LABORATOIRE DE PROBABILITÉS STATISTIQUE & MODÉLISATION

A Multilevel Stochastic Approximation Algorithm for Value-at-Risk and Expected Shortfall Estimation

Azar Louzi (joint work with Stéphane Crépey and Noufel Frikha)

LPSM, CNRS, Université Paris Cité

Introduction

Consider the financial loss

$$\boxed{X=\mathbb{E}[\psi(Y,Z)|Y]=\Psi(Y),}\quad Y\perp\!\!\!\perp Z.$$

For a confidence level $\alpha\in(0,1)$, we want to estimate the VaR $\xi_\star$ and the ES $\chi_\star$:

$$\boxed{\mathbb{P}(X\le\xi_\star)=\alpha,\quad \chi_\star=\mathbb{E}[X|X>\xi_\star].}$$

Stochastic Approximation Approach when Ψ Is Analytical

According to [5],

$$\boxed{\min_{\xi\in\mathbb{R}}\Big\{V(\xi)=\xi+\frac{1}{1-\alpha}\mathbb{E}[(X-\xi)^+]\Big\},\quad \xi_\star\in\arg\min V,\quad \chi_\star=\min V.}$$

We take the step size $\gamma_n=\frac{\gamma_1}{n^\beta}$, $\gamma_1>0$, $\beta\in(0,1]$.

As in [1], we consider the algorithm

$$\boxed{\xi_{n+1}=\xi_n-\gamma_{n+1}H_1(\xi_n,X^{(n+1)}),\quad \chi_0=0,\quad \chi_{n+1}=\chi_n-\frac{1}{n+1}H_2(\chi_n,\xi_n,X^{(n+1)}),}$$

where $\xi_0\perp\!\!\!\perp(X^{(n)})_{n\ge1}\overset{\text{iid}}{\sim}X$ s.t. $\mathbb{E}[|\xi_0|^2]<\infty$, and

$$H_1(\xi,x)=1-\frac{\mathbf{1}_{x>\xi}}{1-\alpha},\quad H_2(\chi,\xi,x)=\chi-\Big\{\xi+\frac{1}{1-\alpha}(x-\xi)^+\Big\}.$$

Computationally, for a prescribed accuracy $\varepsilon^2>0$,

$$\boxed{\mathrm{Cost}_{\mathsf{SA}}=C\varepsilon^{-2}.}$$

However Ψ is not always analytical to be able to simulate directly $X=\Psi(Y)$.

Nested Stochastic Approximation Approach Otherwise

According to [2], we can approximate $X=\mathbb{E}[\psi(Y,Z)|Y]$ by

$$\boxed{X_h=\frac{1}{K}\sum_{k=1}^K\psi(Y,Z^{(k)}),}\quad h=\frac{1}{K}\in\mathcal{H}=\Big\{\frac{1}{K'},K'\ge1\Big\},\quad (Z^{(k)})_{1\le k\le K}\overset{\text{iid}}{\sim}Z\perp\!\!\!\perp Y.$$

We approximate the original optimization problem by

$$\boxed{\min_{\xi\in\mathbb{R}}\Big\{V_h(\xi)=\xi+\frac{1}{1-\alpha}\mathbb{E}[(X_h-\xi)^+]\Big\},\quad \xi_\star^h\in\arg\min V_h,\quad \chi_\star^h=\min V_h.}$$

We devise the algorithm

$$\boxed{\xi_{n+1}^h=\xi_n^h-\gamma_{n+1}H_1(\xi_n^h,X_h^{(n+1)}),\quad \chi_0^h=0,\quad \chi_{n+1}^h=\chi_n^h-\frac{1}{n+1}H_2(\chi_n^h,\xi_n^h,X_h^{(n+1)}),}$$

where $\xi_0^h\perp\!\!\!\perp(X_h^{(n)})_{n\ge1}\overset{\text{iid}}{\sim}X_h$ s.t. $\mathbb{E}[|\xi_0^h|^2]<\infty$.

Theorem 1. The global error is majorized by

$$\mathbb{E}[(\xi_n^h-\xi_\star)^2]+\mathbb{E}[(\chi_n^h-\chi_\star)^2]\le C(h^2+\gamma_n).$$

For a prescribed accuracy $\varepsilon^2>0$, one has to choose

$$h=\varepsilon,\quad n=\lceil\varepsilon^{-\frac{2}{\beta}}\rceil\quad\Rightarrow\quad \mathrm{Cost}_{\mathsf{Nested\ SA}}=C\frac{n}{h}=C\varepsilon^{-\frac{2}{\beta}-1},$$

which is optimal when $\beta=1$:

$$\boxed{\mathrm{Cost}_{\mathsf{Nested\ SA}}=C\varepsilon^{-3}.}$$

The implied nested Monte Carlo computational cost should be countered.

Multilevel Acceleration of the Latter

Let $h_0=\frac{1}{K}\in\mathcal{H}$ and $M>1$. Using telescopic summation,

$$\xi_\star^{h_L}=\xi_\star^{h_0}+\sum_{\ell=1}^L\xi_\star^{h_\ell}-\xi_\star^{h_{\ell-1}},\quad \chi_\star^{h_L}=\chi_\star^{h_0}+\sum_{\ell=1}^L\chi_\star^{h_\ell}-\chi_\star^{h_{\ell-1}},\quad h_\ell=\frac{h_0}{M^\ell}\in\mathcal{H}.$$

As in [4] and [3], we consider $N_0\ge\cdots\ge N_L\ge1$ and define

$$\boxed{\xi_{\mathrm{ML}}^L=\xi_{N_0}^{h_0}+\sum_{\ell=1}^L\xi_{N_\ell}^{h_\ell}-\xi_{N_\ell}^{h_{\ell-1}},\quad \chi_{\mathrm{ML}}^L=\chi_{N_0}^{h_0}+\sum_{\ell=1}^L\chi_{N_\ell}^{h_\ell}-\chi_{N_\ell}^{h_{\ell-1}}.}$$

Each level $0\le\ell\le L$ is simulated independently.
Within each level $1\le\ell\le L$, X_{h_ℓ} and $X_{h_{\ell-1}}$ are perfectly correlated:

$$X_{h_{\ell-1}}=\frac{1}{KM^{\ell-1}}\sum_{k=1}^{KM^{\ell-1}}\psi(Y,Z^{(k)}),\quad \boxed{X_{h_\ell}=\frac{1}{M}X_{h_{\ell-1}}+\frac{1}{KM^\ell}\sum_{k=KM^{\ell-1}+1}^{KM^\ell}\psi(Y,Z^{(k)}).}$$

Assumption 1. We consider 3 different frameworks:

1. $\exists p_*>1$, $\mathbb{E}\big[\big|\psi(Y,Z)-\mathbb{E}[\psi(Y,Z)|Y]\big|^{p_*}\big]<\infty$.
2. $\exists C>0$, $\forall\lambda\in\mathbb{R}$, $\mathbb{E}\big[\exp\big(\lambda\big(\psi(Y,Z)-\mathbb{E}[\psi(Y,Z)|Y]\big)\big)\big]\le\exp(C\lambda^2)$.
3. $G_\ell=h_\ell^{-\frac{1}{2}}(X_{h_\ell}-X_{h_{\ell-1}})$ and $K_\ell=\|F_{X_{h_{\ell-1}}|G_\ell}\|_{\mathrm{Lip}}$ satisfy $\sup_{\ell\ge1}\mathbb{E}[K_\ell|G_\ell|]<\infty$.

Theorem 2. The global error is majorized by

$$\mathbb{E}[(\xi_{\mathrm{ML}}^L-\xi_\star)^2]\le C\Big(h_L^2+\sum_{\ell=0}^L\gamma_{N_\ell}\varphi(h_\ell)\Big),\quad \mathbb{E}[(\chi_{\mathrm{ML}}^L-\chi_\star)^2]\le C\Big(h_L^2+\sum_{\ell=0}^L\frac{h_\ell}{N_\ell}\Big),$$

where $\varphi(h)=h^{\frac{p_*}{2(1+p_*)}}$ (Asp 1.1), $h^{\frac{1}{2}}|\ln h|^{\frac{1}{2}}$ (Asp 1.2), $h^{\frac{1}{2}}$ (Asp 1.3).
To achieve a prescribed accuracy $\varepsilon^2>0$, one must choose

$$h_L=\frac{h_0}{M^L}=\varepsilon\quad\Leftrightarrow\quad L=\left\lceil\frac{\ln\left(\frac{h_0}{\varepsilon}\right)}{\ln M}\right\rceil.$$

VaR Focused Parametrization. We optimize

$$\begin{array}{ll}\min_{N_0,\dots,N_L>0} & \mathrm{Cost}_{\mathsf{MLSA}}=C\sum_{\ell=0}^L\frac{N_\ell}{h_\ell},\\ \text{subject to} & \sum_{\ell=0}^L\gamma_{N_\ell}\varphi(h_\ell)=\varepsilon^2.\end{array}$$

We infer the optimal iterations

$$N_\ell^{\mathrm{VaR}}=\left\lceil\varepsilon^{-\frac{2}{\beta}}\Big(\sum_{\ell'=0}^L h_{\ell'}^{-\frac{\beta}{1+\beta}}\varphi(h_{\ell'})^{\frac{1}{1+\beta}}\Big)^{\frac{1}{\beta}}h_\ell^{\frac{1}{1+\beta}}\varphi(h_\ell)^{\frac{1}{1+\beta}}\right\rceil\quad\Rightarrow\quad\mathrm{Cost}_{\mathsf{MLSA}}^{\mathsf{VaR}}=C\varepsilon^{-\frac{2}{\beta}-1}\varphi(\varepsilon),$$

which is minimal when $\beta=1$:

$$\boxed{\mathrm{Cost}_{\mathsf{MLSA}}^{\mathsf{VaR}}=C\varepsilon^{-\frac{5p_*+6}{2(1+p_*)}}\ \text{(Asp 1.1)},\quad C\varepsilon^{-\frac{5}{2}}|\ln\varepsilon|^{\frac{1}{2}}\ \text{(Asp 1.2)},\quad C\varepsilon^{-\frac{5}{2}}\ \text{(Asp 1.3)}.}$$

ES Focused Parametrization. We optimize

$$\begin{array}{ll}\min_{N_0,\dots,N_L>0} & \mathrm{Cost}_{\mathsf{MLSA}}=C\sum_{\ell=0}^L\frac{N_\ell}{h_\ell},\\ \text{subject to} & \sum_{\ell=0}^L\frac{h_\ell}{N_\ell}=\varepsilon^2.\end{array}$$

We infer the optimal iterations

$$N_\ell^{\mathrm{ES}}=\lceil\varepsilon^{-2}Lh_\ell\rceil=\left\lceil\varepsilon^{-2}\left\lceil\frac{\ln\left(\frac{h_0}{\varepsilon}\right)}{\ln M}\right\rceil h_\ell\right\rceil.$$

The achieved complexity is

$$\boxed{\mathrm{Cost}_{\mathsf{MLSA}}^{\mathsf{ES}}=C\varepsilon^{-2}|\ln\varepsilon|^2.}$$

Financial Case Study: Swap

Model Setting. Consider a swap on a Black-Scholes rate:

$$P_t=\bar{N}\,\mathbb{E}\Big[\sum_{i=i_t}^d\mathrm{e}^{-r(T_i-t)}\Delta T_i(S_{T_{i-1}}-\bar{S})\Big|\mathcal{F}_t\Big].$$

Set $\alpha=85\%$. Define the loss (for a short position issued at par)

$$X=\mathrm{e}^{-r\tau}P_\tau.$$

Approximation. The loss can be directly simulated:

$$X\overset{\mathcal{L}}{=}A\Big(\exp\Big(-\frac{\sigma^2\tau}{2}+\sigma\sqrt{\tau}Y\Big)-1\Big),\quad Y\sim\mathcal{N}(0,1),\quad A=\bar{N}S_0\sum_{i=2}^d\mathrm{e}^{-rT_i}\Delta T_i\mathrm{e}^{\kappa T_{i-1}}.$$

The loss can also be written

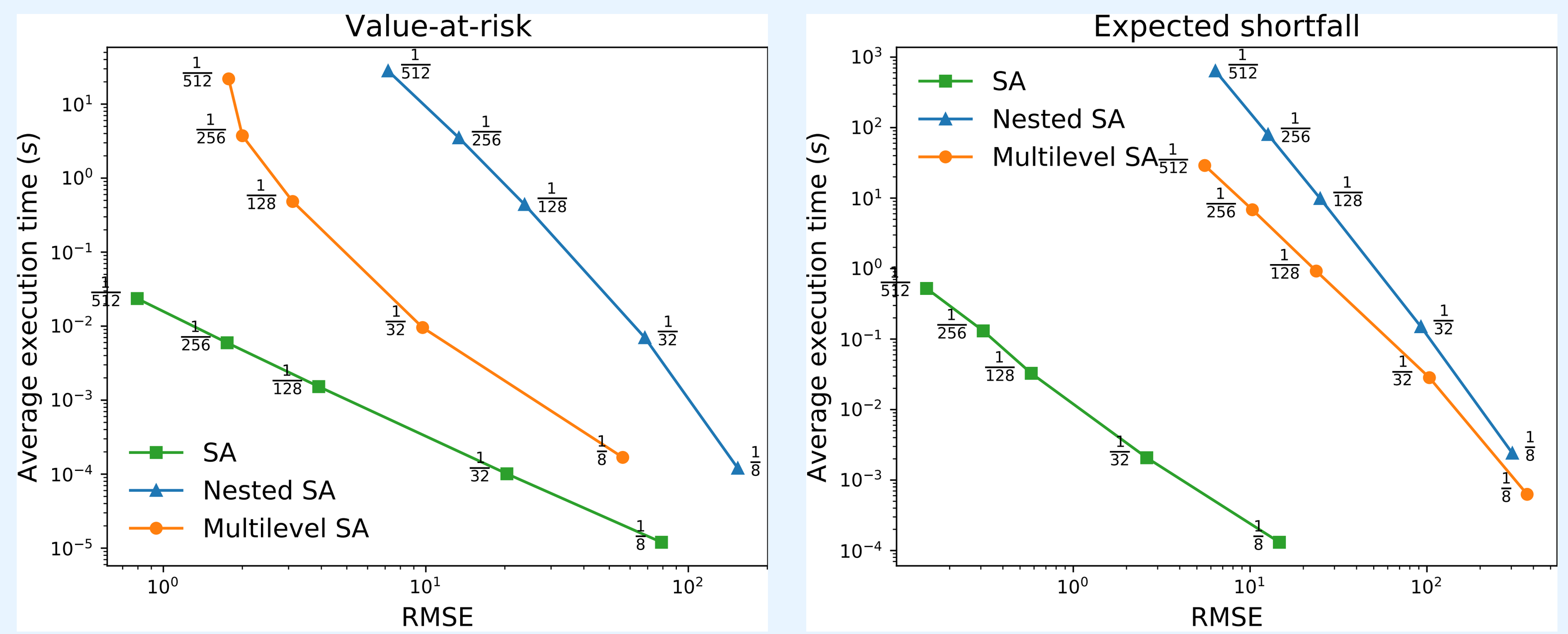
$$X\overset{\mathcal{L}}{=}\mathbb{E}[\psi(Y,Z)|Y],\quad Y\perp\!\!\!\perp Z=(Z_1,\dots,Z_{d-1}),$$

$$Y=\exp\Big(-\frac{\sigma^2\tau}{2}+\sigma\sqrt{\tau}U_0\Big),\quad Z_1=\exp\Big(-\frac{\sigma^2(T_1-\tau)}{2}+\sigma\sqrt{T_1-\tau}\,U_1\Big),$$

$$Z_i=\exp\Big(-\frac{\sigma^2\Delta T_i}{2}+\sigma\sqrt{\Delta T_i}\,U_i\Big),\quad U_0,\dots,U_{d-1}\overset{\text{iid}}{\sim}\mathcal{N}(0,1),$$

$$\psi(y,z)=\bar{N}S_0\sum_{i=2}^d\mathrm{e}^{-rT_i}\Delta T_i\mathrm{e}^{\kappa T_{i-1}}\Big(y\prod_{j=1}^{i-1}z_j-1\Big).$$

Comparative Complexity Study. We work under Asp 1.1 with $p_\star=8$.



Conclusion

For a prescribed accuracy $\varepsilon^2>0$,

$$\boxed{\mathrm{Cost}_{\mathsf{SA}}=C\varepsilon^{-2}\ll\begin{cases}\mathrm{Cost}_{\mathsf{MLSA}}^{\mathsf{VaR}}=C\varepsilon^{-2-\delta},\ \delta<1\\ \mathrm{Cost}_{\mathsf{MLSA}}^{\mathsf{ES}}=C\varepsilon^{-2}|\ln\varepsilon|^2\end{cases}\ll\mathrm{Cost}_{\mathsf{Nested\ SA}}=C\varepsilon^{-3}.}$$

References

[1] O. Bardou, N. Frikha, and G. Pagès. Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling. *Monte Carlo Methods and Applications*, 15(3):173–210, 2009.

[2] D. Barrera, S. Crépey, B. Diallo, G. Fort, E. Gobet, and U. Stazhynski. Stochastic Approximation Schemes for Economic Capital and Risk Margin Computations. *ESAIM: Proceedings and Surveys*, 65:182–218, 2019.

[3] N. Frikha. Multi-level stochastic approximation algorithms. *The Annals of Applied Probability*, 26(2):933 – 985, 2016.

[4] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[5] R. T. Rockafellar and S. Uryasev. Optimization of Conditional Value-at-Risk. *Journal of risk*, 2(3):21–41, 2000.