Some multivariate risk indicators; minimization by using stochastic algorithms

Véronique Maume-Deschamps, université Lyon 1 - ISFA,
Joint work with P. Cénac and C. Prieur.

AST&Risk (ANR Project)

Multivariate risk process

Consider a vectorial risk process

$$X_n = \begin{pmatrix} X_n^1 \\ \vdots \\ X_n^d \end{pmatrix},$$

X_n^k = gains of the kth branch of a compagny, during the nth period, *i.e.* $X_n^k = G_n^k - L_n^k$ where G_n^k is the gain and L_n^k the loss. There may be dependence: *vectorial* (with respect to k) and/or *temporal* (with respect to n).

Multivariate risk process

Consider a vectorial risk process

$$X_n = \begin{pmatrix} X_n^1 \\ \vdots \\ X_n^d \end{pmatrix},$$

X_n^k = gains of the kth branch of a compagny, during the nth period, *i.e.* $X_n^k = G_n^k - L_n^k$ where G_n^k is the gain and L_n^k the loss.

There may be dependence: *vectorial* (with respect to k) and/or *temporal* (with respect to n).

$u > 0$: total capital of the compagny.

How to allocate $u = u_1 + \cdots + u_d$ between the d branches in an optimal way? (u_k is allocated to the kth branch).

Solvency 2 rules

New european rules Solvency 2: insurance companies have to better take into account dependencies in order to compute their solvency margin.

Standard formula vs Internal models.

Once the main risk drivers for the overall company have been identified and the global solvency capital requierement has been computed, it is necessary to split it into marginal capitals for each line of business, avoid as far as possible that some lines of business become insolvent too often.

$\Rightarrow$ Minimize a risk indicator.

Ruin probability

Consider $Y_j^k = \sum_{p=1}^{j} X_p^k$ the total gain of the kth branch on j periods and $R_j^k = u_k + Y_j^k$. The kth branch is insolvent at time j if $R_j^k < 0$.

- Ruin probability widely studied in dimension 1 both for continuous and discrete models (Cramer, Lundberg, Gerber, De Finetti, ...):

$$\Psi(u) = \mathbb{P}(\exists j = 1, \ldots, n \mid \sum_{p=1}^{j} X_p + u < 0).$$

Ruin probability

Consider $Y_j^k = \sum_{p=1}^{j} X_p^k$ the total gain of the kth branch on j periods and $R_j^k = u_k + Y_j^k$. The kth branch is insolvent at time j if $R_j^k < 0$.

- Ruin probability widely studied in dimension 1 both for continuous and discrete models (Cramer, Lundberg, Gerber, De Finetti, ...):

$$\Psi(u) = \mathbb{P}(\exists j = 1, \ldots, n \mid \sum_{p=1}^{j} X_p + u < 0).$$

- In higher dimension (d), we may consider the probability that one branch fails:

$$\Psi(u_1, \ldots, u_d) = \mathbb{P}(\exists k = 1, \ldots, d;\ \exists j = 1, \ldots, n \mid R_j^k < 0),$$

not easy to manipulate (not convex).

Multivariate risk indicators (1)

S. Loisel (2004) introduced the two following risk indicators (continuous time:

- Ruin cost:

$$A(u_1, \dots, u_d) = -\sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} R_p^k \mathbb{1}_{\{R_p^k < 0\}} \right),$$

Does not take into account the dependence structure.

Multivariate risk indicators (1)

S. Loisel (2004) introduced the two following risk indicators (continuous time:

- Ruin cost:

$$A(u_1, \dots, u_d) = -\sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} R_p^k \mathbb{1}_{\{R_p^k<0\}}\right),$$

  Does not take into account the dependence structure.

- The following indicator takes into account the dependence structure:

$$B(u_1, \dots, u_d) = \sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} \mathbb{1}_{\{R_p^k<0\}} \mathbb{1}_{\{\sum_{\ell=1}^{d} R_p^\ell>0\}}\right).$$

  not convex and does not take into account the ruin cost.

Multivariate risk indicators (2)

We consider:

$$I(u_1,\dots,u_d)=\sum_{k=1}^{d}\mathbb{E}\left(\sum_{p=1}^{n} g_k(R_p^k)\mathbb{1}_{\{R_p^k<0\}}\mathbb{1}_{\{\sum_{\ell=1}^{d} R_p^\ell>0\}}\right).$$

$g_k : \mathbb{R} \to \mathbb{R}$ is a C^1 convex function, with $g_k(0)=0$, $g_k(x)\leq 0$ for $x\geq 0$ and $g_k(x)\geq 0$ for $x\leq 0$, $k=1,\dots,d$ it is a penalty function to the kth branch when it becomes insolvable.

Multivariate risk indicators (2)

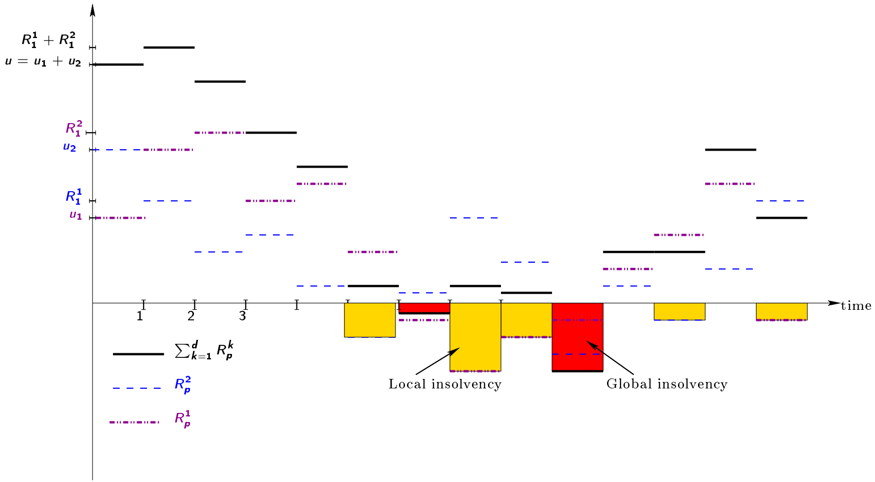
We consider:

$$I(u_1, \dots, u_d) = \sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} g_k(R_p^k) \mathbb{1}_{\{R_p^k < 0\}} \mathbb{1}_{\{\sum_{\ell=1}^{d} R_p^\ell > 0\}} \right).$$

$g_k : \mathbb{R} \to \mathbb{R}$ is a C^1 convex function, with $g_k(0) = 0$, $g_k(x) \leq 0$ for $x \geq 0$ and $g_k(x) \geq 0$ for $x \leq 0$, $k = 1, \dots, d$ it is a penalty function to the kth branch when it becomes insolvable.

Remark: $g_k(x) = -x$ is a possible choice (then we consider the ruin amount).

I is the expected sum of penalties that each line of business would have to pay due to its temporary potential insolvency (orange area).



The red area has been studied by R. Biard, S. Loisel, C. Macci and N. Veraverbeke (asymptotic properties as $u \to \infty$) for a continuous model.

Minimization of I

- Problem: find the minimum $u^\star \in \mathbb{R}_+^d$ with constraint $v_1 + \cdots + v_d = u$:

$$I(u^\star) = \inf_{v_1+\ldots+v_d=u} I(v), \;\; v \in \mathbb{R}_+^d.$$

- Tool: a Kiefer-Wolfowitz version of the stochastic mirror algorithm.
- Avantages of the stochastic algorithms approach:
  - no parametric hypothesis on the law of the X_i's,
  - dependence allowed over one period,
  - high dimension (d) allowed.

Convexity of I

In what follows, it is assumed that for all i, k, (Y_i^k, S_i) admits a density, $S_i = \sum_{\ell=1}^{d} Y_i^\ell$: cumulated gain for the ith period.

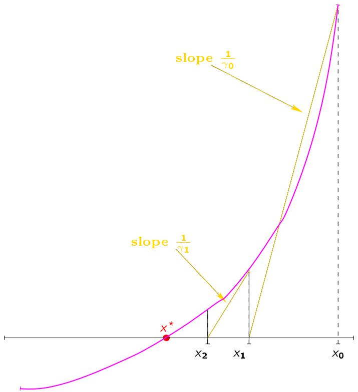
Property

- *g_k are differentiable and convex,*
- *$g_k : \mathbb{R} \to \mathbb{R}$ with $g_k(0) = 0$, $g_k(x) \geq 0$ for $x < 0$, $g_k(x) \leq 0$ for $x > 0$.*

Then the risk indicator I is convex on the convex set $\mathcal{U}_u = \{(v_1, \ldots, v_d) \in (\mathbb{R}^+)^d \ / \ v_1 + \cdots + v_d = u\}$.

A deterministic version of the Robbins-Monro algorithm

$$x_{n+1} = x_n - \gamma_n f(x_n).$$



Robbins-Monro algorithm

Often, we have access only to a perturbation of f. $\Rightarrow$ Robbins-Monro type algorithms:

$$\chi_{n+1} = \chi_n - \gamma_{n+1}\xi_{n+1},$$

where $\xi_{n+1} = f(\chi_n) + \varepsilon_{n+1}$, $(\varepsilon_n)_n$ is an centred i.i.d. sequence, with ε_{n+1} independent of $\sigma(\chi_0, \ldots, \chi_n) = \mathcal{F}_n$.
Wide literature on this algorithm and its variants: from Robbins-Monro (1951), Duflo, Küchner and Yin, ...

$\Rightarrow$ convergence of the algorithm a.e. with TCL ... under various hypothesis.

Kiefer-Wolfowitz algorithm

- Aim: Minimizing a (strictly) convex C^1 function $f \Longrightarrow$ zero of ∇f.
- ∇f is usually unknown
- Example $f(x) = \mathbb{E}(F(x, \xi))$
- ∇F is approximated by

$$D_{c_n} = \frac{F(\chi_n + c_n, \xi_n^1) - F(\chi_n - c_n, \xi_n^2)}{2c_n},$$

  with (ξ_n^1) and (ξ_n^2) two independent i.i.d. sequences of random variables of law ξ. D_{c_n} is seen as a perturbation of ∇f.
- Kifer Wolfowitz algorithm: consider

$$\chi_{n+1} = \chi_n - \gamma_n \frac{F(\chi_n + c_n, \xi_n^1) - F(X_n - c_n, \xi_n^2)}{2c_n}.$$

  Under standard conditions, the algorithm converges a.e. + TCL ...

Why can't we use K-W algorithm ?

Linear constraint : $u_1 + \cdots + u_d = u \Rightarrow$

Lagrange multipliers (= affine part)

$\Longrightarrow$ bad convergence properties.

Mirror algorithm

- Deterministic mirror descent algorithm introduced by Nemirovski and Yudin (1983)
- Stochastic version of this algorithm proposed by C. Tauvel (2008) in her thesis.
- Optimization problem: f a C^1 convex function,

$$\min_{x \in C} f(x), \quad C \text{ is a compact and convex set of } E.$$

- Observations: we have a noisy observation of ∇f

$$\psi(x) = \nabla f(x) + \eta(x),$$

with a martingale difference hypothesis for η.

Auxiliary functions for the mirror descent algorithm

- Let us fix
  - x_0 an initial point,
  - δ a strongly convex function on C, which is differentiable on x_0

  Construct an auxiliary function V:

  $$V(x) = \delta(x) - \delta(x_0) - \langle x_0 - x, \delta'(x_0) \rangle,$$

  which will be used to *push* the trajectory into C. W_β denotes the Fenchel-Legendre transform of βV: for $z \in E^\star$

  $$W_\beta(z) = \sup_{x \in C} \left\{ \langle z, x \rangle - \beta V(x) \right\}.$$

- $(\gamma_n)_n$ and $(\beta_n)_n$ are two positive sequences, $(\beta_n)_n$ is non decreasing.
- Important property of W_β: $\nabla W_\beta(\xi) \in C$.

Description of the stochastic mirror algorithm

The algorithm constructs two random sequences:

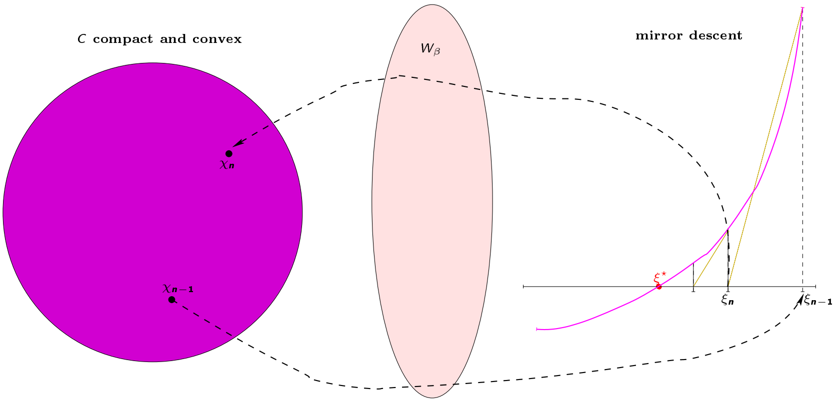
- (χ_n) in C
- (ξ_n) in the dual space E^*

Algorithm

- *Initialisation:* $\xi_0 = 0 \in E^*$, $\chi_0 \in C$
- *Update: for* $n = 1$ *to* N *do*
  - $\xi_n = \xi_{n-1} - \gamma_n \psi(\chi_{n-1})$
  - $\chi_n = \nabla W_{\beta_n}(\xi_n)$
- *Ouput:*

$$S^N = \frac{\sum_{n=1}^{N} \gamma_n \chi_{n-1}}{\sum_{n=1}^{N} \gamma_n}$$

Mirror descent algorithm



Convergence of the stochastic miror algorithm

Under the hypothesis

1. martingale difference hypothesis:

$$\mathbb{E}[\eta(\chi_{n+1})|\mathcal{F}_n] = 0, \quad \text{with} \quad \mathcal{F}_n = \sigma(\chi_0, \ldots, \chi_n).$$

   where $\psi(x) = \nabla f(x) + \eta(x)$ is observable,
2. moment of order 2: $\exists \sigma > 0$ such that for all n,

$$\mathbb{E}(\|\eta(\chi_{n+1})\|^2|\mathcal{F}_n) < \sigma^2$$

3. f admits a unique minimum $x^\star$.

C. Tauvel proved the convergence to 0 of $\mathbb{E}(f(S^N)) - \mathbb{E}(f(x^\star))$.
With an hypothesis of exponential moment on η, see gets the a.e. convergence of $f(S^N) - f(x^\star)$ to 0.

The risk indicator I

Recall that we consider

$$I(u_1, \dots, u_d) = \sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} g_k(R_p^k) \mathbb{1}_{\{R_p^k < 0\}} \mathbb{1}_{\{\sum_{\ell=1}^{d} R_p^\ell > 0\}} \right).$$

We are looking for the minimum $u^\star \in \mathbb{R}_+^d$ under the constraint $v_1 + \cdots + v_d = u$:

$$I(u^\star) = \inf_{v_1 + \ldots + v_d = u} I(v), \;\; v \in \mathbb{R}_+^d.$$

We shall use a Kiefer-Wolfowitz version of the mirror descent algorithm.

The mirror descent algorithm

- The set of constraint is
  $\mathcal{U}_u = \{v \in \mathbb{R}^d \ / \ v_i \geq 0 \ , \ v_1 + \ldots + v_d = u\}$.
- A possible choice for V is the entropy function

$$V(x) = \ln d + \sum_{i=1}^{d} \frac{x_i}{u} \ln\left(\frac{x_i}{u}\right) = \ln d + \delta(x).$$

  which is a strongly convex function.
- The Legendre-Fenchel transform may be computed easily:

$$W_\beta(\xi) = \beta \ln\left(\frac{1}{d}\sum_{i=1}^{d} \exp\left[\xi_i \frac{u}{\beta}\right]\right).$$

Approximate gradient

$$I(u_1, \ldots, u_d) = \mathbb{E}\left(\mathcal{I}(u_1, \ldots, u_d, \mathcal{Y})\right) \quad \text{where} \quad \mathcal{Y} = \begin{pmatrix} Y_1^1 & \cdots & Y_n^1 \\ \cdots & \cdots & \cdots \\ Y_1^d & \cdots & Y_n^d \end{pmatrix}.$$

Denote

$$\begin{aligned} \mathcal{I}^k\left(c_i^+, \mathcal{Y}\right) &= \mathcal{I}(\chi_{i-1}^1, \ldots, \chi_{i-1}^{k-1}, \chi_{i-1}^k + c_i, \chi_{i-1}^{k+1}, \ldots, \chi_{i-1}^d, \mathcal{Y}), \\ \mathcal{I}^k\left(c_i^-, \mathcal{Y}\right) &= \mathcal{I}(\chi_{i-1}^1, \ldots, \chi_{i-1}^{k-1}, \chi_{i-1}^k - c_i, \chi_{i-1}^{k+1}, \ldots, \chi_{i-1}^d, \mathcal{Y}), \end{aligned}$$

Consider the random vector $D_{c_i}\mathcal{I}$ whose kth coordinate $D_{c_i}^k\mathcal{I}(u_1, \ldots, u_d, \mathcal{Y})$ is

$$\frac{\mathcal{I}^k\left(c_i^+, \mathcal{Y}\right) - \mathcal{I}^k\left(c_i^-, \mathcal{Y}\right)}{2c_i}.$$

Our algorithm

Algorithm

Initialisation:
$$\begin{cases} \xi^0 = 0 \in (\mathbb{R}^m)^\star \\ \chi^0 \in C \end{cases}$$

Update: for $i = 1, \dots, N$ do

$$\begin{cases} \xi_i = \xi_{i-1} - \gamma_i \Psi_{c_i}(\chi_{i-1}, \mathcal{Y}_i) \\ \chi_i = \nabla W_{\beta_i}(\xi_i) \end{cases}$$

Output:
$$S^N = \frac{\sum_{i=1}^N \gamma_i \chi_{i-1}}{\sum_{i=1}^N \gamma_i}$$

with $\Psi_{c_i}(\chi_{i-1}, \mathcal{Y}_i) = D_{c_i}\mathcal{I}(\chi_{i-1}, \mathcal{Y}_i)$ and $\mathcal{Y}_i$ independent copies of $\mathcal{Y}$.

Conditions for the convergence (1)

Condition (on I)

1. *I is a convex function from $\mathbb{R}^d$ to $\mathbb{R}$,*
2. *I is C^2 on $\mathcal{U}_u$,*
3. *I has a unique minimum $u^\star$ in $\mathcal{U}_u$,*
4. *$\exists \sigma > 0$ such that for all $v \in \mathcal{U}_u$*

$$\mathbb{E}\left(\mathcal{I}(v_1, \ldots, v_d, \mathcal{Y})^2 | \mathcal{F}_{i-1}\right) \leq \sigma^2.$$

$\mathcal{F}_{i-1}$ is generated by $(\mathcal{Y}^0, \ldots, \mathcal{Y}^{i-1})$.

Conditions for the convergence (2)

Condition (on the sequences)

Let $(\beta_n)_{n\geq 0}$, $(\gamma_n)_{n\geq 0}$ and $(c_n)_{n\geq 0}$ be positive sequences, $(\beta_n)_{n\geq 0}$ is non decreasing and:

1. $\beta_N / \sum_{i=1}^{N} \gamma_i \xrightarrow[N\to+\infty]{} 0,$
2. $\sum_{i=1}^{N} \gamma_i c_i / \sum_{i=1}^{N} \gamma_i \xrightarrow[N\to+\infty]{} 0,$
3. $\sum_{i=1}^{N} \frac{\gamma_i^2}{c_i^2 \beta_{i-1}} / \sum_{i=1}^{N} \gamma_i \xrightarrow[N\to+\infty]{} 0,$
4. $\sum_{i=1}^{+\infty} \left(\frac{\gamma_i}{c_i}\right)^2 < \infty.$

Result

Theorem

With the above conditions,

$$S_N \xrightarrow{L^1} x^\star.$$

With an hypothesis of moment of order > 2 *on* $\mathcal{I}$,

$$S_N \xrightarrow{a.s.} x^\star.$$

Idea of the proof (1)

Based on the decomposition

$$\Psi_{c_n}(\chi_{n-1}, \mathcal{Y}_n) = \nabla I(\chi_{n-1}) + \eta_{c_n}(\chi_{n-1}, \mathcal{Y}_n) + r_{c_n}(\chi_{n-1}).$$

The condition of C. Tauvel fails because of the $r_{c_n}(\chi_{n-1})$ term.
$\eta_{c_n}(\chi_{n-1}, \mathcal{Y}_n)$ is a martingale difference.

$$\begin{aligned} \eta_{c_n}(\chi_{n-1}, \mathcal{Y}_n) &= D_{c_n}\mathcal{I}(\chi_{n-1}, \mathcal{Y}_n) - D_{c_n}I(\chi_{n-1}), \\ r_{c_n}(\chi_{n-1}) &= D_{c_n}I(\chi_{n-1}) - \nabla I(\chi_{n-1}). \end{aligned}$$

and

- Law of large numbers for martingales
- Chow Lemma for martingales

Idea of the proof (2)

Let $\varepsilon_N = I(S^N) - I(x^*) \geq 0$

Lemma

$$\left(\sum_{i=1}^{N} \gamma_i\right) \varepsilon_N \leq \beta_N V(x^\star) - \sum_{i=1}^{N} \gamma_i \left\langle \eta^{c_i}(\chi_{i-1}, \mathcal{Y}^i), \chi^{i-1} - x^\star \right\rangle$$

$$- \sum_{i=1}^{N} \gamma_i \left\langle r^{c_i}(\chi^{i-1}), \chi^{i-1} - x^\star \right\rangle$$

$$+ \sum_{i=1}^{N} \frac{\gamma_i^2}{2\alpha\beta_{i-1}} \left\| \Psi^{c_i}(\chi^{i-1}, \mathcal{Y}^i) \right\|_\star^2$$

Idea of the proof (3)

$$\sum_{i=1}^{N} \gamma_i \left\langle \eta^{c_i}(\chi_{i-1}, \mathcal{Y}^i), \chi^{i-1} - x^\star \right\rangle$$ is a martingale

$$\sum_{i=1}^{N} \gamma_i \left\langle r^{c_i}(\chi^{i-1}), \chi^{i-1} - x^\star \right\rangle$$ controlled by the fact that $r^{c_i}(\chi^{i-1})$ goes to 0 a.e.

$$\sum_{i=1}^{N} \frac{\gamma_i^2}{2\alpha\beta_{i-1}} \left\| \Psi^{c_i}(\chi^{i-1}, \mathcal{Y}^i) \right\|_\star^2$$ has bounded expectation and controlled a.s. by using the Chow Lemma.

Hypothesis on I

1. Unicity of $u^\star$: it is sufficient that for some k and all $u < w_k < v_k$,

$$-\mathbb{E}\left[g_k(Y_p^k + v_k)\mathbf{1}_{\{\sum_{k=1}^d Y_p^k > -u\}} \cdot \mathbf{1}_{\{-v_k < Y_p^k < -w_k\}}\right] < 0.$$

   This is satisfied if for some k, $f_{S_p, Y_p^k} > 0$ on $]-u, \infty[\times]-\infty, -u[$.

2. In the case $g_k(x) = -x$, the moment condition for $\mathcal{I}$ is equivalent to a moment condition (of the same order) on X_n.

Simulations

Some simulations for some simple models. We considered

- Normal laws,
- $n = 1$ observation of several periods of length 1, so that $X_p^k = Y_p^k$,
- $X_p \in \mathbb{R}^d$ are independent and identically distributed random vectors. Some dependencies on the coordinates of X_p are allowed.

Parameters

The algorithm has been performed with the following sequences $(\gamma_n)_{n\in\mathbb{N}^\star}$, $(c_n)_{n\in\mathbb{N}^\star}$ and $(\beta_n)_{n\in\mathbb{N}^\star}$:

- $\gamma_n = \dfrac{1}{(n+1)^\alpha}$ with $\alpha = \frac{3}{4} + \frac{1}{10}$,
- $c_n = \dfrac{1}{(n+1)^\delta}$ with $\delta = \frac{1}{4}$,
- $\beta_n = 1$.

Also, we have chosen $u = 2$ and the initialization of the algorithm (χ^0) is done at random uniformly in the simplex $\mathcal{U}_u$.

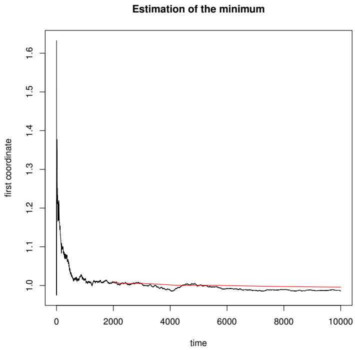
No dependence between the coordinates

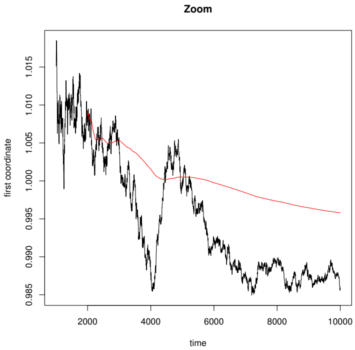
$d = 2$, $n = 1$, X_i^1 and X_i^2 are independent and the vectors X_i are also independent.

- first X_i^1 and X_i^2 have the same normal laws,
- then we consider different normal laws.

No dependence between the coordinates

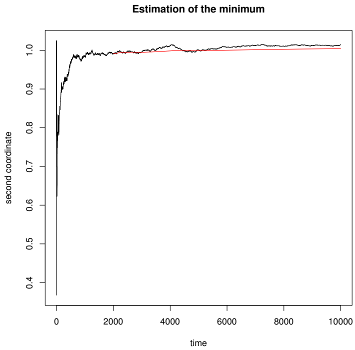
Same normal laws: $X_i^d \rightsquigarrow \mathcal{N}(0.3, 1)$. For $N = 10000$ independent simulations we obtain : $S^N = (0.996; 1.004)$

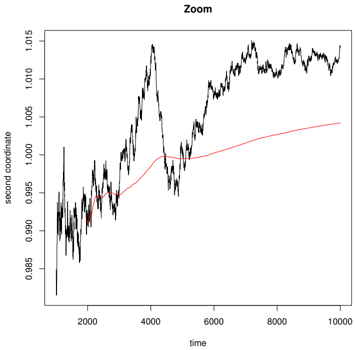




No dependence between the coordinates

Same normal laws: $X_i^d \rightsquigarrow \mathcal{N}(0.3, 1)$. For $N = 10000$ independent simulations we obtain : $S^N = (0.996; 1.004)$





No dependence between the coordinates

Same normal laws: For 50 simulations of length $N = 1000$, with the same parameters as above. The table below gives for each of the two coordinates, the mean of the estimation $(\widehat{u_1}, \widehat{u_2})$ of the minimum (u_1, u_2) and the standard error.

	first coord.	second coord.
mean	1.01	0.99
standard error	0.04	0.04

No dependence between the coordinates

Different normal laws: 50 simulations of length $N = 1000$, with two independent normal laws. $X_p^1 \rightsquigarrow \mathcal{N}(0.3, 1)$ and $X_p^2 \rightsquigarrow \mathcal{N}(0.8, 1)$. The table below gives for each of the two coordinates, the mean of the estimation $(\widehat{u}_1, \widehat{u}_2)$ of the minimum (u_1, u_2) and the standard error.

	first coord.	second coord.
mean	1.226	0.774
standard error	0.051	0.051

No dependence between the coordinates

Different normal laws: 50 simulations of length $N = 1000$, with two independent normal laws. $X_p^1 \rightsquigarrow \mathcal{N}(0.3, 1)$ and $X_p^2 \rightsquigarrow \mathcal{N}(0.3, 4)$. The table below gives for each of the two coordinates, the mean of the estimation $(\widehat{u}_1, \widehat{u}_2)$ of the minimum (u_1, u_2) and the standard error.

	first coord.	second coord.
mean	0.787	1.213
standard error	0.067	0.067

Some dependence between the coordinates

Two simple models of vectorial dependence:

- a comonotonic example: random vectors in $\mathbb{R}^3$ with $X_p^1 \stackrel{\mathcal{L}}{=} X_p^2 \rightsquigarrow \mathcal{N}(0.3, 1)$ and $X_p^3 = 2X_2$,
- gaussian vectors.

Some dependence between the coordinates

Comonotonic example. 50 simulations of length $N = 1000$ of this dimension 3 model. As before, we consider $n = 1$ (the periods are of length 1).
The table below gives for each of the two coordinates, the mean of the estimation $(\widehat{u}_1, \widehat{u}_2, \widehat{u}_3)$ of the minimum (u_1, u_2, u_3) and the standard error.

	first coord.	second coord.	third coord.
mean	0.8	0.43	0.77
standard error	0.06	0.02	0.05

In this model, X_2 and X_3 fail simultaneously and when they fail, X_3 is 2 two times worse than X_2.

Some dependence between the coordinates

Gaussian vectors. We have also performed simulations for a Gaussian vector in $\mathbb{R}^3$ with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.9 \\ 0 & 0.9 & 1 \end{pmatrix}$$

and expectation $m = (0.3, 0.3, 0.3)$. As above, we have performed 50 simulations of length $N = 1000$, of the Gaussian vector X.

Some dependence between the coordinates

The table below gives for each of the three coordinates, the mean of the estimation $(\widehat{u_1}, \widehat{u_2}, \widehat{u_3})$ of the minimum (u_1, u_2, u_3) and the standard error.

	first coord.	second coord.	third coord.
mean	0.785	0.604	0.612
standard error	0.045	0.03	0.028

A example with some temporal dependence

The first two coordinates are independant $AR(1)$ with the same law: $X_i = 0.4X_{i-1} + \varepsilon_i$ with $(\varepsilon_i)_{i\in\mathbb{N}}$ gaussian white noise ($\mathcal{N}(0,1)$).
The third coordinate is 2 times the second one.
We have simulated 5 times 500 independent periodsof length $n = 5$.

	first coord.	second coord.	third coord.
mean	0.85	0.39	0.76
standard error	0.035	0.012	0.025

To do ...

- A deeper study involving other laws (with heavy tail e.g.), more realistic models and temporal dependencies.
- Asymptotic theory for $u \longrightarrow \infty$.
- Some dependence on the $\mathcal{Y}_i$ (using Benveniste, Métivier, Priouret results or methods for example).

Thanks for your attention

Recall the Chow Lemma

Theorem

Suppose $(a_N)_{N\in\mathbb{N}}$ *is a bounded sequence of positive numbers ,* *suppose that* $1 < p \leq 2$. *For* $N \in \mathbb{N}$, *let* $A_N = 1 + \sum_{k=0}^{N} a_k$ *and* $A_\infty = \lim_{N\to\infty} A_N$. *Suppose that* $(Z_N)_{N\in\mathbb{N}}$ *is a positive sequence of random variables adapted to* $\mathcal{F}_N$ *and* K *is a constant such that*

$$\mathbb{E}(Z_{N+1}|\mathcal{F}_N) \leq K \text{ and } \sup_N \mathbb{E}(Z_{N+1}^p|\mathcal{F}_N) < \infty$$

then we have the following properties almost surely :

$$\text{on } \{A_\infty < \infty\} \quad \sum_{k=1}^{\infty} A_k Z_{k+1} \text{ converges}$$

Recall the Chow Lemma

Theorem

Suppose $(a_N)_{N\in\mathbb{N}}$ *is a bounded sequence of positive numbers , suppose that* $1 < p \leq 2$. *For* $N \in \mathbb{N}$, *let* $A_N = 1 + \sum_{k=0}^{N} a_k$ *and* $A_\infty = \lim_{N\to\infty} A_N$. *Suppose that* $(Z_N)_{N\in\mathbb{N}}$ *is a positive sequence of random variables adapted to* $\mathcal{F}_N$ *and* K *is a constant such that*

$$\mathbb{E}(Z_{N+1}|\mathcal{F}_N) \leq K \text{ and } \sup_N \mathbb{E}(Z_{N+1}^p|\mathcal{F}_N) < \infty$$

then we have the following properties almost surely :

$$\text{on } \{A_\infty = \infty\} \;\; A_{n-1}^{-1} \sum_{k=1}^{\infty} A_k Z_{k+1} \leq K.$$

Recall the Chow Lemma

Theorem

Suppose $(a_N)_{N\in\mathbb{N}}$ *is a bounded sequence of positive numbers , suppose that* $1 < p \leq 2$. *For* $N \in \mathbb{N}$, *let* $A_N = 1 + \sum_{k=0}^{N} a_k$ *and* $A_\infty = \lim_{N\to\infty} A_N$. *Suppose that* $(Z_N)_{N\in\mathbb{N}}$ *is a positive sequence of random variables adapted to* $\mathcal{F}_N$ *and* K *is a constant such that*

$$\mathbb{E}(Z_{N+1}|\mathcal{F}_N) \leq K \text{ and } \sup_N \mathbb{E}(Z_{N+1}^p|\mathcal{F}_N) < \infty$$

then we have the following properties almost surely :

Back to idea of proof.