A variance reduction method for computing VaR

1. Computing Value at Risk by Monte Carlo simulations

2. Importance Sampling for variance reduction

3. Interacting Particle Systems for Importance Sampling (IPS-IS)

4. Simulation results

Nadia Oudjane - EDF R&D- Journées MAS 2008
1. Computing Value at Risk by Monte Carlo simulations

**Value at Risk and quantile**

- **P&L of a portfolio on** $[0, T]$  
  \[ \Delta V(X) = V_T - V_0 + \int_0^T CF \]

  with \( X \in \mathbb{R}^d \) the risk factors impacting the portfolio value on \([0, T]\)

- **Value at Risk**  
  \[ \text{VaR}_\alpha = \left| \inf \{ s \in \mathbb{R} \mid \mathbb{P}(\Delta V \leq s) \geq 1 - \alpha \} \right| \]

\[ \text{VaR}_\alpha = |F^{-}(\alpha)|, \quad \text{where} \quad F(s) = \mathbb{P}(\Delta V \leq s), \quad \text{for all} \ s \in \mathbb{R} \]
Monte Carlo method for VaR estimation

The distribution function $F$ can be viewed as an expectation

$$F(s) = \mathbb{E}[I_{\Delta V(X) \leq s}] , \quad \text{for all} \quad s \in \mathbb{R}$$

Traditional Monte Carlo Method for computing VaR

1. Monte Carlo simulations give an approximation of $F(s)$:

$$\hat{F}_N(s) = \frac{1}{N} \sum_{i=1}^{N} I(\Delta V(X_i) \leq s) , \quad \text{for all} \quad s \in \mathbb{R}$$

$\Rightarrow$ Too many evaluations of $\Delta V$ for a given accuracy

2. Inversion of $\hat{F}_N$ and interpolation for approximating VaR
Importance Sampling for variance reduction

\[ m = \mathbb{E}_p[H(X)] = \mathbb{E}_q[H(Y) \frac{p(Y)}{q(Y)}], \quad \text{where } X \sim p \quad \text{and} \quad Y \sim q \]

\[ q^* = \frac{H_p}{\int H(x)p(x) \, dx} = \frac{H_p}{\mathbb{E}_p[H(X)]} = H \cdot p \]

Monte Carlo approximation

\[ \mathbb{E}_p[H(X)] \approx \hat{m}_M^q = \frac{1}{M} \sum_{i=1}^{M} H(Y_i) \frac{p(Y_i)}{q(Y_i)}, \quad \text{where } (Y_1, \cdots, Y_M) \text{ i.i.d. } \sim q \]

⇒ How to simulate and evaluate approximately \( q^* \) ?
Variance of the Importance Sampling estimate

Let $q$ be a (possibly random) importance probability density dominating $q^*$

$$\text{Var}(\hat{m}_M^q) = \mathbb{E}\left[\text{Var}\left(\hat{m}_M^q \mid \mathcal{F}_q\right)\right] + \text{Var}\left[\mathbb{E}\left[\hat{m}_M^q \mid \mathcal{F}_q\right]\right] = 0$$

$\mathcal{F}_q$ denotes the sigma-algebra generated by the random variables involved in $q$

The variance of the IS estimate depends on the "distance" between $q$ and $q^*$

$$\text{Var}(\hat{m}_M^q) = \frac{m^2}{M} \mathbb{E}\left[\int \left((q^* - q) \frac{q^*}{q}\right)(x) dx\right]$$

Idea: use Interacting Particle Systems for Importance Sampling (IPS-IS) to approximate $q^*$ by $q^N$ based on an $N$-particle system to achieve

$$\text{Var}(\hat{m}_M^{q^N}) \leq \frac{C}{MN^\alpha} \quad \text{with} \quad 0 < \alpha < 1/2$$
Some alternative approaches

- **Large deviation approximation** for rare events simulation
- **Approximation of** $H$ **to obtain a simple form for** $q^*$

  \[ \text{ex: } [\text{Glasserman\&al00}] \text{ for computing VaR, } \Delta\Gamma \text{ approximation of the portfolio} \]

- **Cross-entropy**  \[ [\text{Homem-de-Mello\&Rubinstein02}] \]

  $q^\theta$ is chosen in a parametric family such as to minimize the entropy $K(q^\theta, q^*)$

- **Interacting Particle Systems** without Importance Sampling

  \[ [\text{DelMoral\&Garnier05}], [\text{Cerou\&al06}] \]

  Interacting Particle Systems for Importance Sampling (IPS-IS) can be viewed as a non parametric version of cross entropy approach
Progressive correction [Musso&al01]

- We introduce a sequence of non negative functions \((G_k)_{0 \leq k \leq n}\) such that

\[
\begin{cases}
G_0(x) = 1 \\
\text{for all } x \in \mathbb{R}^d, \\
\text{The product } G_0(x) \cdots G_n(x) = H(x) \\
\text{If } G_k(x) = 0 \text{ then } G_{k+1}(x) = 0
\end{cases}
\]

- In our case \(H(x) = I(\Delta V(x) \leq s)\) then we choose

\[
G_k(x) = I_{\Delta V(x) \leq s_k} \quad \text{with} \quad s = s_n \leq \cdots \leq s_0 = +\infty
\]

- Dynamical system on the space of probability measures \((\nu_k)_{0 \leq k \leq n}\)

\[
\begin{cases}
\nu_0 = p \, dx \\
\nu_k = \frac{G_k \nu_{k-1}}{\int_{\mathbb{R}^d} G_k(x) \nu_{k-1}(x) \, dx} = G_k \cdot \nu_{k-1}, \quad \text{for all } 1 \leq k \leq n
\end{cases}
\]

\[
\Rightarrow \nu_n = q^* \, dx
\]
2. Importance Sampling for variance reduction

Space exploration

- We introduce a sequence of Markov kernels \( (Q_k)_{0 \leq k \leq n} \) such that

\[
\nu_k \approx \nu_k Q_k \quad \text{i.e.} \quad \nu_k(dx) \approx \int_{\mathbb{R}^d} \nu_k(du) Q_k(u, dx), \quad \text{for all} \quad x \in \mathbb{R}^d
\]

- In our case where \( G_k(x) = I_{\Delta V(x) \leq s_k} \), if \( p \) is Gaussian then \( Q_k \) is easily obtained from a Gaussian kernel \( Q \) reversible for \( p \),

\[
Q_k(x, dx') = Q(x, dx') I_{\Delta V(x) \leq s_k} + \left[ 1 - Q(x, \Delta V^-((-\infty, s_k])) \right] \delta_x(dx')
\]

- Dynamical system on the space of probability measures \( (\nu_k)_{0 \leq k \leq n} \)

\[
\begin{cases}
\nu_0 = p \, dx \\
\nu_k = G_k \cdot (\nu_{k-1} Q_{k-1}) , \quad \text{for all} \quad 1 \leq k \leq n
\end{cases}
\]

\[ \Rightarrow \nu_n = q^* \, dx \]
Approximation of the dynamical system

- The idea is to replace at each iteration $k$, $\nu_{k-1} Q_{k-1}$ by its $N$-empirical measure $S^N(\nu_{k-1} Q_{k-1})$ such that

$$S^N(\nu_{k-1} Q_{k-1}) = \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i_k}$$

where $(X^1_k, \cdots, X^N_k)$ are i.i.d. $\sim \nu_{k-1} Q_{k-1}$

- Dynamical system on the space of discrete probability measures $(\nu^N_k)_{0 \leq k \leq n}$

$$\begin{cases} 
\nu^N_0 = S^N(\nu) \\
\nu^N_k = G_k \cdot S^N(\nu^N_{k-1} Q_{k-1}) \text{, for all } 1 \leq k \leq n
\end{cases}$$

⇒ One can show that $\nu^N_n \approx q^* dx$ [DelMoral]
Algorithm

★ **Initialization** : Generate independently

\[
(X_1^0, \ldots, X_N^0) \text{ i.i.d. } \sim p \quad \text{then set} \quad \nu_0^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i^0}
\]

★ **Selection** : Generate independently

\[
(\tilde{X}_1^1, \ldots, \tilde{X}_N^1) \text{ i.i.d. } \sim \nu_k^N = \sum_{i=1}^{N} \omega_k^i \delta_{X_k^i}
\]

★ **Mutation** : Generate independently for each \( i \in \{1, \ldots, N\} \),

\[
X_k^{i+1} \sim Q_k(\tilde{X}_k^i, \cdot)
\]

★ **Weighting** : For each particle \( i \in \{1, \ldots, N\} \), compute

\[
\omega_{k+1}^i = \frac{G_{k+1}(X_{k+1}^i)}{\sum_{j=1}^{N} G_{k+1}(X_{k+1}^j)}
\]

then set \( \nu_{k+1}^N = \sum_{i=1}^{N} \omega_{k+1}^i \delta_{X_{k+1}^i} \)
Adaptive choice of the sequence $(G_k)_{0 \leq k \leq n}$

[Musso&al01], [Hommem-de-Mello&Rubinstein02], [Cérou&al06]

- The performance of Interacting particle systems is known to deteriorate when the quantities $\max G_k$ are big

The idea is then to chose $G_k$ such that $\frac{1}{N} \sum_{i=1}^{N} G_k(X^i_k)$ is not to small

- In our case where $G_k(x) = I_{\Delta V(x) \leq s_k}$, the threshold $s_k$ is chosen as a r.v. depending on the current particle system and on a parameter $\rho \in (0, 1)$:

$$s_k = \inf \left\{ s \text{ such that } \sum_{i=1}^{N} I_{\Delta V(X^i) \leq s} \geq \rho N \right\}$$

- This choice of $s_k$ is not proved to guarantee that the algorithms ends in a finite number of iterations but this point does not seem to be a problem in our simulations
At the end of the algorithm, we get \( \nu_n \approx q^* \, dx \)

But Importance Sampling requires a smooth approximation with density \( q^N \)

Kernel of order 2 \( K \)

\[
K \geq 0 \quad \int K = 1 \quad \int x_i K = 0 \quad \int |x_i x_j| \, K < \infty
\]

Rescaled kernel \( K_h \)

\[
K_h(x) = \frac{1}{h^d} K \left( \frac{x}{h} \right)
\]

\[
\nu^N = \sum \omega^i \delta_{X^i} \rightarrow \quad q^{N,h} = \sum \omega^i K_h(\cdot - X^i)
\]

Optimal choice of \( h \) \( \Rightarrow \)

\[
\mathbb{E} \|q^N - q^*\|_1 \leq \frac{C_4}{N^{2(d+4)} + 4}
\]
### Some simulation results: Variance ratio

- Several test cases depending on the form of function $x \mapsto \Delta V(x)$ have been studied: results are all comparable.
- $X$ is a $d$ dimensional Gaussian variable and $m = \mathbb{E}_p[I_{\Delta V(X) \leq s}]$.
- Particles $N = 500$, Iterations $n \approx 10$ to 60, Simulations $M = 10000$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$d = 1$</th>
<th>$d = 2$</th>
<th>$d = 3$</th>
<th>$d = 4$</th>
<th>$d = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-2}$</td>
<td>150 $10^{-1}$</td>
<td>50</td>
<td>50</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>1000 $2$</td>
<td>300</td>
<td>300</td>
<td>200</td>
<td>140</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$2 \cdot 10^5$</td>
<td>$10^5$</td>
<td>$10^5$</td>
<td>$5 \cdot 10^4$</td>
<td>$2 \cdot 10^4$</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>400</td>
<td>300</td>
<td>460</td>
<td>480</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$</th>
<th>$d = 6$</th>
<th>$d = 7$</th>
<th>$d = 8$</th>
<th>$d = 9$</th>
<th>...</th>
<th>$d = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-2}$</td>
<td>22</td>
<td>14</td>
<td>11</td>
<td>8</td>
<td>...</td>
<td>$5 \cdot 10^{-3}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>100</td>
<td>70</td>
<td>55</td>
<td>40</td>
<td>...</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$10^4$</td>
<td>$2 \cdot 10^3$</td>
<td>$2 \cdot 10^3$</td>
<td>$4 \cdot 10^3$</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>480</td>
<td>300</td>
<td>300</td>
<td>...</td>
<td>360</td>
</tr>
</tbody>
</table>
References


