NONPARAMETRIC ESTIMATION OF THE PICKANDS DEPENDENCE FUNCTION USING BERNSTEIN POLYNOMIALS

GIULIA MARCON[1]

giulia.marcon@phd.unibocconi.it

Joint work with: SIMONE PADOAN[1], PHILIPPE NAVEAU[2], PIETRO MULIERE[1]

[1] Bocconi University (Italy)

[2] Laboratoire des Sciences du Climat et l'Environnement (France)

AUSSOIS, 15-19 DECEMBER 2013

Motivation: Modelling extremes with real applications. This also requires **modelling the extremal dependence**

Ex. High-dimensional data (e.g. Pollution)
Spatial processes (e.g. Rainfalls)

↓

Aim: To Provide a non-parametric estimator of the dependence structure

Motivation: Modelling extremes with real applications. This also requires **modelling the extremal dependence**

Ex. High-dimensional data (e.g. Pollution)
Spatial processes (e.g. Rainfalls)

↓

Aim: To Provide a non-parametric estimator of the dependence structure

Outline:

MULTIVARIATE EXTREMES FRAMEWORK

Let $\mathbf{X}_i = (X_{i,1}, \ldots, X_{i,d})$ be a vector of component-wise maxima (Resnick, 1987)

Characterization (de Haan and Ferreira, 2006)

- MEVD with unit Fréchet margins can be written as

$$G(x_1, \ldots, x_d) = \exp\left(-V(x_1, \ldots, x_d)\right)$$

- **exponent measure**

$$V(x_1, \ldots, x_d) = \nu\left(((0, x_1] \times \cdots \times (0, x_d])^c\right) = \int_{S_d} \bigvee_{i=1}^{d} \left(\frac{w_i}{x_i}\right) H(d\mathbf{w})$$

- H is a **probability measure** satisfying the mean constraint

$$\int_{S_d} w_i H(d\mathbf{w}) = \frac{1}{d} \quad \text{for} \quad i = 1, \ldots, d$$

PROBLEMS AND SOLUTIONS

1. In order to get $\mathbf{G}(\mathbf{x})$ we need to integrate out $\mathbf{H(dw)}$ and it can be **DIFFICULT**
2. The family of admissible dependence structures **CAN NOT BE PARAMETRIZED**
3. V needs to satisfy specific properties so that $\mathbf{G}$ is a distribution

→

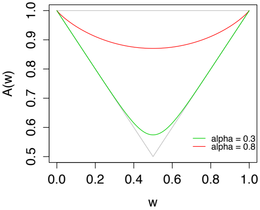
1. Focus on the exponent function $V(\mathbf{x})$
2. Use a **NON-PARAMETRIC** estimator
3. Derive an estimator **SUBJECT** to **CONSTRAINTS**

Pickands Dependence Function

$$G(\mathbf{x}) = \exp\left\{-\left(\frac{1}{x_1} + \ldots + \frac{1}{x_d}\right) A(\mathbf{w})\right\}$$

$$A : \mathcal{S}_{\bar{d}} \to [0,1] \qquad \bar{d} = d - 1$$

$$\mathcal{S}_{\bar{d}} := \left\{(w_1, \ldots, w_{\bar{d}}) \in [0,1]^{\bar{d}} : \sum_{i=1}^{\bar{d}} w_i \leq 1,\ w_i \geq 0\right\}$$



Properties

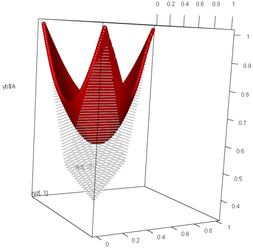
P1) $A(\mathbf{w})$ is continuous and convex

P2) $A(\mathbf{w}) \geq \max\left(w_1, \ldots, w_{\bar{d}}, 1 - \sum_{i \leq \bar{d}} w_i\right)$

$$1/d \leq A(\mathbf{w}) \leq 1$$

P3) $A(\boldsymbol{e_i}) = 1,\ 1 \leq i \leq \bar{d},$

$$A(\mathbf{0}) = 1$$



PROPOSAL

Consider a multivariate estimator based on **BERNSTEIN POLYNOMIALS** (Bayad, 2011)

Let $k \in \mathbb{N}$ be the order of the polynomial and $J = \{0, \ldots, k\}$ be the index set. Let $b_\ell : \mathcal{S}_{\bar{d}} \to [0, 1]$ be the ℓth Bernstein polynomial's base, that is

$$b_\ell(\mathbf{w}; k) = \binom{k}{\alpha_\ell} \mathbf{w}^{\alpha_\ell} (1 - |\mathbf{w}|)^{k - |\alpha_\ell|},$$

where

$$\mathbf{w}^{\alpha_\ell} = \prod_{i=1}^{\bar{d}} w_i^{j_i} \qquad \alpha_\ell! = j_1! \cdots j_{\bar{d}}! \qquad \binom{k}{\alpha_\ell} = \frac{k!}{\alpha_\ell!(n - |\alpha_\ell|)!}.$$

Let $L_k = \{\ell \in L : |\alpha_\ell| \leq k\}$. Then, the Bernstein representation of A is

$$B_A(\mathbf{w}; k) = \sum_{\ell \in L_k} \beta_\ell b_\ell(\mathbf{w}; k), \tag{1}$$

where $\beta_\ell \in \mathbb{R}$, $\ell \in L_k$ are coefficients.

Estimation Procedure

We want to estimate the Pickands Dependence Function $A(\mathbf{w})$

1. Transform data in pseudo-polar coordinates

$$r = \sum_{j=1}^{d} x_j \qquad \text{and} \qquad w_j = \frac{x_j}{r}$$

2. Obtain a first guess $\hat{A}$ with some non-parametric estimator

3. Derive the projection of $\hat{A}$ in the sub-space of $A(\mathbf{w})$ functions satisfying the desired properties

4. It can be done solving the optimization problem

$$\begin{cases} \min & \frac{1}{n}\sum_{i=1}^{n}\left(\mathbf{b}(\mathbf{w}_i;k)\beta - \hat{A}(\mathbf{w}_i)\right)^2 \\ \text{s.t.} & \boldsymbol{C}\beta \geq \mathbf{c} \end{cases}$$

5. The solution is

$$\tilde{A}(\mathbf{w}) = B_A(\mathbf{w};k) = \mathbf{b}(\mathbf{w},k)\tilde{\beta}$$

which coefficients $\tilde{\beta}$ fully characterize the function

Constraints are...

P1 - Convexity.
Hessian matrix must be positive semi-definite $\forall \mathbf{w} \in \mathcal{S}_{\bar{d}}$

$$\Downarrow$$

$$\mathbf{C}^{(1)}$$

P2 - Boundaries.
The optimization problem is verified for each observation

$$\Downarrow$$

$$\mathbf{C}^{(2)}$$

P3 - Endpoints.
The $\bar{d}$ vertices, and respectively some β_ℓ, must be equal to one

$$\Downarrow$$

$$\mathbf{C}^{(3)}$$

Constraints are...

P1 - Convexity.
Hessian matrix must be positive semi-definite $\forall \mathbf{w} \in \mathcal{S}_{\bar{d}}$

$\Downarrow$

$\mathbf{C}^{(1)}$

P2 - Boundaries.
The optimization problem is verified for each observation

$\Downarrow$

$\mathbf{C}^{(2)}$

P3 - Endpoints.
The $\bar{d}$ vertices, and respectively some β_ℓ, must be equal to one

$\Downarrow$

$\mathbf{C}^{(3)}$

Combine **all** restriction matrices which respectively satisfy **P1** - **P3**

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}^{(1)} \\ \mathbf{C}^{(2)} \\ \mathbf{C}^{(3)} \end{bmatrix}$$

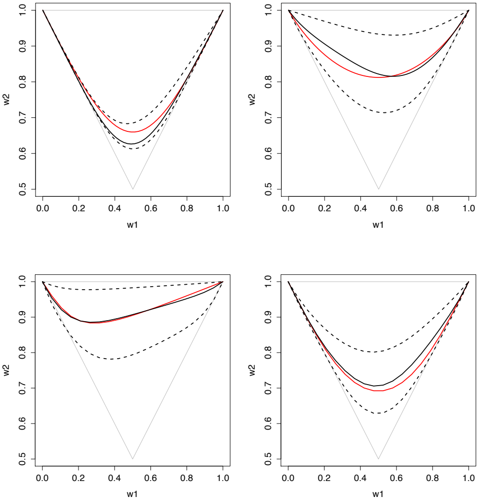
CONFIDENCE BANDS

- We use bootstrap confidence bands (say $R = 500$ or $2,000$ replicates)
- Instead of deriving non-parametric $(100 - \alpha)\%$ individual confidence bands from the quantiles of the bootstrap sampling $\tilde{A}^*(w)$
- We construct them, for each β_ℓ, from the quantiles of the bootstrap sampling distribution of $\tilde{\beta}_\ell^*$

$$\mathbb{P}\left\{ \tilde{\beta}_{\ell}{}^*_{(R\alpha/2)} \le \beta_\ell \le \tilde{\beta}_{\ell}{}^*_{(R(1-\alpha/2))} \right\} = 1 - \alpha \qquad \forall \ell \in L_k,$$

- The benefit of this is that confidence bands are still Pickands dependence function....(not 100% sure)

NUMERICAL RESULTS (d=2)

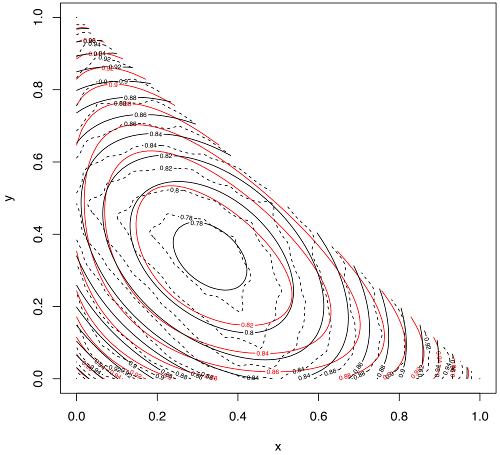


True dependence function of the models (red):

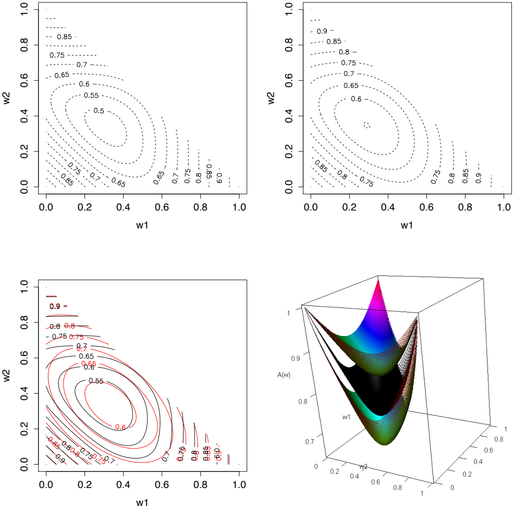
- **Symmetric Logistic Model** with dependence parameter 0.4 and 0.7
- **Asymmetric Logistic Model** with dependence parameter 0.4 and $t_1 = 0.2, t_2 = 0.8$
- **Hüsler-Reiss Model** with $\lambda = 1$
- $n = 50$
- $ngrid = 200$
- $\tilde{A}(\mathbf{w})$ (solid black), Confident Bands (dotted black)

NUMERICAL RESULTS (d=3)

- True dependence function of the **Symmetric Logistic Model** (red) with dependence parameter $\alpha = 0.8$
- $n = 10$
- $ngrid = 200$
- $\hat{A}(\mathbf{w})$ (dashed), $\tilde{A}(\mathbf{w})$ (solid)

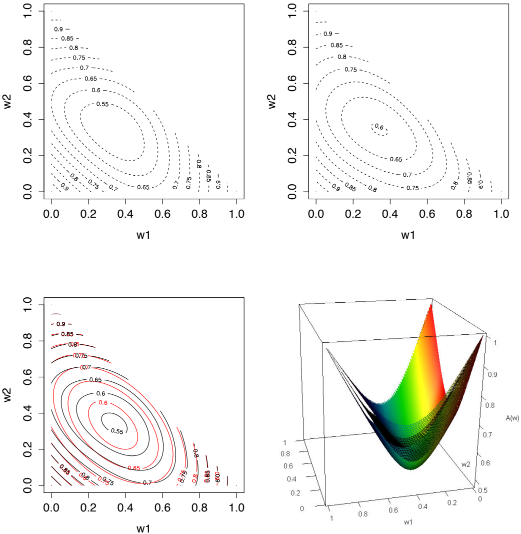


Numerical Results (d=3)



- True dependence function of the **Symmetric Logistic Model** (red) with dependence parameter $\alpha = 0.8$
- $n = 50$
- $ngrid = 200$
- $\tilde{A}(\mathbf{w})$ (solid)
- Confident Bands (dashed)

Numerical Results (d=3)



- True dependence function of the **Symmetric Logistic Model** (red) with dependence parameter $\alpha = 0.5$
- $n = 50$
- $ngrid = 200$
- $\tilde{A}(\mathbf{w})$ (solid)
- Confident Bands (dashed)

Properties

For deriving asymptotic properties of $\tilde{A}$ we need $\tilde{\beta}$ to be bounded

Lemma

- *Consider a d-dimensional vector* $\boldsymbol{X} = (X_1, \ldots, X_d)$ *and let* $k = 0, 1, 2, \ldots$ *be the order of the polynomial*
- *The* ***number of coefficients****,* $p_k \equiv p(k, \bar{d})$ *depend on both k and d*

$$p_k = \begin{cases} k+1 & \text{if } \bar{d} = 1, \\ \sum_{m=1}^{k+1} (k+2-m) & \text{if } \bar{d} = 2, \\ \sum_{m=1}^{k+1} \binom{\bar{d}+m-4}{m-1} \frac{(k+2-m)^2 + (k+2-m)}{2} & \text{if } \bar{d} \geq 3 \end{cases}$$

MAIN RESULT

Weak Consistency

- Assume $\tilde{N} = O(N^r)$ and $k = O(\tilde{N}^\alpha)$, $0 < r, \alpha < 1$
- q_k = rank of the restriction matrix C which is assumed to be of full rank
- For Lemma we have

$$\sum_{\ell \in L_k} |\beta_\ell| \leq p_k \qquad \text{and} \qquad \sup_{\mathbf{w} \in \mathcal{S}_{\bar{d}}} A(\mathbf{w}) < \infty,$$

- If, for some $0 < \delta < 1$ and $N \to \infty$,

$$q_k \to \infty \qquad p_k \to \infty \qquad \frac{q_k p_k^4 log(p_k)}{\tilde{N}} \to 0 \qquad \frac{p_k^4}{k^{1-\delta}} \to 0,$$

- Then

$$\sup_{\mathbf{w} \in \mathcal{S}_{\bar{d}}} \left(\tilde{A}(\mathbf{w}) - A(\mathbf{w}) \right) = O(N^l)$$

where $l \in \mathbb{R}$.

CONCLUSION

1. We proposed a method for deriving a multivariate non-parametric estimator of the Pickands satisfying the required properties

2. The method is based on Bernstein polynomials.
   The benefits of using them are:

- Compared to other polynomials, they have **optimal properties** in shape-preserving estimation problems (Wang and Ghosh, 2012)
- **Fast computation** in high-dimensions
- The resulting estimator has **nice asymptotic properties**, under opportune conditions, as sample size increases
- Bootstrap **confidence bands** are easy to obtain (lower and upper bounds are proper Pickands.... to be formalized)

References

- **de Haan L., Ferreira A.** (2006) *Extreme Value Theory: An Introduction.* Springer, Boston.
- **Resnick S.I.** (1987) *Extreme values, Regular Variation and Point Processes.* Berlin: Springer-Verlag
- **Goldfarb, D. and Idnani, A.** (1983) *A numerically stable dual method for solving strictly convex quadratic programs.* Mathematical programming, 27, 1, 1-33
- **Lai, M.J.** (1993) *Some sufficient conditions for convexity of multivariate Bernstein-Bézier polynomials and box spline surfaces.* Studia Scientiarum Mathematicarum Hungarica, 28, 3, 363-374
- **Wang, J. and Ghosh, S.K.** (2012) *Shape restricted nonparametric regression with Bernstein polynomials.* Computational Statistics and Data Analysis, 56,9, 2729 - 2741


Thank You


THANK YOU