Multivariate L-Moments Based on Transports

Alexis Decurninge
Huawei Technologies

Geometric Science of Information
October 29th, 2015

Outline

Definition of L-moments

L-moments of a distribution :
if $X_1,...,X_r$ are real random variables with common cumulative distribution function F

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k}\mathbb{E}[X_{r-k:r}]$$

with $X_{1:r} \leq X_{2:r} \leq ... \leq X_{r:r}$: **order statistics**

- $\lambda_1 = \mathbb{E}[X]$: localization
- $\lambda_2 = \mathbb{E}[X_{2:2} - X_{1:2}]$: dispersion
- $\tau_3 = \frac{\lambda_3}{\lambda_2} = \frac{\mathbb{E}[X_{3:3} - 2X_{2:3} + X_{1:3}]}{\mathbb{E}[X_{2:2} - X_{1:2}]}$: asymmetry
- $\tau_4 = \frac{\lambda_4}{\lambda_2} = \frac{\mathbb{E}[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}]}{\mathbb{E}[X_{2:2} - X_{1:2}]}$: kurtosis

Existence if $\int |x| dF(x) < \infty$

Characterization of L-moments

L-moments are projections of the **quantile** function on an orthogonal basis

$$\lambda_r \quad = \quad \int_0^1 F^{-1}(t) L_r(t) dt$$

- F^{-1} generalized inverse of F

$$F^{-1}(t) = \inf \{x \in \mathbb{R} \text{ such that } F(x) \geq t\}$$

- L_r Legendre polynomial (orthogonal basis in $L_2([0,1])$)

$$L_r(t) = \sum_{k=0}^{r} (-1)^k \binom{r}{k}^2 t^{r-k} (1-t)^k$$

- L-moments completely **characterize a distribution**

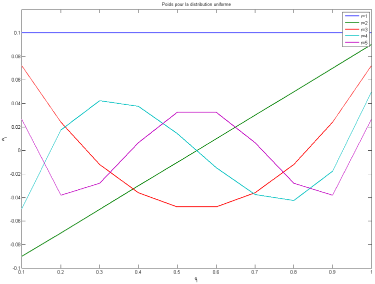
$$F^{-1}(t) = \sum_{r=1}^{\infty} (2r+1) \lambda_r L_r(t)$$

Definition of L-moments (discrete distributions)

L-moments for a multinomial distribution of support $x_1 \leq x_2 \leq \ldots \leq x_n$ and weights $\pi_1, \ldots, \pi_n$ ($\sum_{i=1}^{n} \pi_i = 1$)

$$\lambda_r = \sum_{i=1}^{n} w_i^{(r)} x_i = \sum_{i=1}^{n} \left[K_r \left(\sum_{a=1}^{i} \pi_a \right) - K_r \left(\sum_{a=1}^{i-1} \pi_a \right) \right] x_i$$

with K_r the respective primitive of L_r : $K_r' = L_r$



Empirical L-moments

- U-statistics : mean of all subsequences of size r without replacement

$$\hat{\lambda}_r = \frac{1}{\binom{n}{r}} \sum_{1 \le i_1 < \cdots < i_r \le n} \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k}:n}.$$

- V-statistics : biased estimator (with replacement)

$$\begin{aligned} \hat{\lambda}_r^{(V)} &= \int_0^1 F_n^{-1}(t) L_r(t) dt \\ &= \frac{1}{\binom{r+n-1}{n-1}} \sum_{1 \le i_1 \le \cdots \le i_r \le n} \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k}:n}. \end{aligned}$$

Motivations

- Interest in **multivariate heavy tailed** distributions : multivariate dispersion, skewness, kurtosis ?
- Applications e.g. in portfolio selection by Jurzenko et al., extreme hydrological events
- Desirable properties of univariate L-moments
  - Existence since the existence is finite
  - A distribution is **characterized** by its L-moments
  - An outlier x has a **linear impact** in the estimation of L-moments while it is of order $(x - \bar{x})^k$ for classical moments of order k
  - A characterization as a **scalar product** of the quantile onto an orthogonal basis

$$\lambda_r = \int Q(t) L_r(t) dt$$

Multivariate quantiles : state of the art

Since $\mathbb{R}^d$ is not totally ordered if $d > 1$, many multivariate quantiles has been proposed

- Quantiles coming from depth functions (Tukey, Zuo and Serfling)
- Spatial Quantiles (Chaudhuri)
- Generalized quantile processes (Einmahl and Mason)
- **Quantiles as quadratic optimal transports** (Galichon and Henry)

Multivariate quantiles

- We define a **quantile** related to a probability measure ν as a transport from the uniform measure *unif* on $[0;1]^d$ into ν.

Definition

Let $\boldsymbol{U}$ and $\boldsymbol{X}$ are random variables with respective measure μ and ν.
T is a **transport** from μ into ν if $T(\boldsymbol{U}) = \boldsymbol{X}$ (we note $T\#\mu = \nu$).

- Example of transport families
  - Optimal/monotone transports
  - Rosenblatt transports
  - Moser transports
  - ...

Multivariate L-moments

- $\boldsymbol{X}$ r.v. of interest with related measure ν such that $\mathbb{E}[\|\boldsymbol{X}\|] < \infty$.

Definition

- $\boldsymbol{Q} : [0;1]^d \to \mathbb{R}^d$ a **transport** from *unif* in $[0;1]^d$ into ν.
- **L-moment** λ_α of multi-index $\alpha = (i_1, ..., i_d)$ associated to $\boldsymbol{Q}$:

$$\boldsymbol{\lambda}_\alpha := \int_{[0;1]^d} \boldsymbol{Q}(t_1, ..., t_d) L_\alpha(t_1, ..., t_d) dt_1 ... dt_d \in \mathbb{R}^d.$$

- with $L_\alpha(t_1, ..., t_d) = \prod_{k=1}^d L_{i_k}(t_k)$.

$\Rightarrow$ Definition **compatible with the univariate case** : the univariate quantile is a transport from the uniform measure on $[0;1]$ into the measure of interest ($F^{-1}(U) \overset{d}{=} X$)

Multivariate L-moments

- L-moment of degree 1

$$\boldsymbol{\lambda}_1 (= \lambda_{1,1,\dots,1}) = \int_{[0;1]^d} \boldsymbol{Q}(t_1, \dots, t_d) dt_1 \dots dt_d = \mathbb{E}[\boldsymbol{X}].$$

- L-moments of degree 2 can be regrouped in a matrix

$$\Lambda_2 = \left[\int_{[0;1]^d} Q_i(t_1, \dots, t_d)(2t_j - 1) dt_1 \dots dt_d \right]_{1 \le i,j \le d}.$$

with

$$\boldsymbol{Q}(t_1, \dots, t_d) = \begin{pmatrix} Q_1(t_1, \dots, t_d) \\ \vdots \\ Q_d(t_1, \dots, t_d) \end{pmatrix}$$

Multivariate L-moments : characterization

Proposition

Assume that two quantiles $\boldsymbol{Q}$ *and* $\boldsymbol{Q}'$ *have same multivariate L-moments* $(\lambda_\alpha)_{\alpha \in \mathbb{N}_*^d}$ *then* $\boldsymbol{Q} = \boldsymbol{Q}'$.
Moreover

$$\boldsymbol{Q}(t_1, \ldots, t_d) = \sum_{(i_1,\ldots,i_d) \in \mathbb{N}_*^d} \left(\prod_{k=1}^{d} (2i_k + 1) \right) L_{(i_1,\ldots,i_d)}(t_1, \ldots, t_d) \boldsymbol{\lambda}_{(i_1,\ldots,i_d)}$$

A one-to-one correspondence between quantiles and random vectors is sufficient to guarantee the characteriation of a distribution by its L-moments

Monotone transport

Proposition

Let μ, ν be two probability measures on $\mathbb{R}^d$, such that μ does not give mass to "small sets".
*Then, there is exactly one measurable map $\boldsymbol{T}$ such that $\boldsymbol{T} \# \mu = \nu$ and $\boldsymbol{T} = \nabla \varphi$ for some **convex** function φ.*

- These transports, **gradient of convex functions**, are called **monotone transports** by analogy with the univariate case
- If defined, the transport is solution to the quadratic optimal transport

$$\nabla \varphi^* = \arg \inf_{\boldsymbol{T}:\boldsymbol{T}\#\mu=\nu} \int_{\mathbb{R}^d} \|u - \boldsymbol{T}(u)\|^2 d\mu(u)$$

Example : monotone quantile for a random vector with independent marginals

- $\boldsymbol{X} = (X_1, ..., X_d)$ random vector with **independent marginals**.
- The monotone quantile of $\boldsymbol{X}$ is the collection of its marginals quantiles

$$\boldsymbol{Q}(t_1, ..., t_d) = \begin{pmatrix} Q_1(t_1) \\ \vdots \\ Q_d(t_d) \end{pmatrix} = \begin{pmatrix} \nabla\phi_1(t_1) \\ \vdots \\ \nabla\phi_d(t_d) \end{pmatrix}$$

- Indeed, if $\phi(t_1, ..., t_d) = \phi_1(t_1) + \cdots + \phi_d(t_d)$

$$\nabla\phi = \boldsymbol{Q}$$

- The associated L-moments are then

$$\begin{cases} \boldsymbol{\lambda}_{1,...,1} & = & \mathbb{E}[\boldsymbol{X}] \\ \boldsymbol{\lambda}_{1...1,r,1,...,1} & = & (0, \ldots, 0, \lambda_r(X_i), 0, \ldots, 0)^T \\ \boldsymbol{\lambda}_{\alpha} & = & 0 \;\; \text{otherwise} \end{cases}$$

Monotone transport from the standard Gaussian distribution

- $\boldsymbol{Q}_{\mathcal{N}}$ the monotone distribution from *unif* onto the standard Gaussian distribution $\mathcal{N}(0, I_d)$ defined by

$$\boldsymbol{Q}_{\mathcal{N}}(t_1, .., t_d) = \begin{pmatrix} \mathcal{N}^{-1}(t_1) \\ \vdots \\ \mathcal{N}^{-1}(t_d) \end{pmatrix}$$

- $\boldsymbol{T}_0$ the monotone transport from the standard Gaussian distribution from ν (rotation equivariant)

$$([0;1]^d, du) \overset{\boldsymbol{Q}_{\mathcal{N}}}{\rightarrow} (\mathbb{R}^d, d\mathcal{N}) \overset{\boldsymbol{T}_0}{\rightarrow} (\mathbb{R}^d, d\nu)$$

$\Rightarrow \boldsymbol{Q} = \boldsymbol{T}_0 \circ \boldsymbol{Q}_{\mathcal{N}}$ is then a **quantile**.

Monotone transport from the standard Gaussian distribution : Gaussian distribution with a random covariance

- For $\boldsymbol{x} \in \mathbb{R}^d$, A positive symmetric matrix

$$\begin{cases} \varphi(\boldsymbol{x}) & = & \boldsymbol{m}.\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^T A \boldsymbol{x} \\ \boldsymbol{T}_0(\boldsymbol{x}) & = & \nabla\varphi(\boldsymbol{x}) = \boldsymbol{m} + A\boldsymbol{x} \end{cases}$$

$\Rightarrow \boldsymbol{T}_0(\mathcal{N}_d(0, I_d)) \stackrel{d}{=} \mathcal{N}_d(\boldsymbol{m}, AA^T)$.

- The L-moments of a Gaussian with mean $\boldsymbol{m}$ and covariance AA^T are :

$$\boldsymbol{\lambda}_\alpha = \begin{cases} \boldsymbol{m} & \text{if } \alpha = (1, ..., 1) \\ A\boldsymbol{\lambda}_\alpha(\mathcal{N}_d(0, I_d)) & \text{otherwise} \end{cases}$$

- In particular, the L-moments of degree 2 :

$$\Lambda_2 = (\lambda_{2,1...,1} \ldots \lambda_{1,...,1,2}) = \frac{1}{\sqrt{\pi}} A.$$

Monotone transport from the standard Gaussian distribution : quasi-elliptic distribution

- For $\boldsymbol{x} \in \mathbb{R}^d$, u convex

$$\left\{ \begin{array}{rcl} \varphi(\boldsymbol{x}) & = & \boldsymbol{m}.\boldsymbol{x} + \frac{1}{2}u(\boldsymbol{x}^T A\boldsymbol{x}) \\ \boldsymbol{T}_0(x) & = & \boldsymbol{m} + u'(\boldsymbol{x}^T A\boldsymbol{x})A\boldsymbol{x}. \end{array} \right.$$

- The L-moments of this distribution are then

$$\boldsymbol{\lambda}_\alpha = \left\{ \begin{array}{ll} \boldsymbol{m} & \text{if } \alpha = (1, ..., 1) \\ A \int_{\mathbb{R}^d} u'(\boldsymbol{x}^T A\boldsymbol{x}) L_\alpha(\mathcal{N}(\boldsymbol{x}))\boldsymbol{x} d\mathcal{N}(\boldsymbol{x}) & \text{otherwise} \end{array} \right.$$

- Si $A = I_d$, $\boldsymbol{T}_0(\boldsymbol{X})$ follows a spherical distribution

Monotone transport from the standard Gaussian distribution : quasi-elliptic distribution

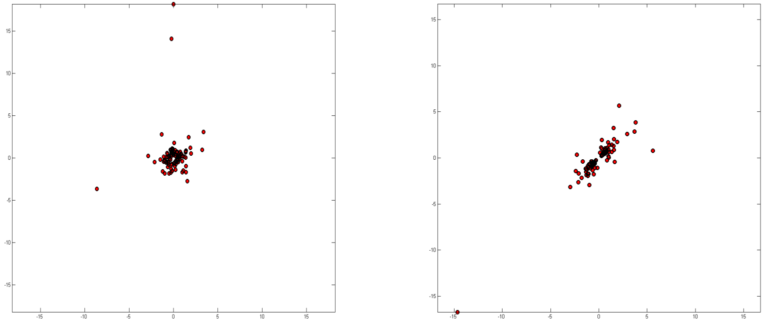


Figure: Samples with $\boldsymbol{T}_0(\boldsymbol{x}) = -\frac{A\boldsymbol{x}}{\boldsymbol{x}^T A \boldsymbol{x}}$ and $A = I_d$ (left) or $A = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$ (right)

Estimation : general case

- $\boldsymbol{x}_1, ..., \boldsymbol{x}_n \in \mathbb{R}^d$ an iid sample issued from a same r.v. $\boldsymbol{X}$ with measure ν of quantile $\boldsymbol{Q}$.
- Empirical measure : $\nu_n = \sum_{i=1}^n \delta_{\boldsymbol{x}_i}$
- Estimation of $\boldsymbol{Q}$: $\boldsymbol{Q}_n$ corresponding transport from *unif* onto ν_n
- Empirical L-moment

$$\hat{\boldsymbol{\lambda}}_\alpha = \int_{[0;1]^d} \boldsymbol{Q}_n(t) L_\alpha(t) dt$$

Estimation of a monotone transport

Monotone transport of an absolutely continuous measure μ (of support Ω) onto the discrete measure ν_n

- **Power diagrams** of $(\boldsymbol{x}_1, w_1), ..., (\boldsymbol{x}_n, w_n)$

$$\bigcup_{1 \leq i \leq n} \left\{ \boldsymbol{u} \in \Omega \text{ s.t. } \|\boldsymbol{u} - \boldsymbol{x}_i\|^2 + w_i \leq \|\boldsymbol{u} - \boldsymbol{x}_j\|^2 + w_j \;\; \forall j \neq i \right\}$$

- **Piecewise linear functions (PL)**

$$\text{for any } \boldsymbol{u} \in \Omega, \quad \phi_h(\boldsymbol{u}) = \max_{1 \leq i \leq n} \{\boldsymbol{u}.\boldsymbol{x}_i + h_i\}.$$

- Areas of gradient of a PL function
  = power diagrams with weights $w_i = \|\boldsymbol{x}_i\|^2 + 2h_i$

$$W_i(\boldsymbol{h}) = \{\boldsymbol{u} \in \Omega \text{ s.t. } \nabla\phi_h(u) = \boldsymbol{x}_i\}.$$

Estimation of a monotone transport

- Gradient of PL functions $\Rightarrow$ Monotone transport

Theorem

$\nabla\phi_h$ is a monotone transport from μ onto ν_n for some $\boldsymbol{h} = \boldsymbol{h}^*$, unique up to constant $(b, ..., b)$, verifying

$$\boldsymbol{h}^* = \arg\min_{\boldsymbol{h}\in\mathbb{R}^n} \int_\Omega \phi_h(u)d\mu - \frac{1}{n}\sum_{i=1}^{n} h_i.$$

- For any $1 \leq i \leq n$

$$\int_{W_i(\boldsymbol{h}^*)} d\mu(x) = \frac{1}{n}$$

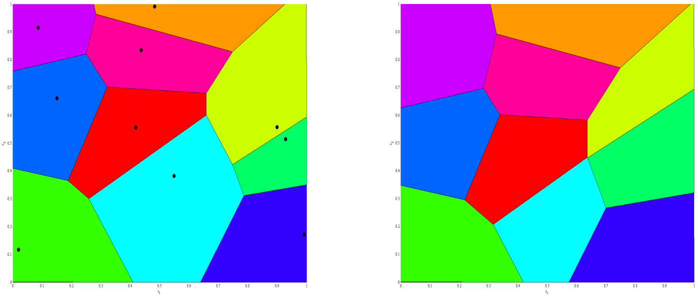
Estimation of a monotone transport : Newton descent

- Computation of h^* : minimization of

$$E(\boldsymbol{h}) = \int_\Omega \phi_h(u) d\mu - \frac{1}{n}\sum_{i=1}^{n} h_i$$

- Gradient descent :
  - **while** $|\nabla E(\boldsymbol{h}_t)| > \eta$
  - $\quad \boldsymbol{h}_{t+1} = \boldsymbol{h}_t - \gamma(\nabla^2 E(\boldsymbol{h}_t))^{-1}\nabla E(\boldsymbol{h}_t)$
  - $\quad t \leftarrow t+1$
  - **end while**
- **However** : delicate Hessian computation

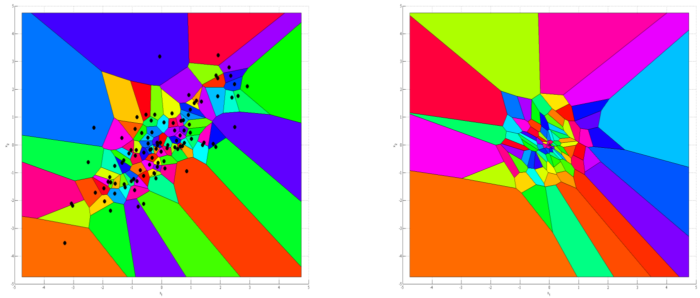
Estimation of a monotone transport : sample in $[0;1]^2$



Voronoi cells of the sample   Optimal power diagram

Figure: Monotone transport for a sample of size 10 onto the uniform distribution on $[0;1]^2$

Estimation of a monotone transport : Gaussian sample



Voronoi cells of the sample — Optimal power diagram

Figure: Monotone transport for a sample of size 100 onto the standard Gaussian

Estimation of a monotone transport : consistency

- $\boldsymbol{T}$ transport from μ onto ν
- $\boldsymbol{T}_n$ transport from μ onto ν_n

Theorem

If ν verifies $\int \|x\| d\nu(x) < +\infty$,

$$\|\boldsymbol{T} - \boldsymbol{T}_n\|_1 = \int_{\mathbb{R}^d} \|\boldsymbol{T}(x) - \boldsymbol{T}_n(x)\| d\mu(x) \overset{a.s.}{\to} 0.$$

- $\boldsymbol{Q}$, $\boldsymbol{Q}_n$ monotone quantiles having μ as a reference measure

Theorem

For $\alpha \in \mathbb{N}_*^d$.

$$\hat{\boldsymbol{\lambda}}_\alpha = \int_{[0;1]^d} \boldsymbol{Q}_n(u) L_\alpha(u) du \overset{a.s.}{\to} \lambda_\alpha = \int_{[0;1]^d} \boldsymbol{Q}(u) L_\alpha(u) du$$

Numerical applications

We simulate a linear combination of independent vectors in $\mathbb{R}^2$

$$\boldsymbol{X} = P \begin{pmatrix} \sigma_1 Z_1 \\ \sigma_2 Z_2 \end{pmatrix}$$

with

- P a rotation matrix

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

- Z_1, Z_2 are drawn from a symmetrical Weibull distribution ϵW_ν of scale parameter equal to 1 and shape parameter $\nu = 0.5$.

Numerical applications

We perform $N = 100$ estimations of the second L-moment matrix Λ_2 and the covariance matrix Σ for a sample of size $n = 30$ or 100.

- The **mean** of the different estimates
- The **median** of the different estimates
- The **coefficient of variation** of the estimates $\hat{\theta}_1, ..., \hat{\theta}_N$ (for an arbitrary parameter θ)

$$CV = \frac{\left(\sum_{i=1}^{N}\left(\hat{\theta}_i - \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_i\right)^2\right)^{1/2}}{\frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_i}$$

		$n = 30$			$n = 100$		
Parameter	True Value	Mean	Median	CV	Mean	Median	CV
$\Lambda_{2,11}$	0.38	0.28	0.27	**0.30**	0.38	0.37	**0.18**
$\Lambda_{2,12}$	0.19	0.14	0.13	**0.65**	0.20	0.20	**0.33**
Σ_{11}	0.69	**0.70**	0.48	1.23	**0.69**	0.59	0.55
Σ_{12}	0.55	**0.55**	0.29	1.62	**0.55**	0.47	0.67

Thank you for your attention !