \approx 18\,000$ rainfall measurements at $d = 40$ locations in southern Australia (Le et al., 2018)
- Typical problem: estimate

$$\mathbb{P}(\text{rainfall at location } j > x_j, \text{for each } j \in J),$$

  for large values x_j and finite collection of locations J
- Both the locations J and the levels x_j could be unobserved
- Marginal tails are estimated by univariate EVT methods, so we focus on the dependence structure

Spatial dependence model

- Inverted Brown–Resnick process on $\mathbb{R}^2$ (or a subset)

Spatial dependence model

- Inverted Brown–Resnick process on $\mathbb{R}^2$ (or a subset)
- Constructed from normalized maxima of iid Gaussian processes over a "shrinking" region (Brown and Resnick, 1977)

Spatial dependence model

- Inverted Brown–Resnick process on $\mathbb{R}^2$ (or a subset)
- Constructed from normalized maxima of iid Gaussian processes over a "shrinking" region (Brown and Resnick, 1977)
- Parametrized by a *variogram function* $\gamma : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$

Spatial dependence model

- Inverted Brown–Resnick process on $\mathbb{R}^2$ (or a subset)
- Constructed from normalized maxima of iid Gaussian processes over a "shrinking" region (Brown and Resnick, 1977)
- Parametrized by a *variogram function* $\gamma : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$
- If the marginal distributions are correctly transformed, variogram of $Z := \{Z(u) : u \in \mathbb{R}^2\}$ is

$$\gamma(u, u') = \mathbb{V}\mathrm{ar}(Z(u) - Z(u'))$$

Spatial dependence model

- Inverted Brown–Resnick process on $\mathbb{R}^2$ (or a subset)
- Constructed from normalized maxima of iid Gaussian processes over a "shrinking" region (Brown and Resnick, 1977)
- Parametrized by a *variogram function* $\gamma : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$
- If the marginal distributions are correctly transformed, variogram of $Z := \{Z(u) : u \in \mathbb{R}^2\}$ is

$$\gamma(u, u') = \mathbb{V}\mathrm{ar}(Z(u) - Z(u'))$$

- Most popular model: *fractal variogram*

$$\gamma_\vartheta(u, u') = (\|u - u'\|/\beta)^\alpha,$$

$\vartheta := (\alpha, \beta) \in (0, 2) \times (0, \infty)$

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$t^{-2\theta^{(s)}} \mathbb{P}(F_u(Z(u)) \geq 1 - tx, F_{u'}(Z(u')) \geq 1 - ty) \longrightarrow (xy)^{\theta^{(s)}},$$

as $t \downarrow 0$

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$t^{-2\theta^{(s)}}\mathbb{P}(F_u(Z(u)) \geq 1 - tx, F_{u'}(Z(u')) \geq 1 - ty) \longrightarrow (xy)^{\theta^{(s)}},$$

  as $t \downarrow 0$

- $\theta^{(s)} := \Phi(\sqrt{\gamma(u, u')}/2) \in (1/2, 1]$

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$t^{-2\theta^{(s)}} \mathbb{P}(F_u(Z(u)) \geq 1 - tx, F_{u'}(Z(u')) \geq 1 - ty) \longrightarrow (xy)^{\theta^{(s)}},$$

  as $t \downarrow 0$
- $\theta^{(s)} := \Phi(\sqrt{\gamma(u, u')}/2) \in (1/2, 1]$
- So $\{\text{all the } \theta^{(s)}\} \Rightarrow \gamma$

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$t^{-2\theta^{(s)}} \mathbb{P}(F_u(Z(u)) \geq 1 - tx, F_{u'}(Z(u')) \geq 1 - ty) \longrightarrow (xy)^{\theta^{(s)}},$$

  as $t \downarrow 0$
- $\theta^{(s)} := \Phi(\sqrt{\gamma(u, u')}/2) \in (1/2, 1]$
- So $\{$all the $\theta^{(s)}\} \Rightarrow \gamma$
- But knowing $\gamma \in \{\gamma_\vartheta\}$, $\{$a small number of $\theta^{(s)}\} \Rightarrow \vartheta$

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$t^{-2\theta^{(s)}}\mathbb{P}(F_u(Z(u)) \geq 1 - tx, F_{u'}(Z(u')) \geq 1 - ty) \longrightarrow (xy)^{\theta^{(s)}},$$

  as $t \downarrow 0$
- $\theta^{(s)} := \Phi(\sqrt{\gamma(u, u')}/2) \in (1/2, 1]$
- So $\{$all the $\theta^{(s)}\} \Rightarrow \gamma$
- But knowing $\gamma \in \{\gamma_\vartheta\}$, $\{$a small number of $\theta^{(s)}\} \Rightarrow \vartheta$
- So only need to estimate the pairwise parameters $\theta^{(s)}$

Identification via bivariate tails

- For an IBR process Z, and any pair of locations $s = (u, u') \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$t^{-2\theta^{(s)}} \mathbb{P}(F_u(Z(u)) \geq 1 - tx, F_{u'}(Z(u')) \geq 1 - ty) \longrightarrow (xy)^{\theta^{(s)}},$$

  as $t \downarrow 0$
- $\theta^{(s)} := \Phi(\sqrt{\gamma(u, u')}/2) \in (1/2, 1]$
- So $\{$all the $\theta^{(s)}\} \Rightarrow \gamma$
- But knowing $\gamma \in \{\gamma_\vartheta\}$, $\{$a small number of $\theta^{(s)}\} \Rightarrow \vartheta$
- So only need to estimate the pairwise parameters $\theta^{(s)}$
- Note: every pair is *asymptotically independent*

Bivariate approach

- Fix an observed pair $s = (u, u')$, say $(X, Y) = (Z(u), Z(u'))$

Bivariate approach

- Fix an observed pair $s = (u, u')$, say $(X, Y) = (Z(u), Z(u'))$
- Wish to estimate

Bivariate approach

- Fix an observed pair $s = (u, u')$, say $(X, Y) = (Z(u), Z(u'))$
- Wish to estimate

$$c(x, y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

Bivariate approach

- Fix an observed pair $s = (u, u')$, say $(X, Y) = (Z(u), Z(u'))$
- Wish to estimate

$$c(x, y) := \lim_{t \downarrow 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- The function c characterizes bivariate tail dependence much more general than the bivariate margins of IBR processes (asympt. dep. or indep.)

Estimation of c

- (iid) observations $(X_1, Y_1), \ldots, (X_n, Y_n)$

Estimation of c

- (iid) observations $(X_1, Y_1), \ldots, (X_n, Y_n)$
- Recall

$$c(x, y) := \lim_{t \to 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

Estimation of c

- (iid) observations $(X_1, Y_1), \ldots, (X_n, Y_n)$
- Recall

$$c(x, y) := \lim_{t \to 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- For a sequence $t_n \downarrow 0$,

Estimation of c

- (iid) observations $(X_1, Y_1), \dots, (X_n, Y_n)$
- Recall

$$c(x, y) := \lim_{t \to 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- For a sequence $t_n \downarrow 0$,

$$\widehat{c}_n(x, y) := q(t_n)^{-1} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left\{ \widehat{F}_1(X_i) \geq 1 - t_n x, \widehat{F}_2(Y_i) \geq 1 - t_n y \right\},$$

where $\widehat{F}_j$ are the empirical d.f.

Estimation of c

- (iid) observations $(X_1, Y_1), \ldots, (X_n, Y_n)$
- Recall

$$c(x,y) := \lim_{t \to 0} q(t)^{-1} \mathbb{P}(F_1(X) \geq 1 - tx, F_2(Y) \geq 1 - ty)$$

- For a sequence $t_n \downarrow 0$,

$$\widehat{c}_n(x,y) := q(t_n)^{-1} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\left\{ \widehat{F}_1(X_i) \geq 1 - t_n x, \widehat{F}_2(Y_i) \geq 1 - t_n y \right\},$$

where $\widehat{F}_j$ are the empirical d.f.

- Issue: q is unknown

Estimation of c

- We know $c(x,y) = c_\theta(x,y) := (xy)^\theta$, for $\theta \in (1/2, 1]$

Estimation of c

- We know $c(x,y) = c_\theta(x,y) := (xy)^\theta$, for $\theta \in (1/2, 1]$
- Estimate θ by

$$\min_{\theta,\sigma} \left\| \sigma \int_{[0,T]^2} g(x,y) c_\theta(x,y) dxdy - q(t_n) \int_{[0,T]^2} g(x,y) \widehat{c}_n(x,y) dxdy \right\|,$$

for any weight function $g : [0, T]^2 \to \mathbb{R}^p$

Estimation of c

- We know $c(x,y) = c_\theta(x,y) := (xy)^\theta$, for $\theta \in (1/2, 1]$
- Estimate θ by

$$\min_{\theta,\sigma} \left\| \sigma \int_{[0,T]^2} g(x,y) c_\theta(x,y) dxdy - q(t_n) \int_{[0,T]^2} g(x,y) \widehat{c}_n(x,y) dxdy \right\|,$$

for any weight function $g : [0, T]^2 \to \mathbb{R}^p$

- We use $g = (\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_p})$

Estimation of c

- We know $c(x,y) = c_\theta(x,y) := (xy)^\theta$, for $\theta \in (1/2, 1]$
- Estimate θ by

$$\min_{\theta,\sigma} \left\| \sigma \int_{[0,T]^2} g(x,y) c_\theta(x,y) dxdy - q(t_n) \int_{[0,T]^2} g(x,y) \widehat{c}_n(x,y) dxdy \right\|,$$

  for any weight function $g : [0, T]^2 \to \mathbb{R}^p$

- We use $g = (\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_p})$
- Minimzer $(\widehat{\theta}, \widehat{\sigma})$

Estimation of c

- We know $c(x,y) = c_\theta(x,y) := (xy)^\theta$, for $\theta \in (1/2, 1]$
- Estimate θ by

$$\min_{\theta,\sigma} \left\| \sigma \int_{[0,T]^2} g(x,y) c_\theta(x,y) dxdy - q(t_n) \int_{[0,T]^2} g(x,y) \widehat{c}_n(x,y) dxdy \right\|,$$

  for any weight function $g : [0,T]^2 \to \mathbb{R}^p$
- We use $g = (\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_p})$
- Minimzer $(\widehat{\theta}, \widehat{\sigma})$
- If $\widehat{c}_n \approx c$, then hopefully $\widehat{\theta} \approx \theta$ and $\widehat{\sigma} \approx q(t_n)$

Spatial estimation

- Get estimates $\widehat{\theta}^{(s)}$ for each observed pair s (or a subset thereof)

Spatial estimation

- Get estimates $\widehat{\theta}^{(s)}$ for each observed pair s (or a subset thereof)
- Recall that if $s = (u, u')$, $\theta^{(s)} = \Phi(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2})$

Spatial estimation

- Get estimates $\widehat{\theta}^{(s)}$ for each observed pair s (or a subset thereof)
- Recall that if $s = (u, u')$, $\theta^{(s)} = \Phi(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2})$
- Estimate the spatial parameters by

$$\min_{\alpha,\beta} \sum_{s} \left\| \Phi\left(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2}\right) - \widehat{\theta}^{(s)} \right\|^2$$

Spatial estimation

- Get estimates $\widehat{\theta}^{(s)}$ for each observed pair s (or a subset thereof)
- Recall that if $s = (u, u')$, $\theta^{(s)} = \Phi(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2})$
- Estimate the spatial parameters by

$$\min_{\alpha,\beta} \sum_s \left\| \Phi\left(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2}\right) - \widehat{\theta}^{(s)} \right\|^2$$

- In the paper, we prove joint asymptotic normality of the collection of estimators $\widehat{c}_n^{(s)}$

Spatial estimation

- Get estimates $\widehat{\theta}^{(s)}$ for each observed pair s (or a subset thereof)
- Recall that if $s = (u, u')$, $\theta^{(s)} = \Phi(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2})$
- Estimate the spatial parameters by

$$\min_{\alpha,\beta} \sum_{s} \left\| \Phi\left(\frac{1}{2}(\|u - u'\|/\beta)^{\alpha/2}\right) - \widehat{\theta}^{(s)} \right\|^2$$

- In the paper, we prove joint asymptotic normality of the collection of estimators $\widehat{c}_n^{(s)}$
- Leads to CLT's for $\widehat{\theta}^{(s)}$ and for (α, β)

Some illustration

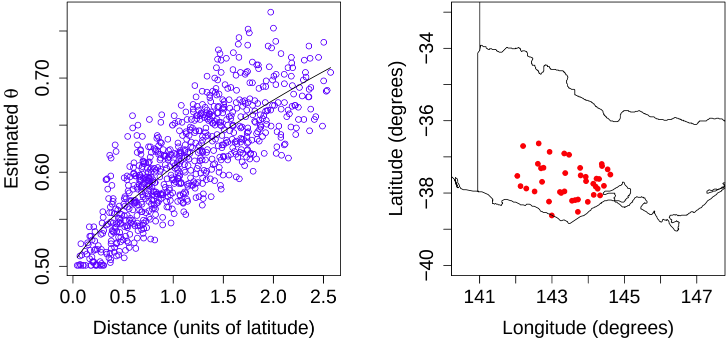

Figure: Left: Estimated parameters $\widehat{\theta}^{(s)}$ against the Euclidean distance. Right: The 40 sampled locations in the state of Victoria, southeastern Australia.

Thanks for your attention!

A few references

Brown, B. M. and S. I. Resnick (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability 14*(4), 732–739.

Lalancette, M., S. Engelke, and S. Volgushev (2021+). Rank-based estimation under asymptotic dependence and independence, with applications to spatial extremes. Ann. Stat., to appear.

Le, P. D., A. C. Davison, S. Engelke, M. Leonard, and S. Westra (2018). Dependence properties of spatial rainfall extremes and areal reduction factors. *J. Hydrol. 565*, 711–719.

mic-lalancette.github.io