Portfolio optimization under CV@R constraint with stochastic mirror descent

Manon Costa

Institut de Mathématiques de Toulouse,
joint work with Sébastien Gadat (TSE, IUF) and Lorick Huang (INSA Toulouse).

Model

Consider a portfolio with assets

$$Z = (Z_1, \cdots, Z_m),$$

where $Z_i = \frac{A_i(T)}{A_i(0)} - 1$ relative return at fixed time horizon.

An allocation strategy is a vector $u = (u_1, \cdots, u_m) \in \Delta_m$ such that

$$0 \le u_i \le 1, \qquad \sum_{i=1}^{m} u_i = 1$$

and is associated with a mean return

$$\mathbb{E}(\langle Z, u \rangle) = \sum_{i=1}^{m} u_i \mathbb{E}(Z_i)$$

Model

What is the best way to allocate resources to optimize the mean relative return at a fixed time horizon ?

$$\mathbb{E}(\langle Z, u\rangle) = \sum_{i=1}^{m} u_i \mathbb{E}(Z_i)$$

- Without constrain, we choose the asset with maximal expected return.
- **But** this could lead to large losses between gains.

$\longrightarrow$ Risk management constraint

Risk management constraint

CV@R (conditional value at risk)
Fix $\alpha > 0$, the quantile (or *value at risk* V@R) is

$$V@R_\alpha(u) = \sup\{q \in \mathbb{R} \,:\, \mathbb{P}(\langle Z, u\rangle \leq q) \leq \alpha\},$$

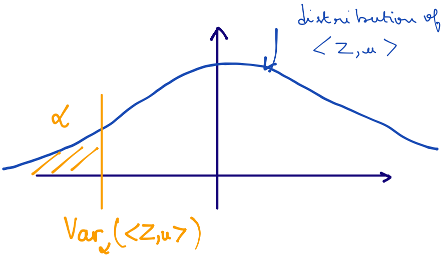
and

$$CV@R_\alpha(u) = \mathbb{E}[-\langle Z, u\rangle \,|\, \langle Z, u\rangle \leq V@R_\alpha(u)].$$

We choose α such that

- $V@R_\alpha(u) < 0$.
- $CV@R_\alpha(u) \geq 0$.

Expected absolute value of large losses.



Risk management constraint

CV@R (conditional value at risk)
Fix $\alpha > 0$, the quantile (or *value at risk* V@R) is

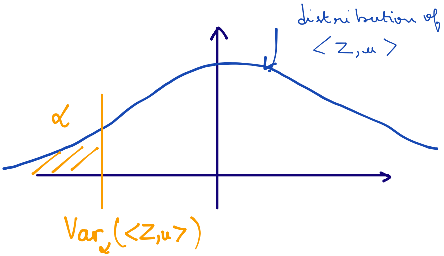
$$V@R_\alpha(u) = \sup\{q \in \mathbb{R} : \mathbb{P}(\langle Z, u \rangle \leq q) \leq \alpha\},$$

and

$$CV@R_\alpha(u) = \mathbb{E}[-\langle Z, u \rangle \mid \langle Z, u \rangle \leq V@R_\alpha(u)].$$

We impose that

$$CV@R_\alpha(u) \leq M.$$



The optimization problem

$$\mathcal{P}_M := \arg\max_{u\in\Delta_m} \left\{ \sum_{i=1}^{m} u_i \mathbb{E}[Z^i] \ : \ CV@R_\alpha(u) \leq M \right\}$$
$$= \arg\min_{u\in\Delta_m} \left\{ -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] \ : \ CV@R_\alpha(u) \leq M \right\}$$

Lagragian formulation of the optimization problem

$$\mathcal{Q}_\lambda \quad := \arg\min_{u\in\Delta_m} \left\{ -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda CV@R_\alpha(u) \right\}.$$

The optimization problem

$$\mathcal{P}_M = \arg\min_{u \in \Delta_m} \left\{ -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] \ : \ CV@R_\alpha(u) \leq M \right\}$$

Proposition

For any feasible constraint $M > 0$, a solution u_M^* exists to $\mathcal{P}_M$ such that

$$\exists \lambda_M^\star > 0 \quad u_M^\star = \arg\min_{u \in \Delta_m} \left\{ -\sum_{i=1}^{m} u_i \mathbb{E}[S_i] + \lambda_M^\star CV@R_\alpha(u) \right\}.$$

Moreover, $\lambda_M^\star$ is a decreasing function of M.
Oppositely, any solution v_λ of $\mathcal{Q}_\lambda$ solves $\mathcal{P}_M$ with $M = CV@R_\alpha(v_\lambda)$.

[P. Krokhmal, J. Palmquist, and S. Uryasev. *Portfolio optimization with conditional value-at-risk objective and constraints.* (2001)]

Convex representation of the CV@R

$$\mathcal{Q}_\lambda = \arg\min_{u \in \Delta_m} \left\{ -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda CV@R_\alpha(u) \right\}.$$

Convex representation of the CV@R

$$\mathcal{Q}_\lambda = \arg\min_{u \in \Delta_m} \left\{ -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda CV@R_\alpha(u) \right\}.$$

As introduced by Rockafeller and Uryasev (2000),

$$\begin{aligned} CV@R_\alpha(u) &= \arg\min_{\theta \in \mathbb{R}} \psi_\alpha(u, \theta), \\ &= \arg\min_{\theta \in \mathbb{R}} \theta + \frac{1}{1-\alpha} \mathbb{E}\left[\lfloor \langle Z, u \rangle - \theta \rfloor_+ \right], \end{aligned}$$

where $\lfloor x \rfloor_+ = \max(0, x)$.

ψ_α is the convex coercive Lipschitz continuous and differentiable function.

Convex unconstrained problem

$$\mathcal{Q}_\lambda = \arg \min_{(u,\theta)\in\Delta_m\times\mathbb{R}} \{p_\lambda(u,\theta)\}, \tag{1}$$

where the key function p_λ is defined by :

$$p_\lambda(u,\theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{\theta + \frac{1}{1-\alpha}\mathbb{E}\left[\lfloor\langle Z, u\rangle - \theta\rfloor_+\right]\right\} \tag{2}$$

Convex unconstrained problem

$$\mathcal{Q}_\lambda = \arg \min_{(u,\theta)\in\Delta_m\times\mathbb{R}} \{p_\lambda(u,\theta)\}, \quad (1)$$

where the key function p_λ is defined by :

$$p_\lambda(u,\theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{ \theta + \frac{1}{1-\alpha} \mathbb{E}\left[\lfloor \langle Z, u\rangle - \theta \rfloor_+ \right] \right\} \quad (2)$$

- The function $p_\lambda(u,\theta)$ writes as an expectation, **Robbins-Monro** stochastic algorithms are available,

Robbins Monro algorithms (1951)

Aim : Find y^* such that $h(y^*) = 0$ when the function $h(y) = \mathbb{E}(H(y, Z))$.

Here, we search for (u, θ) such that $\nabla p_\lambda(u, \theta) = 0$.

$$p_\lambda(u, \theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{ \theta + \frac{1}{1-\alpha} \mathbb{E}\left[\lfloor \langle Z, u \rangle - \theta \rfloor_+ \right] \right\}$$

Assume $(Z_n)_{n \geq 1}$ is a sequence of *i.i.d* random variables, and $(\gamma_n)_{n \geq 1}$ is a sequence of step sizes such that

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

Then the Robbins Monro algorithm writes

$$y_{n+1} = y_n - \gamma_n H(y_n, Z_{n+1}).$$

Optimization problem

$$\mathcal{Q}_\lambda = \arg \min_{(u,\theta)\in\Delta_m\times\mathbb{R}} \{p_\lambda(u,\theta)\},$$

where the key function p_λ is defined by :

$$p_\lambda(u,\theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{\theta + \frac{1}{1-\alpha}\mathbb{E}\left[\lfloor\langle Z, u\rangle - \theta\rfloor_+\right]\right\}$$

- The function $p_\lambda(u,\theta)$ writes as an expectation, **Robbins-Monro** stochastic algorithms are available, **but**
  - → the functions inside the expectation are not smooth.
    sub-gradients techniques
  - → we have only access to biaised simulation of the random variables Z.

Optimization problem

$$\mathcal{Q}_\lambda = \arg \min_{(u,\theta)\in\Delta_m\times\mathbb{R}} \{p_\lambda(u,\theta)\},$$

where the key function p_λ is defined by :

$$p_\lambda(u,\theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{\theta + \frac{1}{1-\alpha}\mathbb{E}\left[\lfloor\langle Z, u\rangle - \theta\rfloor_+\right]\right\}$$

- We optimize on $u \in \Delta_m$
  - $\rightarrow$ either project the stochastic gradient descent
  - $\rightarrow$ use a different strategy that takes the geometry into account. *Stochastic mirror descent*

Deterministic mirror descent

The mirror descent was introduced by A. Nemirovkij and D. Yudin (1983).

Main idea :

- avoid projection by changing the metric in the space
- choose a "metric" in which the mirror/proximal mapping is explicit.

A. Beck and M.Teboulle (2003), *Mirror descent and nonlinear projected subgradient methods for convex optimization*,
G. Lan, A. Nemirovskij, and A. Shapiro. (2012.) *Validation analysis of mirror descent stochastic approximation method.*
Z. Zhou, P. Mertikopoulos, N. Bambos, S. Boyd, and P. Glynn. (2017). *Stochastic mirror descent in variationally coherent optimization problems.*

Deterministic miror descent

When we consider the question of minimizing a convex smooth f, the gradient descent writes :

$$x_{k+1} = x_k - \frac{1}{2\eta_k}\nabla f(x_k),$$

it is equivalent to the proximal problem

$$x_{k+1} = \arg min_x\{\langle x, \nabla f(x_k)\rangle + \frac{1}{2\eta_k}||x - x_k||^2\}$$

The mirror descent considers

$$x_{k+1} = \arg min_x\{\langle x, \nabla f(x_k)\rangle + \frac{1}{\eta_k}D(x, x_k)\}$$

where D is a Bregman distance function. [A. Beck and M. Teboulle (2003)]

Bregman distance

We consider the strongly convex negative entropy on Δ_m and the L^2 norm on $\mathbb{R}$:

$$\Phi(u, \theta) = \sum_{i=1}^{m} u_i \log(u_i) + \frac{\theta^2}{2}$$

The Bregman distance is defined as

$$D_\Phi(u, v) = \Phi(u) - \Phi(v) - \langle \nabla\Phi(v), u - v \rangle.$$

Deterministic mirror descent

Let $x = (u, \theta)$, then the deterministic mirror descent writes

$$X_{k+1} = \arg \min_{x \in \Delta_m \times \mathbb{R}} \left\{ \langle \nabla p_\lambda(X_k), x - X_k \rangle + \frac{1}{\eta_{k+1}} D_\Phi(x, X_k) \right\}.$$

This minimization can be made explicit :

$$X_{k+1} = \begin{pmatrix} U^{k+1} \\ \theta^{k+1} \end{pmatrix} \quad \text{with} \quad \begin{cases} U^{k+1} = \dfrac{U^k e^{-\eta_{k+1} \partial_u p_\lambda(U^k, \theta^k)}}{\| U^k e^{-\eta_{k+1} \partial_u p_\lambda(U^k, \theta^k)} \|_1} \\ \theta^{k+1} = \theta^k - \eta_{k+1} \partial_\theta p_\lambda(U^k, \theta^k) \end{cases},$$

where the first equation has to be understood within a m dimensional vector structure.

Subgradient

Recall that

$$p_\lambda(u,\theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{ \theta + \frac{1}{1-\alpha} \mathbb{E}\left[\lfloor \langle Z, u \rangle - \theta \rfloor_+ \right] \right\}$$

Subgradient

If f is a convex function, η is a sub-gradient of f in x_0 if

$$\forall x, \qquad f(x) \geq f(x_0) + \langle \eta, x - x_0 \rangle.$$

For $f(x) = \max(x, 0)$ we obtain

$$\partial f(x) = \begin{cases} 1\,, \text{if } x > 0 \\ 0\,, \text{if } x < 0 \\ [0,1]\,, \text{if } x = 0 \end{cases}$$

Subgradient

Recall that

$$p_\lambda(u,\theta) = -\sum_{i=1}^{m} u_i \mathbb{E}[Z_i] + \lambda \left\{ \theta + \frac{1}{1-\alpha} \mathbb{E}\left[\lfloor \langle Z, u \rangle - \theta \rfloor_+ \right] \right\}$$

We will therefore choose

$$\partial_u p_\lambda(u,\theta) = \mathbb{E}(g_1(Z,u,\theta)), \qquad g_1(Z,u,\theta) = -Z + \frac{\lambda}{1-\alpha} Z 1_{\langle Z,u \rangle \geq \theta}$$

and

$$\partial_\theta p_\lambda(u,\theta) = \mathbb{E}(g_2(Z,u,\theta)), \qquad g_2(Z,u,\theta) = \lambda \left[1 - \frac{1}{1-\alpha} 1_{\langle Z,u \rangle \geq \theta} \right]$$

Our algorithm

Data Step-size sequence $(\eta_n)_{n\geq 0}$ and $U_0 \in \mathbb{R}$, $\theta_0 \in \mathbb{R}$; $\alpha \in (0,1)$
Results Two sequences : $X_k = (U_k, \theta_k)_{k\geq 0}$
for $k = 0, \ldots,$ **do**

Simulate the random variable $\hat{Z}^{k+1}$
Compute a stochastic approximation $\hat{g}_{k+1}$ of $\nabla p_\lambda(U_k, \theta_k)$ with :

$$\begin{cases} \hat{g}_{k+1,1} & = -\hat{Z}^{k+1} + \frac{\lambda}{1-\alpha}\hat{Z}^{k+1} 1_{\langle \hat{Z}^{k+1}, U_k\rangle \geq \theta_k} \\ \hat{g}_{k+1,2} & = \lambda\left[1 - \frac{1}{1-\alpha} 1_{\langle \hat{Z}^{k+1}, U_k\rangle \geq \theta_k}\right] \end{cases}.$$

Update the algorithm
$X_{k+1} = \arg\min_{x\in\Delta_m\times\mathbb{R}} \left\{\langle \hat{g}_{k+1}, x - X_k\rangle + \frac{1}{\eta_{k+1}} D_\Phi(x, X_k)\right\}$ using :

$$X_{k+1} = (U_{k+1}, \theta_{k+1}), \qquad \begin{cases} U^{k+1} & = \frac{U^k e^{-\eta_{k+1}\hat{g}_{k+1,1}}}{\|U^k e^{-\eta_{k+1}\hat{g}_{k+1,1}}\|_1} \\ \theta^{k+1} & = \theta^k - \eta_{k+1}\hat{g}_{k+1,2} \end{cases}.$$

A first recursion step

A key argument to study the algorithm is to write a recursion inequality on $D_\phi(x^*, X_k)$.

Starting from

$$X_{k+1} = \arg\min_{x \in \mathcal{X}} \left\{ \langle \hat{g}_{k+1}, x - X_k \rangle + \frac{D_\phi(x, X_k)}{\eta_{k+1}} \right\},$$

we can obtain

$$D_\phi(x, X_{k+1}) \leq D_\phi(x, X_k) + \eta_{k+1}^2 \left[C_\alpha + \|\hat{g}_{k+1,1}\|^2 \right] - \eta_{k+1} \langle \hat{g}_{k+1}, X_k - x \rangle.$$

Drift term

$$\begin{aligned}\hat{g}_{k+1} &= \nabla p_\lambda(X_k) + (\mathbb{E}[\hat{g}_{k+1} \mid \mathcal{F}_k] - \nabla p_\lambda(X_k)) + (\hat{g}_{k+1} - \mathbb{E}[\hat{g}_{k+1} \mid \mathcal{F}_k]) \\ &= \nabla p_\lambda(X_k) - \mathfrak{b}_{k+1} + \Delta M_{k+1},\end{aligned}$$

$\mathfrak{b}_{k+1}$ stands for the bias :

$$\mathfrak{b}_{k+1} := \nabla p_\lambda(U_k, \theta_k) - \mathbb{E}[\hat{g}_{k+1} \mid \mathcal{F}_k]$$

$$= \begin{pmatrix} \mathbb{E}[\hat{Z}^{k+1} \mid \mathcal{F}_k] - \mathbb{E}[Z] + \frac{\lambda}{1-\alpha}\left(\mathbb{E}[Z 1_{\langle Z, U_k\rangle \geq \theta_k}] - \mathbb{E}[\hat{Z}^{k+1} 1_{\langle \hat{Z}^{k+1}, U_k\rangle \geq \theta_k} \mid \mathcal{F}_k]\right. \\ \mathbb{P}(\langle Z, U_k\rangle \geq \theta_k) - \mathbb{E}[1_{\langle \hat{Z}^{k+1}, U_k\rangle \geq \theta_k} | \mathcal{F}_k] \end{pmatrix}$$

Assumptions on the biased simulations

The sequence $(\hat{Z}^k)_{k\geq 0}$ satisfies both :

$$\mathcal{W}_1(\mathcal{L}(\hat{Z}^{k+1}), \mathcal{L}(Z)) \leq \delta_{k+1},$$

where $\mathcal{W}_1$ stands for the Wasserstein-1 distance.

$\forall u \in \Delta_m, \quad \forall \theta \in \mathbb{R},$

$$\left\| \mathbb{E}\left[\langle Z, u\rangle 1_{\langle Z,u\rangle \geq \theta} - \langle \hat{Z}^{k+1}, u\rangle 1_{\langle \hat{Z}^{k+1},u\rangle \geq \theta} \,\middle|\, \mathcal{F}_k \right] \right\| \leq \upsilon_{k+1}.$$

Assumptions on the biased simulations

The sequence $(\hat{Z}^k)_{k\geq 0}$ satisfies both :

$$\mathcal{W}_1(\mathcal{L}(\hat{Z}^{k+1}), \mathcal{L}(Z)) \leq \delta_{k+1},$$

where $\mathcal{W}_1$ stands for the Wasserstein-1 distance.

$\forall u \in \Delta_m, \quad \forall \theta \in \mathbb{R},$

$$\left\| \mathbb{E}\left[\langle Z, u\rangle 1_{\langle Z,u\rangle \geq \theta} - \langle \hat{Z}^{k+1}, u\rangle 1_{\langle \hat{Z}^{k+1},u\rangle \geq \theta} \,\middle|\, \mathcal{F}_k \right] \right\| \leq \upsilon_{k+1}.$$

We deduce that

$$\mathbb{E}[\|\mathfrak{b}_{k+1}\|] \leq 2\sqrt{\delta_{k+1}} + \delta_{k+1} + \frac{\lambda}{1-\alpha}\upsilon_{k+1}.$$

Convergence

(Theorem - Almost sure convergence of the biased SMD)

Assume that $\sum_{k\geq 0} \eta_{k+1} = +\infty$, and $\sum_{k\geq 0} \eta_{k+1}^2 < +\infty$, and that

$$\sum_{k\geq 0} \eta_{k+1}(\sqrt{\delta_{k+1}} + \upsilon_{k+1}) < +\infty,$$

then the Cesaro average $\bar{X}_k^\eta$ defined by

$$\bar{X}_k^\eta := \left(\sum_{i=0}^{k} \eta_i\right)^{-1} \left(\sum_{i=0}^{k} \eta_i X_i\right) \quad (3)$$

converges *a.s.* and

$$p_\lambda(\bar{X}_k^\eta) \longrightarrow min(p_\lambda) \quad a.s.$$

Finite horizon controls

$$\mathrm{D}_{\Phi}^{k} = \mathbb{E}\mathcal{D}_{\Phi}(x_{\lambda}^{*}, X_k).$$

Coming back to the recursion we can obtain

$$\mathrm{D}_{\Phi}^{k} \le \mathrm{D}_{\Phi}^{0} \prod_{i=1}^{k}(1+a_i) + \left(\sum_{j=1}^{k} \frac{b_j}{\prod_{i=1}^{j}(1+a_i)} \right) \prod_{i=1}^{k}(1+a_i). \tag{4}$$

where

$$\mathrm{D}_{\Phi}^{0} \le \frac{(\theta_0 - V@R_{\alpha}(u_{\lambda}^{\star}))^2}{2} + \log m := \{\Delta_{\Phi}^{0}\}^2,$$

$$\begin{cases} a_{k+1} = 2\eta_{k+1}\left(2\sqrt{\delta_{k+1}} + \delta_{k+1} + \dfrac{\lambda \upsilon_{k+1}}{1-\alpha}\right) \\ b_{k+1} = C\left(\eta_{k+1}^2 + \eta_{k+1}\left(2\sqrt{\delta_{k+1}} + \delta_{k+1} + \dfrac{\lambda \upsilon_{k+1}}{1-\alpha}\right)\right) \end{cases}.$$

Finite horizon controls

(Finite-time guarantees)

Recall that $\bar{X}_k^\eta := \left(\sum_{i=0}^k \eta_i\right)^{-1}\left(\sum_{i=0}^k \eta_i X_i\right)$ then for any $n > 1$,

$$\mathbb{E}[p_\lambda(\bar{X}_n^\eta)] - p_\lambda(x_\lambda^\star) \leq \left(\sum_{j=0}^{n-1} \eta_{j+1}\right)^{-1}\left(\mathrm{D}_\Phi^0 + \sum_{k=0}^{n-1}\left[a_{k+1}\mathrm{D}_\Phi^k + b_{k+1}\right]\right)$$

where

$$a_{k+1} = 2\eta_{k+1}\left(2\sqrt{\delta_{k+1}} + \delta_{k+1} + \frac{\lambda\upsilon_{k+1}}{1-\alpha}\right), \quad b_{k+1} = C\left(\eta_{k+1}^2 + a_{k+1}\right)$$

SMD with a constant step-size sequence

Consider the step sequence

$$\eta_{k+1} = \eta > 0, \quad \forall\, 0 \leq k \leq n.$$

We will also assume a constant upper bound of the bias in the simulation

$$2\sqrt{\delta_{k+1}} + \delta_{k+1} + \frac{\lambda \upsilon_{k+1}}{1-\alpha} = \omega > 0, \qquad \forall 1 \leq k \leq n.$$

For a given $n \in \mathbb{N}$, if (η, ω) are chosen such that
$\eta = \frac{\Delta_\Phi^0}{2\sqrt{n+1}}$ and $\omega = \frac{1}{\sqrt{n+1}\Delta_\Phi^0}$ with $\{\Delta_\Phi^0\}^2 = \frac{(\theta_0 - V@R_\alpha(u_\lambda^\star))^2}{2} + \log m$,
then there exists $C > 0$ large enough such that :

$$\mathbb{E}\left[p_\lambda\left(\hat{X}_n^\eta\right)\right] - p_\lambda(x_\lambda^\star) \leq C \frac{|\theta_0 - V@R_\alpha(u_\lambda^\star)| + \sqrt{\log m}}{\sqrt{n+1}}.$$

Portfolio model

The portfolio Z contains $m = m' + 1$ assets

- a return Y obtained as the baseline short-term interest rate Cox-Ingersoll-Ross process $(r_t)_{t \geq 0}$ with no drift,
- a family $\mathbf{S} = (S^1, \ldots S^{m'})$ of $m' = m - 1$ geometric Brownian motions that encode some risky assets in the portfolio.

Portfolio model

The portfolio Z contains $m = m' + 1$ assets

- a return Y obtained as the baseline short-term interest rate Cox-Ingersoll-Ross process $(r_t)_{t \geq 0}$ with no drift,
- a family $\mathbf{S} = (S^1, \dots S^{m'})$ of $m' = m - 1$ geometric Brownian motions that encode some risky assets in the portfolio.

The CIR short rate model depends on a triple (a, b, σ_0)

$$dr_t = a(b - r_t)dt + \sigma_0 \sqrt{r_t} dB_0(t), \qquad (5)$$

where $(B_0(t))_{t \geq 0}$ stands for a standard real Brownian motion.

- b stands for the long-time mean of the short rate
- a quantifies the strength of the mean-reversion effect.
- σ_0 is the volatility.

Portfolio model - 2

$$dr_t = a(b - r_t)dt + \sigma_0 \sqrt{r}_t dB_0(t),$$

Then the portfolio $Z_t = (Y_t, S_t^1, \cdots, S_t^{m'})$ writes

$$\forall t \geq 0 \qquad \begin{cases} dY_t &= r_t Y_t dt, \\ dS_t^i &= \mu_i S_t^i dt + \sigma_i S_t^i dB_i(t), \quad \forall i \in \{1, \dots, m'\}, \end{cases} \tag{6}$$

where $\mathrm{B} = (B_0, B_1, \dots, B_{m'})$ refers to a multivariate Brownian motion with correlated components with

$$\mathbb{E}[B_i(t)B_j(t)] = \rho_{i,j} t,$$

Assumptions on the portfolio parameters *We assume that :*

i) the CIR parameters satisfy $ab > \sigma_0^2$ *and* $a > 2\sqrt{2}\sigma_0$.

ii) the correlation matrix of the Brownian motions Σ *in invertible.*

Simulation of the portfolio

Geometric Brownian motion can be exactly simulated.

- Let $W(t) = (W_0(t), W_1(t), \ldots, W_{m'}(t))$ be independent standard Brownian paths
- Cholesky decomposition :

$$LL^T = \Sigma \qquad \text{and} \quad B(t) = LW(t).$$

- Finally for $i \in \{1, \cdots m'\}$

$$S_1^i = S_0^i \exp\left(\left(\mu_i - \frac{(\sigma_i)^2}{2}\right) + \sigma_i B_1^i\right)$$

Simulation of the CIR

Main issues

$$dr_t = a(b - r_t)dt + \sigma_0 \sqrt{r}_t dB_0(t),$$
$$Y(t) = Y_0 \exp\left(\int_0^t r_s ds \right).$$

- Even if there exists method to simulate the CIR exactly, the integral needs to be approximated.
- We need to choose an discretization scheme that allows the control of the error in $Z\mathbb{1}_{\langle Z,u\rangle \geq \theta}$.

Simulation of the CIR

Discrete time grid $(kh)_{\{1\leq k\leq N\}}$ sur $[0,1]$

Drift implicit Euler scheme

$$\hat{r}_{(k+1)h} = \left(\frac{\sqrt{\hat{r}_{kh}} + \frac{\sigma_0}{2}\Delta B_0^{(k)}}{2(1+\frac{ah}{2})} + \sqrt{\frac{\left(\sqrt{\hat{r}_{kh}} + \frac{\sigma_0}{2}\Delta B_0^{(k)}\right)^2}{4(1+\frac{ah}{2})^2} + \frac{(4ab-\sigma_0^2)h}{8(1+\frac{ah}{2})}} \right)^2 .$$

[Alfonsi (2005), (2013), S. Dereich, A. Neuenkirch, and L. Szpruch (2012)]

The method derives from the SDE satisfied by $y_t = \sqrt{r_t}$:

$$\hat{y}_{(k+1)h} \quad = \quad \hat{y}_{kh} + \left(\frac{4ab-\sigma_0^2}{8\hat{y}_{(k+1)h}} - \frac{a}{2}\hat{y}_{(k+1)h} \right) h + \frac{\sigma_0}{2}\Delta B_0^{(k)},$$

where $\Delta B_0^{(k)} = B_0((k+1)h) - B_0(kh)$.

Simulation of the CIR

Recall

$$Y_1 = Y_0 \exp(\int_0^1 r_s ds)$$

We use Rieman integral approximation of $\int_0^1 r_s ds$:

$$\hat{I}_h = \frac{1}{N} \sum_{k=1}^{N} \hat{r}_{kh}$$

and define

$$\hat{Y}_1^{(h)} := Y_0 \exp(\hat{I}_h) = Y_0 \exp\left(\frac{1}{N} \sum_{k=1}^{N} \hat{r}_{kh} \right). \tag{7}$$

Results on the approximation scheme

(Alfonsi (2013))

Assume $2ab > \sigma^2$ then for any $p \in [1, \frac{2ab}{\sigma^2})$,

$$\left(\mathbb{E}\left[\max_{0\leq k\leq N} |\hat{r}_{kh} - r_{rh}|^p\right]\right)^{1/p} \leq K_p\sqrt{h}$$

We need to control

$$\Delta_h = \int_0^1 r_s ds - \frac{1}{N}\sum_{k=1}^{N} \hat{r}_{kh}.$$

(Proposition)

Assume $ab > \sigma_0^2$ and $a > 2\sqrt{2}\sigma_0$, then

$$\mathbb{E}\left(|\Delta_h|^2\right) \leq C\sqrt{h} \qquad \mathbb{E}\left(e^{q|\delta_h|}\right) < \infty \quad \forall q < 2$$

Results on the approximation scheme

We have constructed $\hat{Z}_1 = (\hat{Y}_1^{(h)}, S_1^1, \cdots S_1^{m'})$ using a grid $(kh)_{\{k \leq N\}}$ on $[0, 1]$.

(Approximation results)

- A constant C exists (dependent on the CIR parameters) such that :

$$\mathcal{W}_1(\mathcal{L}(\hat{Z}_1), \mathcal{L}(Z_1)) = \mathcal{W}_1(\mathcal{L}(\hat{Y}_1^{(h)}), \mathcal{L}(Y_1)) \leq C\sqrt{h}.$$

- For any $\epsilon > 0$, there exists a constant K_ϵ independent of h and m such that :

$$\left\| \mathbb{E}\left[\langle Z_1, u \rangle 1_{\langle Z_1, u \rangle \geq \theta} - \langle \hat{Z}_1, u \rangle 1_{\langle \hat{Z}_1, u \rangle \geq \theta} \right] \right\|_2 \leq K_\epsilon \sqrt{m} e^{\frac{\{\sigma^+\}^2 m^2}{4\epsilon^2}} h^{\frac{1}{6} - \epsilon}.$$

SMD at fixed time horizon n

Recall that we assume

- a constant step-size sequence $\eta_k = \eta$ for $1 \leq k \leq n$
- a constant discretizaton step-size h

Then h should be chosen as :

$$h^{1/4} + h^{\frac{1}{6}-\epsilon} \sim n^{-1/2},$$

which entails that we could choose a discretization step-size close to n^{-3}.

SMD with decreasing step-size sequence

The second interesting case is choosing

$$\eta_k = k^{-\alpha}, \quad \alpha \in (\frac{1}{2}, 1] \qquad h_k = h_0^{(m)} k^{-\beta}, \quad \beta > 0.$$

This lead to

$$\delta_k = k^{-\frac{\beta}{2}} \qquad \text{and} \qquad v_k = K_\epsilon \sqrt{m} e^{\frac{\{\sigma^+\}^2 m^2}{4\epsilon^2}} h^{-\beta(\frac{1}{6}-\epsilon)}.$$

The condition for the convergence of the algorithm reads

$$\sum_{k\geq 1} k^{-\alpha} \left(k^{-\frac{\beta}{4}} + k^{-\beta(\frac{1}{6}-\epsilon)} \right) < \infty,$$

which is equivalent to :

$$\alpha + \frac{\beta}{6} > 1, \qquad \text{with} \quad \alpha \in (\frac{1}{2}, 1].$$

SMD with decreasing step-size sequence

The second interesting case is choosing

$$\eta_k = k^{-\alpha}, \quad \alpha \in (\frac{1}{2}, 1) \qquad h_k = h_0^{(m)} k^{-\beta}, \quad \beta > 0.$$

This lead to

$$\delta_k = k^{-\frac{\beta}{2}} \qquad \text{and} \qquad v_k = K_\epsilon \sqrt{m} e^{\frac{\{\sigma^+\}^2 m^2}{4\epsilon^2}} k^{-\beta(\frac{1}{6}-\epsilon)}.$$

In this case, we can obtain the finite horizon controls

$$\mathbb{E}\left[p_\lambda\left(\hat{X}_n^\eta\right)\right] - p_\lambda(x_\lambda^\star) \lesssim n^{\alpha-1}\mathrm{D}_\Phi^0 + n^{-\alpha\wedge\frac{\beta}{6}}.$$

SMD with decreasing step-size sequence

The second interesting case is choosing

$$\eta_k = k^{-\alpha}, \quad \alpha \in (\frac{1}{2}, 1) \qquad h_k = h_0^{(m)} k^{-\beta}, \quad \beta > 0.$$

This lead to

$$\delta_k = k^{-\frac{\beta}{2}} \qquad \text{and} \qquad v_k = K_\epsilon \sqrt{m} e^{\frac{\{\sigma^+\}^2 m^2}{4\epsilon^2}} k^{-\beta(\frac{1}{6}-\epsilon)}.$$

In this case, we can obtain the finite horizon controls

$$\mathbb{E}\left[p_\lambda\left(\hat{X}_n^\eta\right)\right] - p_\lambda(x_\lambda^\star) \lesssim n^{\alpha-1} \mathrm{D}_\Phi^0 + n^{-\alpha \wedge \frac{\beta}{6}}.$$

Choosing $\alpha = 1/2$ and $\beta > 3$ allows to recover the convergence rate in $\sqrt{n}$.

Simulated data

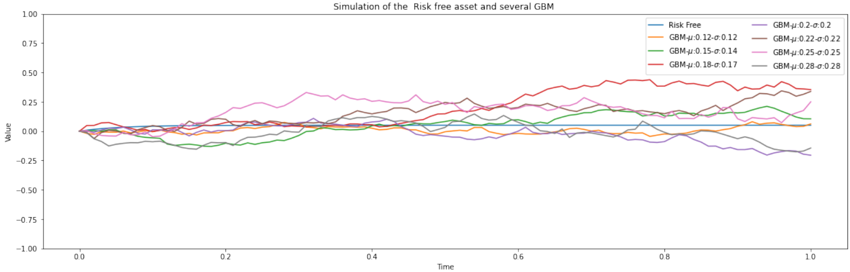


Figure – Time evolution of the return of the discretized trajectories associated to the assets of the synthetic portfolio (CIR + GBM).

Simulated data

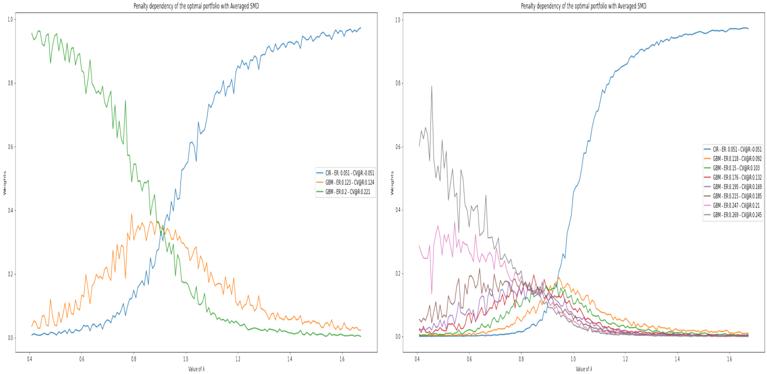


Figure – Composition of the optimal portfolio when λ increases. Left : 3 assets, Right : 8 assets. ER and CV@R are indicated in the legend box of each subfigure.

Some references

- M.C., S. Gadat and L. Huang *Portfolio optimization under CV@R constraint with stochastic mirror descent*. Preprint hal 2022.
- A. Beck and M.Teboulle , *Mirror descent and nonlinear projected subgradient methods for convex optimization.* Operations Research Letters, 2003.
- G. Lan, A. Nemirovskij, and A. Shapiro.*Validation analysis of mirror descent stochastic approximation method.* Math. Program., Ser. A, 2012.,
- Z. Zhou, P. Mertikopoulos, N. Bambos, S. Boyd, and P. Glynn. *Stochastic mirror descent in variationally coherent optimization problems*. Advances in Neural Information Processing Systems, 30, 2017.
- A. Alfonsi. *On the discretization schemes for the cir (and bessel squared) processes*. Monte Carlo Methods and Applications, 2005.
- A. Alfonsi. *Strong order one convergence of a drift implicit euler scheme : Application to the cir process.* Statistics & Probability Letters, 2013.
- S. Dereich, A. Neuenkirch, and L. Szpruch. *An Euler-type method for the strong approximation of the cox ingersoll ross process.* Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences, 2012.

Thank you for your attention !