Input-to-output propagation of extreme events

Thomas Opitz
BioSP, INRAE, Avignon, France

Séminaire Geolearning

Fréjus
31 mars – 3 avril 2025

GEOLEARNING
CHAIRE /// Data Science for the Environment

Input-to-output propagation

$$\boldsymbol{Y} = f(\boldsymbol{X})$$

- Vector of (simple) inputs $\boldsymbol{X} = (X_1, \ldots, X_m)^T$, e.g. white noise
- Vector of (complex) outputs $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^T$
- Transformation by algorithm f, e.g. multi-layer neural network

What is the univariate and joint tail behavior of $\boldsymbol{Y}$ given the distribution of $\boldsymbol{X}$?

Common elementary operations in generative models

Based on m independent stochastic inputs $X_1, \ldots, X_m$, we can consider:

- (Weighted) sums: $\sum_{j=1}^{m} \omega_j X_j$
- (Weighted) maxima: $\max_{j=1}^{m} \omega_j X_j$
- Threshold excess: $(X_1 - u)_+$
- Mixtures: X_J, $J \sim \mathrm{Unif}(1, \ldots, m)$

Plain vanilla neural network layer: weighted sum and thresholding (*ReLU activation*)

$$Y_i = \left(\sum_{j=1}^{m} \omega_j X_j + b_i \right)_+$$

with *weights* ω_j and *bias* b_i

The realms of light and heavy tails

Random variable $X \sim F$
Distribution function $F(x) = \Pr(X \leq x)$
Survival function $\overline{F}(x) = 1 - F(x)$

Heavy tails

F is heavy-tailed if

$$\exp(\lambda x) \times \overline{F}(x) \to \infty, \quad x \to \infty, \quad \text{for all } \lambda > 0$$

- *Exponential-tailed distributions* (within the light-tailed realm) form the **frontier**

$$\frac{\overline{F}(t+x)}{\overline{F}(t)} \to \exp(-\alpha x), \quad t \to \infty \quad (\textit{rate } \alpha)$$

- {Interesting heavy-tailed distributions} = {Subexponential distributions}
(where F is *subexponential* if $\overline{F \star F}(x)/\overline{F}(x) \to 2$, $x \to \infty$)
including {Regularly varying distributions with index $\alpha > 0$}
= {Power laws} = $\{F(\exp(\cdot))$, F exponential-tailed with rate $\alpha\}$

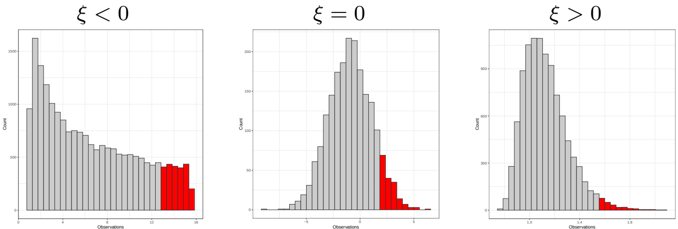
Short, light, heavy tails

Extreme-Value Theory suggests classification based on the **tail index**

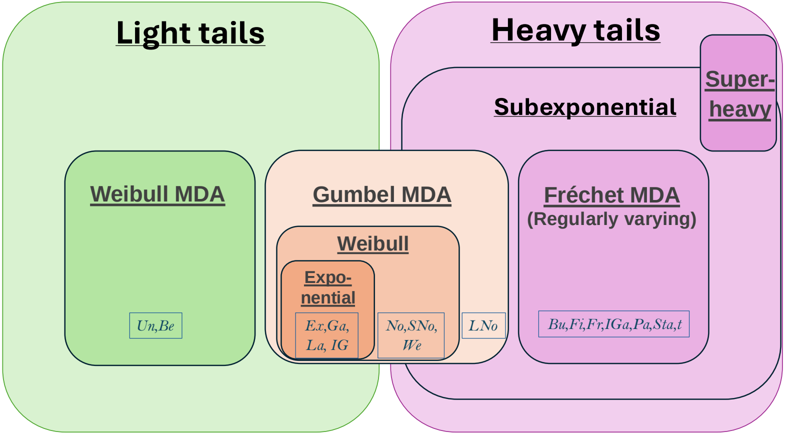
Tail index

The tail index, often denoted ξ (or γ), describes the shape of the distribution tail

- $\xi > 0$: heavy tails (power-law limit) – Fréchet domain of attraction
- $\xi = 0$: light tails (exponential limit) – Gumbel domain of attraction
- $\xi < 0$: short tails (bounded limit) – Weibull domain of attraction



General overview of common univariate tail classes



Be–Beta; Bu–Burr; Ex–Exponential; Fi–Fisher; Fr–Fréchet; Ga–Gamma; IG–Inverse-Gaussian; IGa–Inverse-Gamma; La–Laplace; LNo–Log-Normal; No–Normal; Pa–Pareto; SNo–Skew-Normal; Sta–Stable; t–student-t; Un–Uniform; We–Weibull

⚠ Gumbel domain of attraction is very large
(e.g., it includes both the Normal and the Lognormal distribution)

Focus on linear propagation from inputs to outputs

Consider the linear system

$$
\begin{aligned}
Y_1 &= b_1 + \omega_{11} X_1 + \ldots + \omega_{m1} X_m \\
Y_2 &= b_2 + \omega_{12} X_1 + \ldots + \omega_{m2} X_m \\
\ldots &= \ldots \\
Y_n &= b_n + \omega_{1n} X_1 + \ldots + \omega_{mn} X_m
\end{aligned}
$$

In matrix form: $\boldsymbol{Y} = \boldsymbol{b} + \boldsymbol{\Omega} \boldsymbol{X}$

How is the tail behavior of $\boldsymbol{Y}$ related to the tail behavior of $\boldsymbol{X}$?

- Univariate tails: behavior of $\Pr(Y_i > y)$ for large y?
  $\Rightarrow$ Strongly determined by heaviest tail among $\operatorname{sign}(\omega_{j,i} X_j)\, \omega_{j,i} X_j$, $j = 1, \ldots, m$
- Joint (esp. bivariate) tails: what behavior of $\Pr(Y_{i_1} > y, Y_{y_2} > y)$ for large y?

Tail correlation (Extremogram)

For any two random variables $X_1 \sim F_1, X_2 \sim F_2$, consider the conditional probability

$$\chi_{12}(u) = \Pr\{F_2(X_2) > u \mid F_1(X_1) > u\} = \frac{\Pr\{F_2(X_2) > u, F_1(X_1) > u\}}{\Pr\{F_1(X_1) > u\}}, \quad u \in (0,1).$$

The following limit χ_{12} (if it exists) is called **tail correlation** (or **extremogram**):

$$\chi_{12}(u) \to \chi_{12} \in [0,1], \quad u \to 1,$$

- **Asymptotic dependence** if $\chi_{12} > 0$ ($\rightsquigarrow$ **simultaneous extremes**)
- **Asymptotic independence** if $\chi_{12} = 0$

Numerical illustration of tail correlations

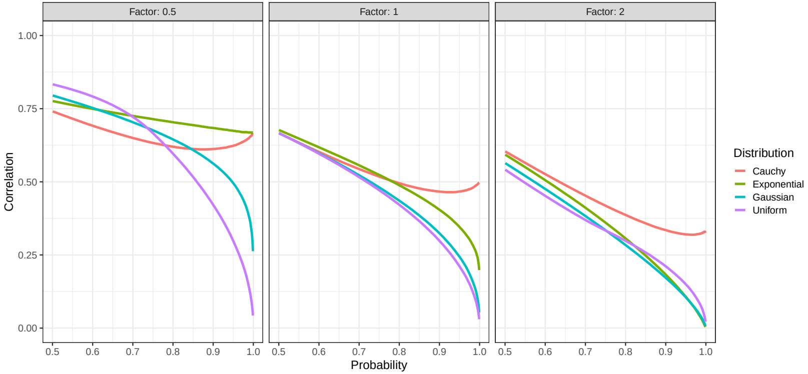
Consider the bivariate system

$$Y_1 = X_0 + \omega \times X_1$$
$$Y_1 = X_0 + \omega \times X_2$$

with

- i.i.d. distribution of X_0, X_1, X_2 among $\{\mathrm{Unif}(0,1), \mathcal{N}(0,1), \mathrm{Exp}(1), \mathrm{Cauchy}(1)\}$
- $\omega \in \{0.5, 1, 2\}$

Tail correlations $\chi_{12}(u)$ (here calculated by Monte–Carlo simulation):



Tail correlation in regularly varying setting

Consider $X_j \overset{\text{ind.}}{\sim} F_j \in \mathrm{RV}_\alpha$ such that $n \Pr(X_j/b_n > x) \to x^{-\alpha}$.
⚠ Lighter-tailed components ($n \Pr(X_j/b_n > x) \to 0$) never contribute to asymptotics!

$$Y_1 = \omega_{11}X_1 + \ldots + \omega_{1J}X_J$$
$$Y_2 = \omega_{21}X_1 + \ldots + \omega_{2J}X_J,$$

Result

Suppose that $\omega_{i,j} \geq 0$ and $\max_{j=1}^{J} \omega_{i,j} > 0$ for $i = 1, 2$. Then, the variables Y_1 and Y_2 are asymptotically dependent with

$$\chi_{12} = \sum_{j=1}^{J} \min\left(\frac{\omega_{1,j}^{\alpha}}{\sum_{k=1}^{J} \omega_{1,k}^{\alpha}}, \frac{\omega_{2,j}^{\alpha}}{\sum_{k=1}^{J} \omega_{2,k}^{\alpha}} \right).$$

⚠ Biases and thresholding (e.g. ReLU activation) do not change results (if we combine several NN layers)

⚠ More generally, the bulk of the distribution F_j has no influence

Tail correlation in exponential-tailed setting

Consider $X_j \overset{\text{ind.}}{\sim} F_j \in \mathrm{ET}_{\beta_j}$ (where $\beta_j = \infty$ if tail is lighter than exponential).

$$Y_1 = \omega_{11} X_1 + \ldots + \omega_{1J} X_J$$
$$Y_2 = \omega_{21} X_1 + \ldots + \omega_{2J} X_J,$$

- $\tilde{\omega}_{ij} = \omega_{ij,+}/\beta_j =$ scale parameter (i.e., inverse rate) of $\omega_{ij} X_j$.
- Maximum scales $\omega_i^\star = \max_j \tilde{\omega}_{ij}$, $i = 1, 2$
- Indices where the maximum is realized: $I_i = \{j : \tilde{\omega}_{ij} = \omega_i^\star\}$.

Results

1. Suppose that $I_1 = I_2$ and $\omega_i^\star > 0$, $i = 1, 2$. Then, the variables Y_1 and Y_2 are asymptotically dependent with

$$\chi_{12} = \mathbb{E}\left[\min\left(\exp(\tilde{Y}_1)/m_2, \exp(\tilde{Y}_2)/m_2\right)\right] > 0,$$

   where $\tilde{Y}_i = \sum_{j \notin I_i} X_j \omega_{ij}/\omega_i^\star$ and $m_i = \mathbb{E}[\exp(\tilde{Y}_i)]$, $i = 1, 2$.

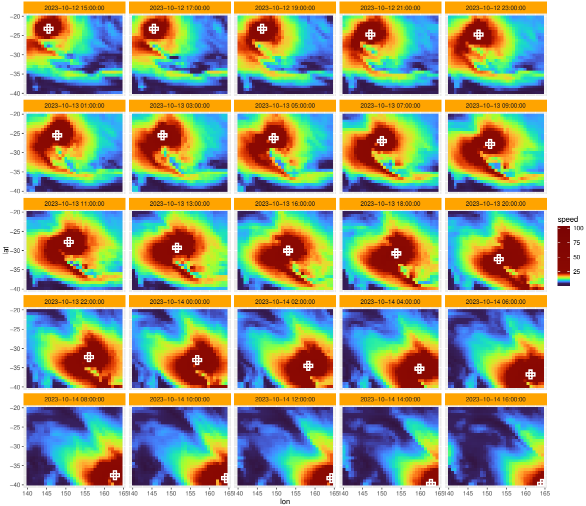
2. If $I_1 \cap I_2 = \emptyset$, then $\chi_{12} = 0$.

⚠ Biases do not change the result
(if we combine several NN layers)

⚠ Intermediate cases between 1. and 2. are more involved

Example: Modeling high and low extremes with a novel dynamic spatio-temporal model

- Work in progress...
- Bayesian dynamic spatiotemporal models
- Joint work with Myungsoo Yoo, Likun Zhang, Chris Wikle (University of Missouri)



Bayesian dynamic spatiotemporal models

Observation model:

$$\boldsymbol{y}_t = \boldsymbol{\Phi}\boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \quad \varepsilon_{t,i} \overset{\text{ind.}}{\sim} \mathcal{N}(0, \sigma^2)$$

- Observations $\boldsymbol{y}_t = (y_t(s_1), \ldots, y_t(s_n))^T$
- $\boldsymbol{\Phi} = (\Phi_{i,j})$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, with basis function values $\Phi_{i,j} = \varphi_j(s_i)$
- We can have $n \gg m$

Latent process model with dynamical evolution via autoregression:

$$\boldsymbol{\alpha}_t = M\boldsymbol{\alpha}_{t-1} + \boldsymbol{\omega}_t$$

- Transition matrix M (autoregression, dilation, translation...)
- Classical dynamic models have Gaussian innovations $\boldsymbol{\omega}_t \sim \mathcal{N}(\mathbf{0}, \sigma_\omega I_m)$
- Estimation of $\boldsymbol{\omega}_t$, M and hyperparameters using MCMC

Goals :

- Regime-switching latent process for detecting extremes
- Asymptotic dependence around locally extreme states
- Theoretical characterization of spatial and temporal dependence

Conclusion

- Work in progress...

- Next steps:

  1. A catalog of univariate and joint tail behavior for elementary operations
  2. Tail behavior when composing several elementary operations
  3. Results for specific models (e.g. dynamic models as outlined here)
  4. Construct new generative models (e.g. with heavy tails)