Stochastic Grid Bundling Method for Backward Stochastic Differential Equations

Ki Wai Chau

Centrum Wiskunde & Informatica

17th Winter school on Mathematical Finance
22 January 2018

(A joint work with Kees Oosterlee (CWI & TU Delft))

Centrum Wiskunde & Informatica

Backward Stochastic Differential Equations

- Settings:
  - A filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$
  - $W := (W_t)_{0 \le t \le T}$ is a d-dimensional Brownian motion adapted to $\mathbb{F}$
- Forward Backward Stochastic Differential Equation

$$\begin{cases} dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \ X_0 = x_0, \\ dY_t = -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t, \ Y_T = \Phi(X_T), \end{cases}$$

  - $\mu : \Omega \times [0, T] \times \mathbb{R}^q \to \mathbb{R}^q$ and $\sigma : \Omega \times [0, T] \times \mathbb{R}^q \to \mathbb{R}^{q \times d}$
  - $f : \Omega \times [0, T] \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$
  - $\Phi : \Omega \times \mathbb{R}^q \to \mathbb{R}$
  - Solution: (Y_t, Z_t) which satisfies the equation, adapts to $\mathbb{F}$ and satisfies some regularity requirements.

Discretization

For a given time grid $\pi = \{0 = t_0 < \ldots < t_N = T\}$, we define the backward time discretizations (Y^π, Z^π) based on the theta-scheme from [Zhao et al., 2012]:

$$\begin{aligned}
Y^\pi_{t_N} =& \Phi(X^\pi_{t_N}), \quad Z^\pi_{t_N} = \nabla \Phi(X^\pi_{t_N}) \sigma(t_N, X^\pi_{t_N}) \\
Z^\pi_{t_k} =& -\theta_2^{-1}(1-\theta_2)\mathbb{E}_{t_k}\left[Z^\pi_{t_{k+1}}\right] + \frac{1}{\Delta_k}\theta_2^{-1}\mathbb{E}_{t_k}\left[Y^\pi_{t_{k+1}}\Delta W_k^T\right] \\
&+ \theta_2^{-1}(1-\theta_2)\mathbb{E}_{t_k}\left[f_{k+1}(Y^\pi_{t_{k+1}}, Z^\pi_{t_{k+1}})\Delta W_k^T\right], \ k = N-1, \ldots, 0 \\
Y^\pi_{t_k} =& \mathbb{E}_{t_k}\left[Y^\pi_{t_{k+1}}\right] + \Delta_k \theta_1 f_k(Y^\pi_{t_k}, Z^\pi_{t_k}) \\
&+ \Delta_k(1-\theta_1)\mathbb{E}_{t_k}\left[f_{k+1}(Y^\pi_{t_{k+1}}, Z^\pi_{t_{k+1}})\right], \ k = N-1, \ldots, 0,
\end{aligned}$$

where $f_k(y, z) := f(t_k, X^\pi_{t_k}, y, z)$, $0 \le \theta_1 \le 1$ and $0 < \theta_2 \le 1$.

Discretization (cont.)

- Note that:
  - the globally Lipschitz driver assumption is in force;
  - we use a Markovian approximation $X^{\pi}_{t_k}$, $t_k \in \pi$:
    - $X^{\pi}_{t_{k+1}} = X^{\pi}_{t_k} + b(t_k, X^{\pi}_{t_k})\Delta_k + \sigma(t_k, X^{\pi}_{t_k})\Delta W_k$;
  - due to the Markovian setting, there exist functions $y_k^{(\theta_1,\theta_2)}(x)$ and $z_k^{(\theta_1,\theta_2)}(x)$ such that

$$Y^{\pi}_{t_k} = y_k^{(\theta_1,\theta_2)}(X^{\pi}_{t_k}),\ Z^{\pi}_{t_k} = z_k^{(\theta_1,\theta_2)}(X^{\pi}_{t_k}).$$

- Question:
  How to approximate $\mathbb{E}^x_{t_k}\left[y_{k+1}^{(\theta_1,\theta_2)}(X^{\pi}_{t_k+1})\right]$, $\mathbb{E}^x_{t_k}\left[y_{k+1}^{(\theta_1,\theta_2)}(X^{\pi}_{t_k+1})\Delta W_k^T\right]$, and other similar quantities along the time grid?

Stochastic Grid Bundling Method

- Non-nested Monte Carlo scheme
  - It starts with the simulation of M independent samples of $(X^{\pi}_{t_k})_{0\leq k\leq N}$, denoted by $(X^{\pi,m}_{t_k})_{1\leq m\leq M, 0\leq k\leq N}$.
  - The simulation is only performed once.
  - The terminal values for each path are:
    $y^{(\theta_1,\theta_2),R,I}_N(X^{\pi,m}_{t_N}) = \Phi(X^{\pi,m}_{t_N})$,
    $z^{(\theta_1,\theta_2),R}_N(X^{\pi,m}_{t_N}) = \nabla\Phi(X^{\pi,m}_{t_N})\sigma(t_N, X^{\pi,m}_{t_N}),\ m = 1, \ldots, M.$

Recurring steps in time (I)

- Non-nested Monte Carlo scheme
- Regress-later
  - The least-squares regression technique for functions is performed on the random variable $X^{\pi}_{t_{k+1}}$
  - Then we use the (analytical) expectation of the resulting approximation in our algorithm.
  - This will remove the "statistical" error at the regression step.
  - To ensure the stability of our algorithm, the regression coefficients must be bounded above.
  - It means that an error notion should be given by the program when the Euclidean norm of any regression coefficient vector is greater than a predetermined constant L.

Regress now and Regress later

- Regress now

$$\begin{pmatrix} \eta_1(X_{t_k}^{\pi,1}) & & \eta_Q(X_{t_k}^{\pi,1}) \\ & \ddots & \\ \eta_1(X_{t_k}^{\pi,\#B}) & & \eta_Q(X_{t_k}^{\pi,\#B}) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_Q \end{pmatrix} = \begin{pmatrix} g(X_{t_{k+1}}^{\pi,1}) \\ \vdots \\ g(X_{t_{k+1}}^{\pi,\#B}) \end{pmatrix}$$

$$\mathbb{E}[g(X_{t_{k+1}}^{\pi})|X_{t_k}^{\pi} = x] \approx \sum_{l=1}^{Q} \alpha_l \eta_l(x)$$

- Regress later

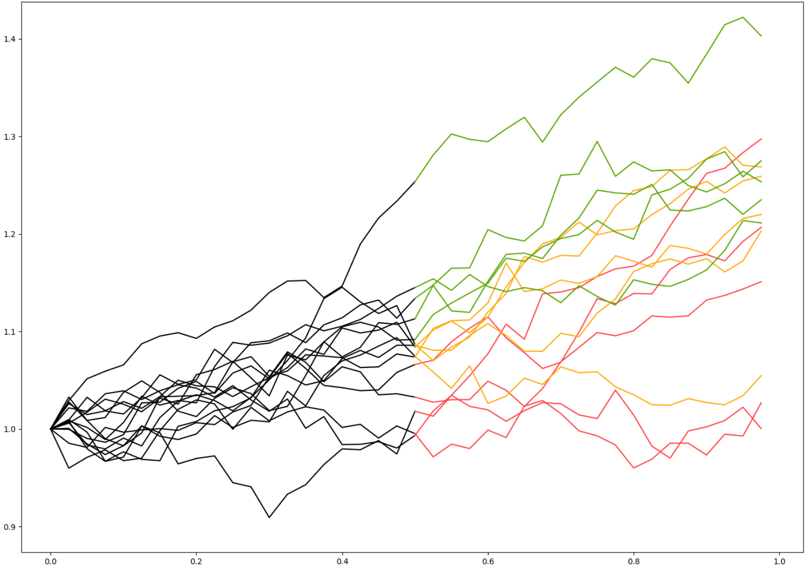
$$\begin{pmatrix} \eta_1(X_{t_{k+1}}^{\pi,1}) & & \eta_Q(X_{t_{k+1}}^{\pi,1}) \\ & \ddots & \\ \eta_1(X_{t_{k+1}}^{\pi,\#B}) & & \eta_Q(X_{t_{k+1}}^{\pi,\#B}) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_Q \end{pmatrix} = \begin{pmatrix} g(X_{t_{k+1}}^{\pi,1}) \\ \vdots \\ g(X_{t_{k+1}}^{\pi,\#B}) \end{pmatrix}$$

$$\mathbb{E}[g(X_{t_{k+1}}^{\pi})|X_{t_k}^{\pi} = x] \approx \sum_{l=1}^{Q} \alpha_l \mathbb{E}[\eta_l(X_{t_{k+1}}^{\pi})|X_{t_k}^{\pi} = x]$$

Recurring steps in time (II)

- Non-nested Monte Carlo scheme
- Regress-later
- Localization (Bundling)
  - At each time period, all paths are bundled into $\mathcal{B}_{t_k}(1), \dots, \mathcal{B}_{t_k}(B)$ (almost) equal-size, non-overlapping partitions based on the result of $(X_{t_k}^{\pi,m})$.
  - We perform the approximation separately at each bundle.

Bundling



Formulation

Specifically, the bundle regression parameters $\alpha_{k+1}(b)$, $\beta_{k+1}(b)$, $\gamma_{k+1}(b)$ are defined as

$$\alpha_{k+1}(b) = \arg\min_{\alpha \in \mathbb{R}^Q} \frac{\sum_{m=1}^{M} (p(X_{t_{k+1}}^{\pi,m})\alpha - y_{k+1}^{(\theta_1,\theta_2),R,I}(X_{t_{k+1}}^{\pi,m}))^2 \mathbf{1}_{\mathcal{B}_{t_k}(b)}(X_{t_k}^{\pi,m})}{\sum_{m=1}^{M} \mathbf{1}_{\mathcal{B}_{t_k(b)}}(X_{t_k}^{\pi.m})}$$

$$\beta_{i,k+1}(b) = \arg\min_{\beta \in \mathbb{R}^Q} \frac{\sum_{m=1}^{M} (p(X_{t_{k+1}}^{\pi,m})\beta - z_{i,k+1}^{(\theta_1,\theta_2),R}(X_{t_{k+1}}^{\pi,m}))^2 \mathbf{1}_{\mathcal{B}_{t_k}(b)}(X_{t_k}^{\pi,m})}{\sum_{m=1}^{M} \mathbf{1}_{\mathcal{B}_{t_k(b)}}(X_{t_k}^{\pi.m})}$$

$$\gamma_{k+1}(b) =$$

$$\arg\min_{\gamma \in \mathbb{R}^Q} \frac{\sum_{m=1}^{M} (p(X_{t_{k+1}}^{\pi,m})\gamma - f_{k+1}(y_{k+1}^{(\theta_1,\theta_2),R,I}, z_{k+1}^{(\theta_1,\theta_2),R}))^2 \mathbf{1}_{\mathcal{B}_{t_k}(b)}(X_{t_k}^{\pi,m})}{\sum_{m=1}^{M} \mathbf{1}_{\mathcal{B}_{t_k(b)}}(X_{t_k}^{\pi.m})}$$

Formulation (cont.)

The approximate functions within the bundle at time k are defined by :

$$
\begin{aligned}
z_{r,k}^{(\theta_1,\theta_2),R}(b,x) &= -\theta_2^{-1}(1-\theta_2)\mathbb{E}_{t_k}^{x}\left[p(X_{t_{k+1}}^{\pi})\right]\beta_{k+1}(b) \\
&\quad + \theta_2^{-1}\mathbb{E}_{t_k}^{x}\left[\frac{\Delta W_{r,k}}{\Delta_k}p(X_{t_{k+1}}^{\pi})\right](\alpha_{k+1}(b)+(1-\theta_2)\Delta_k\gamma_{k+1}(b)), \\
y_k^{(\theta_1,\theta_2),R,0}(b,x) &= \mathbb{E}_{t_k}^{x}\left[p(X_{t_{k+1}}^{\pi})\right]\alpha_{k+1}(b), \\
y_k^{(\theta_1,\theta_2),R,i}(b,x) &= \Delta_k\theta_1 f_k(y_k^{\pi,R,i-1}(x), z_k^{\pi,R}(x)) + h_k(x), \\
h_k(b,x) &= \mathbb{E}_{t_k}^{x}\left[p(X_{t_{k+1}}^{\pi})\right](\alpha_{k+1}(b)+\Delta_k(1-\theta_1)\gamma_{k+1}(b)), \quad i=1,\ldots,I,
\end{aligned}
$$

with

$$
y_k^{(\theta_1,\theta_2),R,I}(x) = \sum_{b=1}^{B}\mathbf{1}_{x\in\mathcal{B}_{t_k}(b)}y_k^{(\theta_1,\theta_2),R,I}(b,x)
$$

and similarly for z.

Refined Regression

Theorem 1

Assume that for a real function v that is bounded in a compact set and $\int v^2(x)\nu(dx) \leq \infty$, then

$$
\begin{aligned}
&\hat{\mathbb{E}}_S\left[\iint |v(y) - \tilde{v}(x,y)|^2 \nu(dx,dy)\right] \\
\leq &\frac{\vartheta(L')}{\hat{\mathbb{E}}[\mathbf{1}_S]}\hat{\mathbb{E}}\left[\sum_{B\in\mathbb{B}}\int_B\int \nu(dx,dy)\frac{(\log(\sum_{m=1}^M \mathbf{1}_\mathcal{B}(X^m))+1)Q}{\sum_{m=1}^M \mathbf{1}_\mathcal{B}(X^m)}\right] \\
+&\frac{8}{\hat{\mathbb{E}}[\mathbf{1}_S]}\hat{\mathbb{E}}\left[\sum_{B\in\mathbb{B}}\int_B\int \nu(dx,dy)(\inf_{\phi\in H}\sup_{x\in B}\mathbb{E}\left[|v(Y)-\phi(Y)|^2|X=x\right]\wedge L')\right] \\
+&\hat{\mathbb{E}}_S\left[\iint |v(y) - \tilde{v}(x,y)|^2(1-\mathbf{1}_\mathcal{A}(y))\nu(dx,dy)\right]
\end{aligned}
$$

Example 1

We consider the BSDE:

$$\begin{cases} dX_t = & dW_t, \\ dY_t = & -(Y_t Z_t - Z_t + 2.5 Y_t - \sin(t + X_t)\cos(t + X_t) \\ & -2\sin(t + X_t))dt + Z_t dW_t, \end{cases}$$

with the initial and terminal conditions $x_0 = 0$ and $Y_T = \sin(X_T + T)$. The exact solution is given by

$$(Y_t, Z_t) = (\sin(X_t + t), \cos(X_t + t)).$$

The terminal time is set to be $T = 1$ and $(Y_0, Z_0) = (0, 1)$.
We run the examples with the basis functions $\eta(x) = (1, x, x^2)$ and bundle based on the value of x.

Test Case	Example	θ_1	θ_2	I	M	N	B	L
D	1	0.5	0.5	4	2^{2J}	2^J	2^J	100
E	1	0.5	0.5	4	2^{2J}	2^J	2^J	10000
F	1	0.5	0.5	4	2^{2J}	2^J	2^J	–

$$|Y_0 - y_0^{(\theta_1,\theta_2),R}(x_0)|$$

J	2	3	4	5
D	NA	9.2870×10^{-2}	1.0114×10^{-1}	8.1415×10^{-2}
E	29.2228	7.8601×10^{-1}	3.9639×10^{-1}	5.2388×10^{-2}
F	2.2154×10^{15}	1.9059×10^{56}	3.4731×10^{-1}	5.8511×10^{-2}
J	**6**	**7**	**8**	
D	3.9920×10^{-3}	1.5486×10^{-2}	NA	
E	1.1931×10^{-2}	1.2395×10^{-2}	1.4347×10^{-3}	
F	2.0485×10^{-3}	6.8277×10^{-3}	2.6705×10^{-3}	

Example 2: European option

We consider a market where the assets satisfy:

$$dS_{i,t} = \mu_i S_{i,t} dt + \sigma_i S_{i,t} dB_{i,t},\ 1 \leq i \leq q$$

with B_t being a correlated q-dimension Wiener process with

$$dB_{i,t} dB_{j,t} = \rho_i j dt.$$

The parameters ρ_{ij} form a symmetric matrix ρ,

$$\rho = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1q} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2q} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{q1} & \rho_{q2} & \rho_{q3} & \cdots & 1 \end{pmatrix},$$

and we assume it is invertible. By performing a Cholesky decomposition on ρ such that $LL^T = \rho$, we relate B_t to standard Brownian motion

$$B_t = LW_t.$$

Example 2: European option (cont.)

For a European option with terminal payoff $g(S_t)$, a replicating portfolio Y_t, containing $\omega_{i,t}$ of asset $S_{i,t}$ and $Z_t = (\omega_{1,t}\sigma_1 S_{1,t}, \ldots, \omega_{q,t}\sigma_q, S_{q,t})L$ solve the BSDE,

$$\begin{cases} dY_t = -\left(-rY_t - Z_t L^{-1}\left(\frac{\mu - r}{\sigma}\right)\right)dt + Z_t dW_t; \\ Y_T = g(S_T), \end{cases}$$

where $\left(\frac{\mu-r}{\sigma}\right) = \left(\frac{\mu_1 - r}{\sigma_1}, \cdots, \frac{\mu_q - r}{\sigma_q}\right)^T$.
In this numerical test, we use the 5-dimensional example from [Reisinger and Wittum, 2007].

Example 2: European option (cont.)

We would consider a European weighted basket put option for 1 year in our test, therefore, the payoff function g is given by

$$g(s) = \left(1 - \sum_{i=1}^{5} w_i s_i\right)^+,$$

where $(w_1, w_2, w_3, w_4, w_5) = (38.1, 6.5, 5.7, 27.0, 22.7)$.
The reference price is given as 0.175866.
We use equal-partitioning and sorting the paths according to $\sum_{i=1}^{5} w_i X^m_{t_p,i}$.
The regression basis is $p_k(x) = \left(\sum_{i=1}^{5} w_i x_i\right)^{k-1}$ for $k = 1, \ldots, K$.

Test Case	Example	θ_1	θ_2	I	M	N	B	L	K
AA	2	0.5	0.5	4	2^{12}	10	2^{2J}	-	3
AB	2	0	1	-	2^{11}	10	2^{2J}	-	2

$$|Y_0 - y_0^{(\theta_1,\theta_2),R}(x_0)|$$

J	0	1	2
AA	2.0321×10^{-3}	2.2567×10^{-3}	1.9883×10^{-3}
AB	2.9314×10^{-3}	1.8934×10^{-3}	2.2151×10^{-4}

References

Reisinger, C. and Wittum, G. (2007).
Efficient hierarchical approximation of highdimensional option pricing problems.
SIAM Journal on Scientific Computing, 29(1):440–458.

Zhao, W., Li, Y., and Zhang, G. (2012).
A generalized θ-scheme for solving backward stochastic differential equations.
Discrete and Continuous Dynamical Systems - Series B, 17(5):1585–1603.

Thank You