Marginal quantization of an Euler diffusion process and its application

Abass SAGNA
abass.sagna@ensiie.fr

Laboratoire Analyse et Probabilités de l’UEVE,
ENSIIE

Advances In Financial Mathematics, January, 07-10, 2014
Applications of optimal quantization

1. Optimal Quantization (OQ) has known its expansion in Information Theory in the early 40’s.
2. Its common use is the conversion of a continuous (or big) set of data into finite (or small) set of data, in an optimal way.
2. In Numerical Probability, where the seminal work go back to [Pagès, (1998)], this means approximating a (continuous) random variable into r.v. valued in a set of finite cardinality.
**Description of the problem**

In almost all these problems we have to estimate quantities like (for a given function $f : \mathbb{R}^d \mapsto \mathbb{R}$)

$$\mathbb{E}[f(X_T)], \quad T > 0,$$

(1)

or involving terms like

$$\mathbb{E}[f(X_t)|X_s], \quad 0 < s < t,$$

(2)

where $(X_t)_{t \in [0,T]}$ is the solution of the SDE

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}^d,$$

(3)

where

- $W$ is a standard $d$-dimensional Brownian motion starting at 0
- $b : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$; $\sigma : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^d$ are measurable and satisfy some appropriate conditions which ensure the existence of a strong solution of the SDE.
Example. How to estimate

- fast
- with a good precision

the quantity

$$\mathbb{E}[f(X_T)], \text{ with } f(x) = e^{-rT}\mathbb{E}(K - x)^+, \text{ where}$$

$$dX_t = rX_t dt + \vartheta \frac{X_t^{\delta+1}}{\sqrt{1 + X_t^2}} dW_t, \quad X_0 = x_0 \in \mathbb{R},$$

(4)

for $\delta \in (0, 1)$ and $\vartheta \in (0, \vartheta)$ with $\vartheta > 0$, $r$ is the interest rate.
Methods for solving the problem

In a general setting we first have to approximate the continuous paths of \((X_t)_{t \in [0,T]}\) by discrete "Euler" paths process \((\tilde{X}_{tk})_k\), at time discretization steps \(t_k = k\Delta, \ k = 0, \ldots, n, \ \Delta = T/n\): Then, the quantity (1) e.g. is estimated by

\[
\mathbb{E}[f(\tilde{X}_{tn})]
\]  

How to estimate (5) by OQ? by \((x^{N_n} = \{x_1^{N_n}, \ldots, x_{N_n}^{N_n}\})\)

\[
\mathbb{E}[f(\hat{X}_{tn}^{x^{N_n}})] = \sum_{i=1}^{N_n} f(x_i^{N_n}) \mathbb{P}(\hat{X}_{tn}^{x^{N_n}} = x_i^{N_n})
\]  

How to get the optimal grids \(x^{N_n}\) and the associated weights?

- Up to now, in the framework of our given example, the only way to get the optimal grids and the associated weights is carrying out stochastic or Lloyd’s type algorithms, even for \(d = 1\).
Goal. Quantize \((\tilde{X}_{t_k})_k\) from a recursive procedure involving the distribution of \(\tilde{X}_{t_{k+1}}|\tilde{X}_{t_k}\) and use Newton algorithm to compute numerically the optimal grids when \(d = 1\).

Advantages of the method.

1. more accurate and in far more fast than stochastic algorithms.
2. \(\forall k, \hat{X}^{x_{N_k}}_{t_k}\) is stationary.
3. The process \((\hat{X}^{x_{N_k}}_{t_k})_k\) remains a Markov chain.
4. simulations show that it may be more competitive than MC.

The main idea is the following result:

\[
D_{k+1}(x^{N_{k+1}}) = \int_{\mathbb{R}^d} \mathbb{E}\left( \min_{j=1,...,N_{k+1}} |Y_k(x) - x_{j}^{N_{k+1}}|^2 \right) \mathbb{P}(\tilde{X}_{t_k} \in dx),
\]

\(D_{k+1}(\cdot)\) is the distortion function, \(Y_k(x) \sim \mathcal{N}(m_k(x); \Sigma_k(x))\).
Marginal quantization

We have to solve

$$\arg \min \{ D_{k+1}(x^{N_k+1}), \quad x^{N_k+1} \in \mathbb{R}^{dN_k+1} \}.$$  

In practice we will solve the problem

$$\arg \min \{ \hat{D}_{k+1}(x^{N_k+1}), \quad x^{N_k+1} \in \mathbb{R}^{dN_k+1} \}.$$  

**Theorem.** Let $\theta : \mathbb{R}^d \to \mathbb{R}_+$ be a nonnegative convex s.t. $
\forall k, \mathbb{E}\theta(\tilde{X}_{t_k}) < \infty and for every \ x, z \in \mathbb{R}^d, \sup_{t \in [0, T]} |\sigma\sigma^T(t, x) - \sigma\sigma^T(t, z)| \leq |x - z|(\theta(x) + \theta(z)).$  

(7)

For any sequence of optimal $N_k$-quantizer $x^{N_k}$ we have

$$|D_{k+1}(x^{N_k+1}) - \hat{D}_{k+1}(x^{N_k+1})| \leq K \|\tilde{X}_{t_k} - \hat{X}_{t_k} \|^2, \quad for \ k = 1, \ldots, n-1.$$  

for some real constant $K > 0.$
Figure: ("Black-Scholes model ") \( dX_t = rX_t dt + \sigma X_t dW_t, \; X_0 = 86.3, \)
\( r = 0.03, \; \sigma = 0.05. \) Abscissa axis: the optimal grids, \( \hat{X}_{tk} = x^i_k, \; t_k = k \Delta, \)
\( \Delta = 0.02, \; k = 1, \ldots, 25, \; i = 1, \ldots, N_k. \) Ordinate axis: \( \mathbb{P}(\hat{X}_{tk} = x^i_k). \) \( \hat{X}_{t1} \) is depicted in dots '•', \( \hat{X}_{t25} \) is represented by the symbol '⋆', \( t_1 = 0.02 \) and
Figure: ("Pseudo-CEV model") \( dX_t = rX_t \, dt + \vartheta (X_t^{\delta+1}/(1 + X_t^2)^{-1/2}) \, dW_t \), \( X_0 = 100, \ r = 0.15, \ \vartheta = 0.7, \ \delta = 0.5 \). Abscissa axis: the optimal grids, \( \hat{X}_{tk} = x_k^i, \ t_k = k\Delta, \ \Delta = 0.02, \ k = 1, \ldots, 25, \ i = 1, \ldots, N_k \). Ordinate axis: the associated weights. \( \hat{X}_{t_1} \) is depicted in dots '•', \( \hat{X}_{t_{25}} \) is represented by the
Our aim. Compute $e^{-rT} \mathbb{E}(K - X_T)^+ = \max(K - X_T, 0)$, where

$$dX_t = rX_t dt + \vartheta \frac{X_t^{\delta + 1}}{\sqrt{1 + X_t^2}} dW_t, \quad X_0 = x_0. \quad (8)$$

We set $\delta = 0.5$, $X_0 = 100$, $r = 0.15$; $n = 120$, $N_k = 400$, $\forall k \geq 1$.

**Remark (on the computational time)**
- All the quantization grids of sizes $N_k = 400$, $k = 1, \ldots, n = 120$, and there weights are obtained in about 1 mn from the Newton algorithm with 5 iterations. Computations are performed using Scilab software on a CPU 2.7 GHz and 4 Go memory computer.
- For the Monte Carlo sample size $M = 10^7$ it takes about 2 minutes and 30 seconds to get a price using the C programming language on the same computer described previously. Then, in this situation, it takes more time to obtain a price by MC method than carrying it out by MQ.
<table>
<thead>
<tr>
<th>$K$</th>
<th>MC ($10^5$)</th>
<th>MC ($10^6$)</th>
<th>MC ($10'$)</th>
<th>MQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>08.89</td>
<td>08.81</td>
<td>08.81</td>
<td>08.81</td>
</tr>
<tr>
<td>Cl</td>
<td>[08.80; 08.99]</td>
<td>[08.78; 08.84]</td>
<td>[08.80; 08.82]</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>10.61</td>
<td>10.60</td>
<td>10.59</td>
<td>10.59</td>
</tr>
<tr>
<td>Cl</td>
<td>[10.51; 10.72]</td>
<td>[10.57; 10.63]</td>
<td>[10.58; 10.60]</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>12.53</td>
<td>12.57</td>
<td>12.57</td>
<td>12.57</td>
</tr>
<tr>
<td>Cl</td>
<td>[12.42; 12.64]</td>
<td>[12.53; 12.60]</td>
<td>[12.56; 12.58]</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>14.72</td>
<td>14.74</td>
<td>14.75</td>
<td>14.75</td>
</tr>
<tr>
<td>Cl</td>
<td>[14.60; 14.84]</td>
<td>[14.70; 14.78]</td>
<td>[14.75; 14.77]</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>17.18</td>
<td>17.10</td>
<td>17.13</td>
<td>17.12</td>
</tr>
<tr>
<td>Cl</td>
<td>[17.04; 17.31]</td>
<td>[17.06; 17.15]</td>
<td>[17.11; 17.14]</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>19.64</td>
<td>19.69</td>
<td>19.67</td>
<td>19.67</td>
</tr>
<tr>
<td>Cl</td>
<td>[19.50; 19.78]</td>
<td>[19.64; 19.73]</td>
<td>[19.65; 19.68]</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** (Pseudo-CEV model) Comparison of the Put prices obtained from MC with associated CI and from the MQ method: $r = 0.15$; $n = 120$; $N_k = 400$, $\forall k = 1, \ldots, n$; $T = 1$; $\vartheta = 4$; $X_0 = 100$. 
Comments

• In all examples the prices obtained by MQ stay in the confidence interval induced by the MC price estimates.

• The prices obtained by the MQ method are more precise (more specifically when $\vartheta = 4$ and $K$ increasing away from 100) than those obtained by the MC when $M$ equals $10^5$ or $10^6$.

• Consequently, the MQ method seems to be more efficient than the MC when the sample size is less than $10^6$.

• However, when increasing the sample size to $M = 10^7$ the two prices become closer (up to $10^{-2}$) but the MC method becomes more time consuming.