HAL open science

Some multivariate risk indicators; minimization by using a Kiefer-Wolfowitz approach to the mirror stochastic algorithm.

Peggy Cenac, Véronique Maume-Deschamps, Clémentine Prieur

To cite this version:

Peggy Cenac, Véronique Maume-Deschamps, Clémentine Prieur. Some multivariate risk indicators; minimization by using a Kiefer-Wolfowitz approach to the mirror stochastic algorithm.. 2010. hal-00484233v1

HAL Id: hal-00484233

https://hal.science/hal-00484233v1

Preprint submitted on 19 May 2010 (v1), last revised 5 Dec 2011 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SOME MULTIVARIATE RISK INDICATORS ; MINIMIZATION BY USING A KIEFER-WOLFOWITZ APPROACH TO THE MIRROR STOCHASTIC ALGORITHM.

P. CÉNAC, V. MAUME-DESCHAMPS, AND C. PRIEUR

Abstract. We consider some risk indicators of vectorial risk processes. These indicators take into account the dependencies between business lines as well as some temporal dependencies. By using stochastic algorithms, we may estimate the minimum of these risk indicators, under a fixed total capital constraint. This minimization may apply to optimal reserve allocation.

1. Introduction

The new regulation rules for insurance industry, namely Solvency 2 in Europe, lead companies to adjust their solvency margins to the underlying risks. Once the overall company capital requirement has been computed, it must be splited into solvency capital for each line of business. In other words, given an initial capital u we assume that it is allocated to each line of business: u_k is the initial capital of the kth line of business, then $u_1 + \ldots + u_d = u$. We aim to optimize the capital allocation with respect to some risk indicator. In this context u is fixed and we search for an optimal choice of the u_k's. The ruin probability is a quite standard risk indicator and has been widely studied in dimension 1 (see [5] for a review). Other risk indicators have been introduced by S. Loisel [10] in a multidimensional setting. These risk indicators either do not take the dependence structure into account or are not convex so that convex analysis cannot be used for minimization. We introduce some new risk indicators, both involving dependence and having nice convexity properties. Considering that the main risk drivers for the overall company have been identified and that the global solvency capital requirement has been computed, they reveal the marginal solvency capitals for each line of business. A way to avoid as far as possible that some lines of business become insolvent too often could be to minimize these risk indicators, under a fixed total capital constraint. This might be achieved if some capital fungibility between lines of business or between entities is possible. One possible way to define optimality of the global reserve allocation is to minimize the expected sum of the penalties that each line of business would have to pay due to its temporary potential insolvency.

Key words and phrases. Multivariate risk processes, risk indicators, stochastic algorithms, optimal allocation.

This work has been partially supported by the French Research National Agency (ANR) under the reference ANR-08-BLAN-0314-01.

We consider a vectorial risk process

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

where X_i^k corresponds to the gains of the kth business line during the ith period. That is, $X_i^k = G_i^k - L_i^k$ where G_i^k denotes the incomes and L_i^k denotes the losses. We are interested in the cumulative gain : $Y_i^k = \sum_{p=1}^{i} X_p^k$.
As a simple example, it might be considered that $G_i^k = c_k$: the income is constant on each period and thus, $Y_i^k = ic_k - S_i^k$ where S_i^k is the cumulative losses.
The ruin probability in a multivariate setting may be defined as the probability that one of the d lines of business makes default during the n first periods :

$$R(u_1, \dots, u_d) = \mathbb{P}(\exists k = 1, \dots, d \ \exists i = 1, \dots, n \ / \ Y_i^k + u_k < 0).$$

In [3] some bounds for ruin probabilities in multivariate compound risk models are provided.
Let $R_p^k = u_k + Y_p^k$. We may also be interested in the total cost of ruin :

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

The asymptotic (as $u \to \infty$) properties of A have been studied in [2]. In the case $d = 1$, they study the asymptotics of $A(u)$. In higher dimension, they provide the asymptotic behavior of the vector $u^\star \in \mathbb{R}^d$ realizing the minimum of A under the constraint $u_1 + \cdots + u_d = u$.
The risk indicator A doesn't take into account the dependence structure between the lines of business. A multivariate risk indicator that takes into account the dependence structure may be :

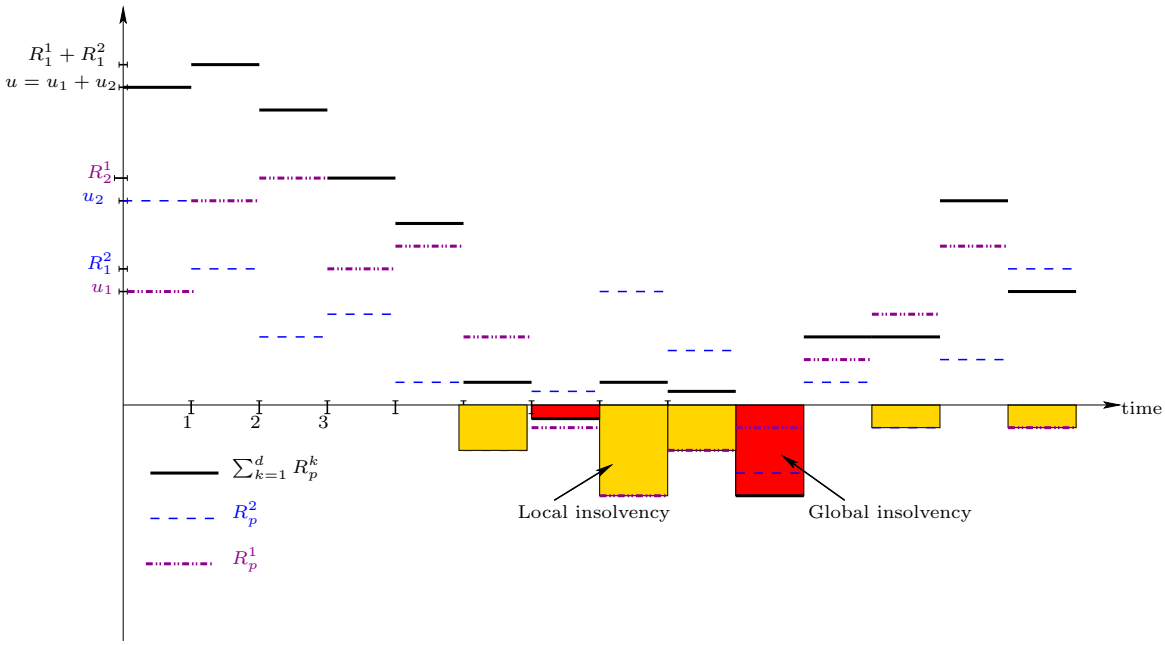
$$B(u_1, \dots, u_d) = \sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} \mathbf{1}_{\{R_p^k<0\}} \mathbf{1}_{\{\sum_{k=1}^{d} R_p^k>0\}}\right).$$

This risk indicator gives an indication on some average time to ruin, it has been introduced in [10]. The risk indicator B, as the ruin probability, is not in general convex. In our approach, we need the convexity property so we introduce the following risk indicator, related to the ruin cost. Given a differentiable and convex function $g_k : \mathbb{R} \to \mathbb{R}$ satisfying $g_k(x) \geq 0$ for $x \leq 0$, $k = 1, \dots, d$, we consider the risk indicator

$$I(u_1, \dots, u_d) = \sum_{k=1}^{d} \mathbb{E}\left(\sum_{p=1}^{n} g_k(R_p^k) \mathbf{1}_{\{R_p^k<0\}} \mathbf{1}_{\{\sum_{k=1}^{d} R_p^k>0\}}\right). \tag{1.1}$$

The function g_k represents the cost that the kth business branch has to pay if it becomes insolvent. In the case $g_k(x) = -x$, the risk indicator I is a generalization of B that takes into account the cost of insolvency. The

following picture represents the risk indicator (light area) for $g_k(x) = -x$ and $d = 2$.



We aim to find a minimum under the constraint $v_1 + \cdots + v_d = u$. Formally, we are looking for $u^\star \in (\mathbb{R}_+)^d$ such that

$$I(u^\star) = \inf_{v_1+\ldots+v_d=u} I(v), \ v \in (\mathbb{R}_+)^d. \tag{1.2}$$

Unless for very specific examples, no explicit solution is available. We propose to solve this problem by using stochastic algorithms (see [4, 1, 6] for a review). Also, the gradient of the function I is not explicitly computable. We are thus led to use approximations of ∇I in stochastic algorithms. Because we have to find a minimum under an affine constraint, it would seem reasonable to use a stochastic algorithm including the Lagrange multipliers. It is known that in this case the algorithm is parameter-sensitive. Indeed, we observed this phenomena in practice. We shall use instead an algorithm introduced in the deterministic form by Nemirovski and Yudin ([9]) and by Tauvel ([11]) in its stochastic form. In her thesis, Tauvel considers that an approximation of the gradient of I by a martingale difference is available, which is not our case. Thus we use a Kiefer-Wolfowitz approach of her algorithm : we approximate the gradient *online*. We construct a stochastic algorithm to estimate $u^\star$, the *optimal allocation*. Under a moment condition (order > 2) we prove the almost sure convergence of the estimator.

Notice that our approach do not require any hypothesis, apart from a moment assumption, on the distribution of X_i and the existence of a density. Also there may be temporal dependence on a period of length n. Nevertheless, in order to achieve the algorithm, we shall need N independent copies of the distribution of X_i on a period of length n.

The paper is organized as follows: in section 2 we prove some convexity property on I. In section 3 we make explicit the algorithm. In section 4 we prove the convergence results. In section 5 we show how to apply these results for our risk indicator and provide some simulations.

2. Convexity property

Let us first prove that, under suitable assumptions on g_k, the risk indicator I is differentiable and convex. Notice that the knowledge of the whole joint distribution $R_p^k,\ p=1,\dots,n,\ k=1,\dots,d$, which should be required to prove the differentiability of the ruin probability, is not needed here.

Let $v^{(-k)}=\sum_{j\neq k} v_j$ and let $S_p=\sum_{k=1}^{d} Y_p^k$. In the following, we assume that (Y_p^i,S_p) has a joint density distribution denoted by $f_{Y_p^i,S_p}$.

Proposition 2.1. *Assume that the functions g_k are differentiable and convex, that $g_k : \mathbb{R} \to \mathbb{R}$ with $g_k(0)=0$, $g_k(x)\geq 0$ for $x<0$ and $g_k(x)\leq 0$ for $x>0$. Then the risk indicator I is convex on the convex set $\mathcal{U}_u=\{(v_1,\dots,v_d)\in(\mathbb{R}^+)^d \ / \ v_1+\dots+v_d=u\}$.*

Proof. In order to prove the convexity of the risk indicator I, it is sufficient to prove the gradient inequality, that is for any $v,w\in\mathcal{U}_u$,

$$(v-w)^T\nabla I(w)\leq I(v)-I(w). \tag{2.1}$$

We have :

$$\begin{aligned}(\nabla I(v))_i &= \sum_{k=1}^{d}\sum_{p=1}^{n}\int_{-\infty}^{-v_k} g_k(y+v_k) f_{Y_p^k,S_p}(y,-u)dy \\ &\quad + \sum_{p=1}^{n}\mathbb{E}(g_i'(Y_p^i+v_i)\mathbf{1}_{\{Y_p^i<-v_i\}}\mathbf{1}_{\{S_p>-u\}}) \\ &\quad - \sum_{p=1}^{n} g_i(0)\int_{-u}^{+\infty} f_{Y_p^i,S_p}(-v_i,s)ds.\end{aligned}$$

Using that $g_i(0)=0$ and $\sum_{k=1}^d v_k=\sum_{k=1}^d w_k=u$ gives

$$(v-w)^T\nabla I(w)=\sum_{k=1}^{d}\sum_{p=1}^{n}(v_k-w_k)\mathbb{E}(g_k'(Y_p^k+w_k)\mathbf{1}_{\{Y_p^k<-w_k\}}\mathbf{1}_{\{S_p>-u\}}).$$

Since the function g_k is convex, it satisfies the gradient inequality, thus

$$(v-w)^T\nabla I(w) \leq \sum_{k=1}^{d}\sum_{p=1}^{n}\mathbb{E}\left[(g_k(Y_p^k+v_k)-g_k(Y_p^k+w_k))\mathbf{1}_{\{Y_p^k<-w_k\}}\mathbf{1}_{\{S_p>-u\}}\right]$$

The second term of this inequality is equal to

$$I(v)-I(w)+\sum_{k=1}^{d}\sum_{p=1}^{n}\mathbb{E}\left[g_k(Y_p^k+v_k)\mathbf{1}_{\{S_p>-u\}}(\mathbf{1}_{\{Y_p^k<-w_k\}}-\mathbf{1}_{\{Y_p^k<-v_k\}})\right].$$

For $k = 1, \ldots, d$, if $v_k > w_k$ then

$$\mathbf{1}_{\{Y_p^k < -w_k\}} - \mathbf{1}_{\{Y_p^k < -v_k\}} = \mathbf{1}_{\{-v_k \leq Y_p^k \leq -w_k\}}$$

and

$$\begin{aligned} &\mathbb{E}\left[g_k(Y_p^k + v_k)\mathbf{1}_{\{S_p > -u\}}(\mathbf{1}_{\{Y_p^k < -w_k\}} - \mathbf{1}_{\{Y_p^k < -v_k\}})\right] \\ &= \mathbb{E}(g_k(Y_p^k + v_k)\mathbf{1}_{\{S_p > -u\}}\mathbf{1}_{\{-v_k \leq Y_p^k < -w_k\}}) \leq 0. \end{aligned}$$

If $w_k > v_k$ then

$$\begin{aligned} &\mathbb{E}\left[g_k(Y_p^k + v_k)\mathbf{1}_{\{S_p > -u\}}(\mathbf{1}_{\{Y_p^k < -w_k\}} - \mathbf{1}_{\{Y_p^k < -v_k\}})\right] \\ &= -\mathbb{E}(g_k(Y_p^k + v_k)\mathbf{1}_{\{S_p > -u\}}\mathbf{1}_{\{-w_k \leq Y_p^k < -v_k\}}) \leq 0. \end{aligned}$$

We conclude that the gradient inequality is satisfied which concludes the proof of the proposition. □

The fact that I is a convex function is crucial to ensure the convergence of the stochastic algorithm. The convexity assumption is also standard for classical stochastic algorithms (as Robbins-Monro or Kiefer-Wolfowitz).

3. Estimation by use of stochastic algorithms

In this section we present our algorithm which is a Kiefer-Wolfowitz version of the mirror algorithm. In section 4 we shall prove its convergence under general assumptions (see Assumption 1) that are satisfied in particular for our specific risk indicator I.

3.1. Auxiliary function for the mirror algorithm.

The stochastic mirror algorithm introduced by Tauvel requires the use of an auxiliary function that will be used to push the trajectory into the set of constraints. We consider that $\mathbb{R}^d$ is endowed with the L^1 norm $\| \ \|$ and the dual space $(\mathbb{R}^d)^\star$ is endowed with the dual norm $\| \ \|_\star$:

$$\|x\| = \sum_{i=1}^{d} |x_i|, \quad \|\xi\|_\star = \sup_{i=1,\ldots,d} |\xi_i|.$$

Let C be a compact convex subset of $\mathbb{R}^d$. Recall that a convex function $\delta \ : C \longrightarrow \ \mathbb{R}$ is strongly convex with parameter $\alpha > 0$ with respect to the norm $\| \ \|$ if for any $\lambda \in [0, 1]$ and any x, y belonging to C,

$$\delta(\lambda x + (1-\lambda)y) \leq \lambda\delta(x) + (1-\lambda)\delta(y) - \frac{\alpha}{2}\lambda(1-\lambda)\|x-y\|^2.$$

The mirror algorithm uses an α-convex function δ differentiable on a point $x_0 \in C$ to define the auxiliary function V :

$$V(x) = \delta(x) - \delta(x_0) - < \nabla\delta(x_0), x - x_0 > .$$

It is then easy to see that the function V is also α-convex.
Setting a positive β, W_β denotes the Legendre-Fenchel transform of βV :

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

Due to the properties of Legendre-Fenchel transform of α-convex functions, we know that W_β is continuously differentiable and that its gradient takes its values in C. Moreover, the gradient of W_β has a Lipschitz constant equal

to $(\alpha\beta)^{-1}$.
We consider a function Ψ which is an approximation of the gradient of I. The mirror algorithm uses two positive sequences β_i and γ_i and a sequence of i.i.d. random vectors $\mathcal{Y}^i$ in the following way (the choices for these sequences will be explicited a bit later).

Algorithm 1

Initialization : $\begin{cases} \xi_0 = 0 \in (\mathbb{R}^d)^\star \\ \chi_0 \in C \end{cases}$

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

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

Output : $S_N = \dfrac{\sum_{i=1}^N \gamma_i \chi_{i-1}}{\sum_{i=1}^N \gamma_i}$

Under assumptions on the function Ψ, we shall prove that S^N goes a.s. to the unique minimum $x^\star$ of I.

3.2. **Approximate gradient.** In our case, we do not have access directly to the gradient of I, ∇I, but we may approximate it *online*.
Recall that

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

d denotes the number of lines of business, and n the number of periods of interest. The problem is to solve (1.2). We will consider an approximation Ψ of ∇I with the following decomposition $\Psi = \nabla I + \eta + r$, where η is a martingale difference and r is negligible with respect to ∇I and η. Arguing as for Kiefer-Wolfowitz, we first notice that

$$I(u_1, \dots, u_d) = \mathbb{E}\left(\mathcal{I}(u_1, \dots, u_d, \mathcal{Y})\right)$$

where

$$\mathcal{Y} = \begin{pmatrix} Y_1^1 & \cdots & Y_n^1 \\ \cdots & \cdots & \cdots \\ Y_1^d & \cdots & Y_n^d \end{pmatrix}$$

and

$$\mathcal{I}(u_1, \dots, u_d, y) = \sum_{k=1}^d \sum_{p=1}^n g_k(y_p^k + u_k)\mathbf{1}_{\{y_p^k+u_k<0\}}\mathbf{1}_{\{\sum_{k=1}^d y_p^k+u_k>0\}}.$$

Let us denote a realization of $\mathcal{Y}$ by

$$y = \begin{pmatrix} y_1^1 & \cdots & y_n^1 \\ \cdots & \cdots & \cdots \\ y_1^d & \cdots & y_n^d \end{pmatrix}.$$

For sake of shortness, we introduce the following notations :

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

We then consider $D_{c_i}\mathcal{I}$ (*resp.* $D_{c_i}I$) the random vector whose kth coordinate $D_{c_i}^k\mathcal{I}(u_1,\ldots,u_d,\mathcal{Y})$ (*resp.* $D_{c_i}^kI(u_1,\ldots,u_d,\mathcal{Y})$) is defined by

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

Considering independent copies of the random matrix $\mathcal{Y}$: $\mathcal{Y}_1,\ldots,\mathcal{Y}_N$, we perform Algorithm 2 below :

Algorithm 2

Initialization : $\begin{cases}\xi_0=0\in(\mathbb{R}^m)^\star\\ \chi_0\in C\end{cases}$

Update : for $i=1,\ldots,N$ 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 : $S_N=\dfrac{\sum_{i=1}^N\gamma_i\chi_{i-1}}{\sum_{i=1}^N\gamma_i}$

with

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

Then defining

$$\begin{aligned}
\eta_{c_i}(\chi_{i-1},\mathcal{Y}_i) &= D_{c_i}\mathcal{I}(\chi_{i-1},\mathcal{Y}_i)-D_{c_i}I(\chi_{i-1}),\\
r_{c_i}(\chi_{i-1}) &= D_{c_i}I(\chi_{i-1})-\nabla I(\chi_{i-1}),
\end{aligned}$$

it enhances the decomposition

$$\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i)=\nabla I(\chi_{i-1})+\eta_{c_i}(\chi_{i-1},\mathcal{Y}_i)+r_{c_i}(\chi_{i-1}). \tag{3.1}$$

In the next section, we give conditions on γ_i, β_i and c_i for the convergence of the algorithm.

4. Convergence of the algorithm

Let us consider C is compact convex subset of $\mathbb{R}^d$. We make also the following assumptions on I :

Assumption 1.

(1) I is a convex function from $\mathbb{R}^d$ to $\mathbb{R}$,
(2) I is of C^2 class,

(3) I admits a unique minimum $x^\star$ on C,

Moreover, we have for any x in $\mathbb{R}^d$ and any $c > 0$

$$\mathbb{E}(\eta_c(x, \mathcal{Y})) = 0 \quad \text{and}$$

$$r_c(x) \text{ converges to zero as } c \text{ goes to } 0.$$

We perform **Algorithm** 2 with χ_0 being any random vector taking values in C, and $\mathcal{F}_i$ being the σ-field generated by $\chi_0, \mathcal{Y}_1, \ldots, \mathcal{Y}_i$. Then ξ_i is $\mathcal{F}_i$ measurable and $\mathcal{Y}_i$ is independent of $\mathcal{F}_{i-1}$. We get therefore

$$\mathbb{E}(\eta_{c_i}(\chi_{i-1}, \mathcal{Y}_i)|\mathcal{F}_{i-1}) = 0.$$

We shall need the following additional assumption on the function $\mathcal{I}$.

Assumption 2. There exists some non negative real number σ such that for all vector $(v_1, \ldots, v_d) \in \mathbb{R}^d$, the random variable $\mathcal{I}(x, \mathcal{Y})$ satisfies the moment condition $\forall i$

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

Remark 1. Assumption 2 yields the existence of a real number $\sigma > 0$ such that for any $i = 1, \ldots, N$,

$$\mathbb{E}\left(\|\eta_{c_i}(\chi_{i-1}, \mathcal{Y}_i)\|_*^2 |\mathcal{F}_{i-1}\right) \leq \frac{\sigma^2}{c_i^2}. \tag{4.1}$$

Indeed, one has

$$\begin{aligned}\mathbb{E}\left(\|\eta_{c_i}(\chi_{i-1}, \mathcal{Y}_i)\|_*^2 |\mathcal{F}_{i-1}\right) &= \sup_{1\leq k\leq d}\Bigg(\frac{\mathrm{Var}\left(\mathcal{I}^k(c_i^+, \mathcal{Y}i)|\mathcal{F}_{i-1}\right)}{4c_i^2} + \frac{\mathrm{Var}\left(\mathcal{I}^k(c_i^-, \mathcal{Y}_i)|\mathcal{F}_{i-1}\right)}{4c_i^2} \\ &\quad -\frac{1}{2c_i^2}\mathbb{E}\left[\left(\mathcal{I}^k(c_i^+, \mathcal{Y}_i) - I(c_i^+)\right)\left(\mathcal{I}^k(c_i^-, \mathcal{Y}_i) - I(c_i^-)\right)\right]\Bigg).\end{aligned}$$

From Assumption (2) together with the Cauchy-Schwartz inequality, one gets

$$\mathbb{E}\left(\left\|\eta_{c_i}(\chi_{i-1}, \mathcal{Y}^i)\right\|_*^2 |\mathcal{F}_{i-1}\right) \leq \frac{\sigma^2}{c_i^2},$$

which concludes the proof of (4.1). Condition (4.1) is sufficient to obtain the L^1-convergence. In order to prove the almost sure convergence, we shall need a higher order moment condition.

Assumption 3. There exists $p > 2$ such that almost surely

$$\sup_{i>0} \mathbb{E}\left(|\mathcal{I}(v_1, \ldots, v_d, \mathcal{Y})|^p |\mathcal{F}_{i-1}\right) < \infty.$$

Remark 2. Assumption 3 yields the existence of a real $p > 2$ such that almost surely

$$\sup_{i>0} \mathbb{E}(\|c_i\eta_{c_i}(\chi_{i-1}, \mathcal{Y}_i)\|_\star^p|\mathcal{F}_{i-1}) < \infty.$$

Our main result is the following.

Theorem 4.1. *Let $(\beta_i)_{i\geq 0}$, $(\gamma_i)_{i\geq 0}$ and and $(c_i)_{i\geq 0}$ be sequences in $\left(\mathbb{R}_+^*\right)^{\mathbb{N}}$. Assume moreover that $(\beta_i)_{i\geq 0}$ is non decreasing and that the following assumptions are satisfied :*

(i) $\beta_N / \sum_{i=1}^N \gamma_i \xrightarrow[N\to+\infty]{} 0$,

(ii) $\sum_{i=1}^N \gamma_i c_i / \sum_{i=1}^N \gamma_i \xrightarrow[N\to+\infty]{} 0$,

(iii) $\sum_{i=1}^N \frac{\gamma_i^2}{c_i^2 \beta_{i-1}} / \sum_{i=1}^N \gamma_i \xrightarrow[N\to+\infty]{} 0$,

(iv) $\sum_{i=1}^{+\infty} \left(\frac{\gamma_i}{c_i}\right)^2 < \infty$.

Then, provided that Assumptions 1 and 2 are satisfied, we have :

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

If moreover, Assumption 3 is satisfied, we have :

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

Remark 3. Let $0 < c < a - \frac{1}{2} < \frac{1}{2}$. One can choose for $(\beta_i)_{i\geq 0}$ the constant sequence equal to 1, for $(c_i)_{i\geq 0}$ the sequence $(i^{-c})_{i\geq 0}$ and for $(\gamma_i)_{i\geq 0}$ the sequence $(i^{-a})_{i\geq 0}$.

Remark 4. For the "classical" Kiefer-Wolfowitz algorithm, the assumptions on the gains are particular cases of our assumption with a constant sequence $(\beta_i)_{i\geq 0}$:

$$\sum \gamma_n = \infty, \quad \sum \gamma_n c_n < \infty, \quad \text{and} \quad \sum \gamma_n^2 c_n^{-2} < \infty.$$

The proof of Theorem 4.1 uses an estimation of $\varepsilon_N = I(S_N) - I(x^\star)$ inspired from [11].

Lemma 4.2. *Assume that $(\beta_i)_{i\in\mathbb{N}}$ is a non decreasing sequence, then*

$$\begin{aligned} 0 \leq \left(\sum_{i=1}^N \gamma_i\right) \varepsilon_N \quad &\leq \quad \beta_N V(x^\star) - \sum_{i=1}^N \gamma_i \left\langle \eta_{c_i}(\chi_{i-1}, \mathcal{Y}_i), \chi_{i-1} - x^\star \right\rangle \\ &\quad - \sum_{i=1}^N \gamma_i \left\langle r_{c_i}(\chi_{i-1}), \chi_{i-1} - x^\star \right\rangle + \sum_{i=1}^N \frac{\gamma_i^2}{2\alpha\beta_{i-1}} \left\| \Psi_{c_i}(\chi_{i-1}, \mathcal{Y}_i) \right\|_\star^2 . \end{aligned}$$

Proof. We remark that because $I(x^\star) = \inf_{x\in C} I(x)$ then $\varepsilon_N \geq 0$ and for any $i = 1, \ldots, N$, $I(\chi_{i-1}) - I(x^\star) \geq 0$. Because I is a convex and differentiable function,

$$\begin{aligned} \left(\sum_{i=1}^N \gamma_i\right) \varepsilon_N \quad &\leq \quad \sum_{i=1}^N \gamma_i \left(I(\chi_{i-1}) - I(x^\star)\right) \\ &\leq \quad \sum_{i=1}^N \gamma_i \left\langle \nabla I(\chi_{i-1}), \chi_{i-1} - x^\star \right\rangle . \end{aligned}$$

Thus,

$$\begin{aligned}\left(\sum_{i=1}^{N}\gamma_i\right)\varepsilon_N &\leq \sum_{i=1}^{N}\gamma_i\left\langle\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}-x^\star\right\rangle \\ &\quad -\sum_{i=1}^{N}\gamma_i\left\langle\eta_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}-x^\star\right\rangle-\sum_{i=1}^{N}\gamma_i\left\langle r_{c_i}(\chi_{i-1}),\chi_{i-1}-x^\star\right\rangle \\ &= \left\langle\xi_N,x^\star\right\rangle+\sum_{i=1}^{N}\gamma_i\left\langle\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}\right\rangle \\ &\quad -\sum_{i=1}^{N}\gamma_i\left\langle\eta_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}-x^\star\right\rangle-\sum_{i=1}^{N}\gamma_i\left\langle r_{c_i}(\chi_{i-1}),\chi_{i-1}-x^\star\right\rangle\end{aligned}$$

Now, because the sequence $(\beta_i)_{i\in\mathbb{N}}$ is non decreasing and using the convexity and the differentiability of $W_{\beta_{i-1}}$ (as the Legendre transform of a strictly convex function), item 2 in Proposition A.3. of Tauvel [11] yields :

$$\begin{aligned}W_{\beta_i}(\xi_i)-W_{\beta_{i-1}}(\xi_{i-1}) &\leq W_{\beta_{i-1}}(\xi_i)-W_{\beta_{i-1}}(\xi_{i-1}) \\ &\leq \left\langle\nabla W_{\beta_{i-1}}(\xi_{i-1}),\xi_i-\xi_{i-1}\right\rangle+\frac{\gamma_i^2}{2\alpha\beta_{i-1}}\left\|\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i)\right\|_*^2 \\ &= -\gamma_i\left\langle\chi_{i-1},\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i)\right\rangle+\frac{\gamma_i^2}{2\alpha\beta_{i-1}}\left\|\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i)\right\|_*^2.\end{aligned}$$

As a consequence, we get:

$$\begin{aligned}\sum_{i=1}^{N}\gamma_i\left\langle\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}\right\rangle &\leq W_{\beta_0}(\xi_0)-W_{\beta_N}(\xi_N) \\ &\quad +\sum_{i=1}^{N}\frac{\gamma_i^2}{2\alpha\beta_{i-1}}\left\|\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i)\right\|_*^2 \\ &= -W_{\beta_N}(\xi_N)+\sum_{i=1}^{N}\frac{\gamma_i^2}{2\alpha\beta_{i-1}}\left\|\Psi_{c_i}(\chi_{i-1},\mathcal{Y}_i)\right\|_*^2.\end{aligned}$$

The result finally follows from:

$$\langle x^\star,\xi_N\rangle-W_{\beta_N}(\xi_N)\leq\beta_N V(x^\star).$$

□

Remark 5. With our assumptions, we have that

$$\mathbb{E}(<\eta_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}-x^\star>)=0,$$

indeed,

$$<\eta_{c_i}(\chi_{i-1},\mathcal{Y}_i),\chi_{i-1}-x^\star>$$

is a martingale difference.

The following lemma may be deduced from Taylor inequality, recalling that $\chi_i\in C\ \forall i$.

Lemma 4.3. *There exists $\kappa>0$ such that, for any i,*

$$\|r_{c_i}(\chi_{i-1})\|_\star\leq\kappa c_i.$$

Proof. It follows from the fact that the Hessian matrix of I on C is bounded. □

Remark 6. As a consequence of Lemmas 4.2 and 4.3, we have that

$$0 \leq \mathbb{E}(\varepsilon_N) \leq \frac{\beta_N V(x^\star)}{\sum_{i=1}^N \gamma_i} + \frac{2\kappa D \sum_{i=1}^N \gamma_i c_i}{\sum_{i=1}^N \gamma_i} + \frac{\sum_{i=1}^N \frac{B\gamma_i^2}{2c_i^2 \alpha \beta_{i-1}}}{\sum_{i=1}^N \gamma_i},$$

where D is the diameter of C and $\frac{B}{c_i^2}$ denotes a bound of $\mathbb{E}\left\|\Psi_{c_i}(\chi_{i-1}, \mathcal{Y}_i)\right\|_\star^2$. Indeed, because of Decomposition (3.1)

(4.2)

$$\|\Psi_{c_i}\|_\star^2 \leq \|\nabla I\|_\star^2 + \|r_{c_i}\|_\star^2 + \|\eta_{c_i}\|_\star^2 + 2\left(\|\nabla I\|_\star + \|r_{c_i}\|_\star\right)\|\eta_{c_i}\|_\star + 2\|\nabla I\|_\star \times \|r_{c_i}\|_\star.$$

Recall that $\|\nabla I\|_\star$ is bounded (because C is compact and ∇I is continuous), also following Lemma 4.3,

$$\|r_{c_i}\|_\star \leq \kappa c_i.$$

Finally, using (4.1),

$$\mathbb{E}(\|\eta_{c_i}\|_\star^2) \leq \frac{\sigma^2}{c_i^2} \text{ recall that } c_i \to 0 \text{ as } i \to \infty.$$

So that, sufficient conditions to have the L^1-convergence of (ε_N) are Assumptions (i), (ii) and (iii) of Theorem 4.1 (let us recall that $\varepsilon_N > 0$).

We shall apply some standard results for martingales. Let us fix some notations :

Definition 1. Let M_N be a square integrable martingale written as

$$M_N = \sum_{i=1}^N e_i$$

with e_i a martingale difference. Let

$$\langle M \rangle_N = \sum_{i=1}^N \mathbb{E}(e_i^2 | \mathcal{F}_{i-1}),$$

with $\mathcal{F}_{i-1} = \sigma(e_1, \ldots, e_{i-1})$.

The following results are well known.

Theorem (Law of large numbers for martingales, see for instance [8])**.** *Let M_N be a square integrable martingale if*

$$\langle M \rangle_\infty = \lim_{N \to \infty} \langle M_N \rangle < \infty \text{ a.s.}$$

then $(M_N)_{N \in \mathbb{N}}$ converges a.s. to M_∞ square integrable.

Theorem (Chow Lemma, see for instance [4] p.22)**.** *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 :

$$\text{on } \{A_\infty < \infty\} \quad \sum_{k=1}^{\infty} a_k Z_{k+1} \text{ converges} \tag{4.3}$$

$$\text{on } \{A_\infty = \infty\} \quad A_N^{-1} \sum_{k=1}^{N} a_k Z_{k+1} \leq K. \tag{4.4}$$

Proof of Theorem 4.1. We have already proved the convergence in L^1 (see Remark 6).
Let us prove the almost sure convergence. We apply the law of large numbers for martingales to

$$M_N = \sum_{i=1}^{N} \gamma_i < \eta_{c_i}(\chi_{i-1}, \mathcal{Y}_i), \chi_{i-1} - x^\star >,$$

M_N is a martingale. Under Assumption (iv), $\langle M_N \rangle$ is bounded and hence converges because it is increasing, thus M_N converges almost surely. Using Lemma 4.2 we get:

$$0 \leq \varepsilon_N \leq \frac{\beta_N V(x^\star)}{\sum_{i=1}^{N} \gamma_i} - \frac{M_N}{\sum_{i=1}^{N} \gamma_i} + \frac{2D\kappa \sum_{i=1}^{N} c_i\gamma_i}{\sum_{i=1}^{N} \gamma_i} \tag{4.5}$$

$$+ \frac{\sum_{i=1}^{N} \frac{\gamma_i^2}{2\alpha\beta_{i-1}} \|\Psi_{c_i}(\chi_{i-1}, \mathcal{Y}_i)\|_\star^2}{\sum_{i=1}^{N} \gamma_i}. \tag{4.6}$$

The three terms in (4.5) goes to zero a.s. by using the convergence of M_N, (i) and (ii), remark that (i) implies that $\sum_{i=1}^{\infty} \gamma_i = +\infty$. In order to prove that the term in (4.6) converges to zero a.e., we use the decomposition (4.2) and we apply the Chow Lemma twice to $Z_N = c_N^2 \|\eta_{c_N}\|_\star^2$ and $a_N = \frac{\gamma_N^2}{\beta_{N-1}c_N^2}$ (resp. $Z_N = c_N \|\eta_{c_N}\|_\star$ and $a_N = \frac{\gamma_N^2}{\beta_{N-1}c_N}$).
Theorem 4.1 is proved. Indeed, we have proved that

$$I(S^N) \xrightarrow{\text{a.s.}} I(x^\star).$$

Now, one gets $S_N \xrightarrow{\text{a.s.}} x^*$ because S_N belongs to the compact set C. If $s^\star$ is an accumulation point then because $I(S_N) - I(x^\star)$ goes to zero, $s^\star$ is also a minimum of I. So that $s^\star = x^\star$ by uniqueness of the minimum. □

Remark 7. If some higher order moment assumption holds we could obtain an estimation of the rate of convergence, by using Markov inequality and the Chow Lemma.

5. Application to risk indicators and simulations

We go back to our initial problem: we aim at estimating the optimal allocation of the function I (defined by (1.1)) on the simplex

$$\mathcal{U}_u = \{u_k \geq 0,\ u_1 + \cdots + u_d = u\}.$$

In order to run our mirror algorithm, we have to find an auxiliary function V, whose Legendre-Fenschel transform is computable. A natural strongly convex function on $\mathcal{U}_u$ is the entropy function :

$$\delta_u(x) = \sum_{k=1}^{d} \frac{x_k}{u} \ln\left(\frac{x_k}{u}\right),$$

it is $\frac{1}{u^2}$-convex and admits a unique minimum $x_0 = (\frac{u}{d}, \ldots, \frac{u}{d})$. As in [11], we consider the function

$$\begin{aligned} V(x) &= \delta_u(x) - \delta_u(x_0) - \langle \nabla \delta_u(x_0), x - x_0 \rangle \\ &= \ln d + \sum_{k=1}^{d} \frac{x_k}{u} \ln\left(\frac{x_k}{u}\right). \end{aligned}$$

This function is also $\frac{1}{u^2}$-convex and its Legendre-Fenschel transform is easily computable :

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

Remark that this function involves an exp part which could be problematic in the computations if u is large. Nevertheless, in that case, we could take $\beta = u$, so that the computations remain feasible.

5.1. Properties of the risk indicator I.

We have proved (Proposition 2.1) that our risk indicator I is convex. If we chose g_k a C^2 function for all k, then I is also C^2. In order to prove that it satisfies Assumption 1, it remains to prove that the minimum $x^\star$ is unique. Let $s^\star$ be another minimum. Going back to the proof of Proposition 2.1, we have:

$$\begin{aligned} &I(s^\star) - I(x^\star) - (s^\star - x^\star)^T \nabla I(x^\star) \geq \\ &\quad - \sum_{k=1}^{d}\sum_{p=1}^{n} \mathbb{E}\left[g_k(Y_p^k + s_k^\star)\mathbf{1}_{\{S_p > -u\}}(\mathbf{1}_{\{Y_p^k < -x_k^\star\}} - \mathbf{1}_{\{Y_p^k < -s_k^\star\}})\right]. \end{aligned}$$

Now, if $x^\star \in \mathrm{int}(\mathcal{U}_u)$ the interior of $\mathcal{U}_u$ then $\nabla I(x^\star) = 0$; if $x^\star$ is in the boundary of $\mathcal{U}_u$, then $(s^\star - x^\star)^T \nabla I(x^\star) \geq 0$. In any cases, we have:

$$I(s^\star) - I(x^\star) \geq -\sum_{k=1}^{d}\sum_{p=1}^{n} \mathbb{E}\left[g_k(Y_p^k + s_k^\star)\mathbf{1}_{\{S_p > -u\}}(\mathbf{1}_{\{Y_p^k < -x_k^\star\}} - \mathbf{1}_{\{Y_p^k < -s_k^\star\}})\right],$$

the right term of the inequality being positive provided that for some $k = 1, \ldots, d$ or some $p = 1, \ldots, n$, for $u < w_k < v_k$,

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

(see the proof of Proposition 2.1 for the proof that all the terms in the sum are positive under these hypothesis).

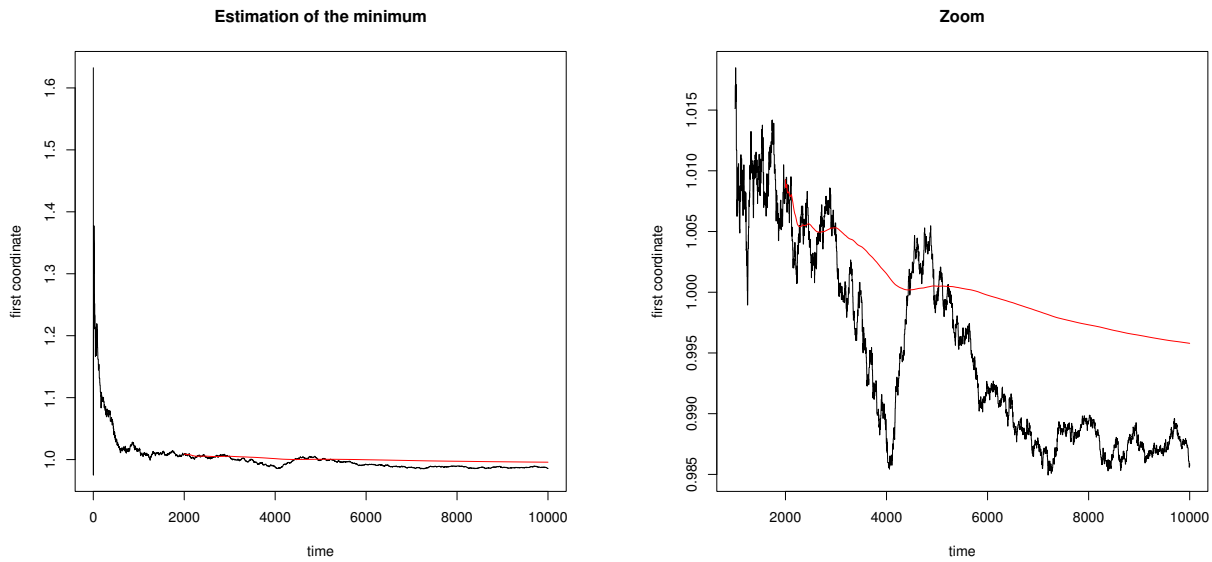
In the sequel we have performed some simulations for some models. For simplicity, we consider only normal laws. A deeper study involving other laws (with heavy tail e.g.), more realistic models and temporal dependencies will be done in a forthcoming work.
In Section 5.2, we consider that $n = 1$ (observation of several periods of length 1) and that $X_p^k = Y_p^k$ are independent ($k = 1, \ldots, d$, $p \in \mathbb{N}$). In Section 5.3, we are still considering that $n = 1$, and that there is no temporal dependencies: the vectors $X_p \in \mathbb{R}^d$ are independent and identically distributed random vectors, but some dependencies on the coordinates of X_p are allowed. In section 5.4, a example with temporal dependency is provided.
Below, the algorithm has been performed with the following sequences $(\gamma_n)_{n\in\mathbb{N}^\star}$, $(c_n)_{n\in\mathbb{N}^\star}$ and $(\beta_n)_{n\in\mathbb{N}^\star}$:

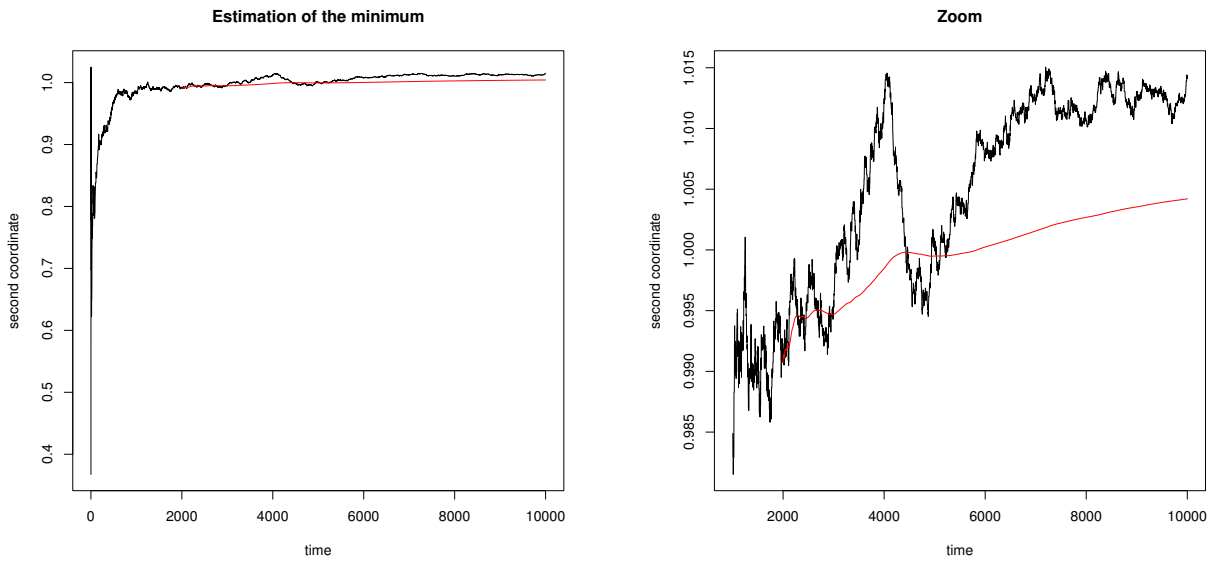
- $\gamma_n = \dfrac{1}{(n+1)^\alpha}$ with $\alpha = \frac{3}{4} + \frac{1}{10}$,
- $c_n = \dfrac{1}{(n+1)^\delta}$ with $\delta = \frac{1}{4}$,
- $\beta_n = 1$.

Also, we have chosen to perform the simulations with $u = 2$. The initialization of the algorithm (χ_0) is done at random uniformly in the simplex $\mathcal{U}_u$.

5.2. **Independent models.** The aim of this section is to provide some benchmark for the use of the algorithm. So we consider the simplest examples : $d = 2$, $n = 1$, the coordinates X_i^1 and X_i^2 are independent and the vectors X_i are also independent. At first, we assume that X_i^1 and X_i^2 have the same normal laws, then we consider different normal laws.

5.2.1. *Same normal laws.* We have chosen independent normal laws with mean equal to 0.3 and standard deviation equal to 1. For $N = 10000$ independent simulations we obtain : $S_N = (0.996; 1.004)$ and the following graphs. Remark that, as expected, it seems that the minimum of our risk indicator is reached for $u_1 = u_2 = \frac{u}{2} = 1$. Below we have plotted for each of the two coordinates the trajectory χ_i (black line) and S_i (gray line); we have plotted the whole trajectories ($i = 1, \ldots, 10000$) and a zoom on the end of the trajectories ($i = 2000, \ldots, 10000$).





We conclude this subsection on independent same normal laws by a study of

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

5.2.2. *Different normal laws.* For this subsection, we have chosen to present the results of 50 simulations of length $N = 1000$, with two independent normal laws. First of all, we consider for the first coordinate a normal law $\mathcal{N}(0.3, 1)$ and for the second coordinate a normal law $\mathcal{N}(0.8, 1)$.
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

We observe that there is a significant difference between the two coordinates. As one could have guest, in order to minimize the risk indicator, one should affect more initial capital to the branch whose expected income is the less (which is the more risky branch in that case).
We terminate this section on independent normal laws with two normal laws with different variance (and same mean). As above, we have performed 50 simulations of length $N = 1000$, with two independent normal laws. For the first coordinate we consider a normal law $\mathcal{N}(0.3, 1)$ and for the second coordinate we consider a normal law $\mathcal{N}(0.3, 4)$.
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

In that case also, we observe that there is a significant difference between the two coordinates. Here, the more risky branch is the second one (which has a higher variance).

5.3. **Model with spatial dependencies.** We propose two simple models of spatial dependence. First of all we consider random vectors in $\mathbb{R}^3$ with the first two coordinates begin independent identically distributed normal laws, the third coordinate begin two times the second coordinate ($X_3 = 2X_2$). A second example of spatial dependence is given by Gaussian vectors.

5.3.1. *Strong dependence.* We consider the following model : two independent normal laws X_1 and X_2 with mean equal to 0.3 and standard deviation equal to 1 and let $X_3 = 2X_2$. We have performed 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 3−uplet (u_1, u_2, u_3) satisfying the minimum of I 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

From our model, X_2 and X_3 fail together and if they fail, X_3 fails two times more than X_2. The loss is larger when X_2 and X_3 fail together. This is revealed by the results of the simulation where $\widehat{u_2} + \widehat{u_3} > \widehat{u_1}$ and $\widehat{u_3} \approx 2\ \widehat{u_2}$.

5.3.2. *Gaussian vectors.* We conclude this paragraph on vectorial dependence with an example of a Gaussian vector in dimension 2 and an example of a Gaussian vector in dimension 3. First, we consider the Gaussian vector with covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

and expectation $m = (0.3, 0.3)$. As above, we have performed 50 simulations of length $N = 1000$, of the Gaussian vector X.
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	0.999	1.001
standard error	0.034	0.034

It seems that the result is the same as in the non correlated case but with a smaller standard deviation.
We have also performed the simulations for a Gaussian vector with expectation $m = (0.3, 0.8)$ and the covariance matrix Σ. 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	1.21	0.79
standard error	0.0.34	0.034

As in the above example where the marginal laws where the same, it seems that the result is the same as in the non correlated case but with a smaller standard error.
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.
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

5.4. **An example with temporal dependency.** We end this simulation section with an example with temporal dependency. The first to coordinates are independent $AR(1)$ processes with same law ($X^i = 0.4X^{i-1} + \varepsilon_i$ with $(\varepsilon_i)_{i\in\mathbb{N}}$ a Gaussian white noise ($\mathcal{N}(0,1)$). Then the third coordinate is two times the second one. We consider that we observe 500 independent periods of length $n = 5$.
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	0.85	0.39	0.76
standard error	0.035	0.012	0.025

Some more work has to be done for realistic models, including temporal dependencies, as well from a theoretical and simulation point of view, in order to propose a useful tool for risk managers. This will be the aim of a forthcoming paper.

References

[1] N. Bartoli, P. Del Moral *Simulation & algorithmes stochastiques.* Cépaduès-éditions (2001).

[2] R. Biard, S. Loisel, C. Macci and N. Veraverbeke *Asymptotic behavior of the finite-time expected time-integrated negative part of some risk processes and optimal reserve allocation.* Journal of Mathematical Analysis and Applications, to appear (2010).

[3] J. Cai, H. Li *Dependence properties and bounds for ruin probabilities in multivariate compound risk models.* Journal of Multivariate Analysis, **98**, (2007), 757-773.

[4] M. Duflo, *Random Iterative Models.* Springer (1997).

[5] H.U. Gerber, *An introduction to Mathematical Risk Theory.* Huebner Foundation monograph n°8, Irwin Homewood IL (1979).

[6] HJ Kushner, G Yin *Stochastic approximation and recursive algorithms and applications.* (2003), Applications of Mathematics Stochastic modelling and applied probability.

[7] J. Lelong, *Almost sure convergence of randomly truncated stochastic algorithms under verifiable conditions.* Statistics & Probability Letters, **78**, 16 (2008), 2632-2636.

[8] J.Neveu *Martingales à temps discret.* Masson et Cie, éditeurs, Paris, 1972.

[9] Nemirovski & Yudin *Problem Complexity and Method Efficiency in Optimization.* John Wiley and Sons, 1983.

[10] S. Loisel, *Differentiation of some functionals of risk processes and optimal reserve allocation.* Journal of Applied Probability, **42**, 2, (2005), 379-392.

[11] C. Tauvel, *Optimisation stochastique à grande échelle*, PhD Thesis, (2008).

Université de Bourgogne, IMB UMR 5584
E-mail address: peggy.cenac@u-bourgogne.fr

Université de Lyon, Université Lyon 1, Laboratoire SAF EA 2429
E-mail address: veronique.maume@univ-lyon1.fr

Université Joseph Fourier, Tour IRMA, MOISE-LJK
E-mail address: clementine.prieur@imag.fr