Convergence of Bender/Denk algorithm

L^2-error estimates and rate of convergence for Monte Carlo projection

Plamen Turkedjiev

Tuesday 26th October, 2010

Approximate the following BSDE using a discretized scheme:

$$-\mathrm{d}Y_s = f(s, X_s, Y_s, Z_s)\mathrm{d}s - Z_s\mathrm{d}W_s,\ 0 \le t < T$$
$$Y_T(= \xi) = \Phi(X_T)$$

- $(X_t)_{0\le t\le T}$ diffusion with Lipschitz data (b, σ), $t \mapsto b(t,0), \sigma(t,0)$ bounded
- $(x, y, z) \mapsto f(t, x, y, z)$ and $x \mapsto \Phi(x)$ are Lipschitz continuous, $t \mapsto f(t, 0, 0, 0)$ bounded

Discretization: Time-grid with N points, $\Delta_k := t_{k+1} - t_k$, $|\pi| := \max_{0\le k\le N-1} \Delta_k$,$\Delta W_{t_k} := W_{k+1} - W_k$

1. Select regression basis $p(x) = (\phi^1(x), \dots, \phi^M(x))$ such that for Euler simulation $(X^{(\pi)}_{t_k})_{0\le k\le N}$, $\phi^i(X^{(\pi)}_{t_k}) \in L^2$ and $\mathbb{E}\Big[p(X^{(\pi)}_{t_k})[p(X^{(\pi)}_{t_k})]^{tr}\Big]$ invertible
   Transform $v_{t_k} := \Big(p(X^{(\pi)}_{t_k}), \frac{\Delta W_{t_k}}{\sqrt{\Delta_k}} p(X^{(\pi)}_{t_k})\Big) \in \mathbb{R}^M \times \mathbb{R}^M$
2. Perform n Picard iterations for regression coefficients $\hat\theta^{(n)}_{t_k} = (\hat\theta^{(n)}_{Y,t_k}, \hat\theta^{(n)}_{Z,t_k}) \in \mathbb{R}^M \times \mathbb{R}^M$:

$$\hat Y^{(m)}_{t_j} := \hat\theta^{(m)}_{Y,t_j} \cdot p(X^{(\pi)}_{t_j}),\ \hat Z^{(m)}_{t_j} := \frac{1}{\sqrt{\Delta_k}} \hat\theta^{(m)}_{Z,t_j} \cdot p(X^{(\pi)}_{t_j})$$

$$\hat\theta^{(m+1)}_{t_k} := \arg\min_\theta \mathbb{E}\left[\left|\xi^{(\pi)} - \sum_{j=k}^{N-1} f(t_j, X^{(\pi)}_{t_j}, \hat Y^{(m)}_{t_j}, \hat Z^{(m)}_{t_j})\Delta_j - \theta \cdot v_{t_k}\right|^2\right]$$

4. Projection estimators:

$$\hat Y^{(n)}_{t_j} := \hat\theta^{(n)}_{Y,t_j} \cdot p(X^{(\pi)}_{t_j}),\ \hat Z^{(n)}_{t_j} := \frac{1}{\sqrt{\Delta_k}} \hat\theta^{(n)}_{Z,t_j} \cdot p(X^{(\pi)}_{t_j})$$

1. Select regression basis $p(x) = (\phi^1(x), \dots, \phi^M(x))$ such that for Euler simulation $(X_{t_k}^{(\pi)})_{0\leq k\leq N}$, $\phi^i(X_{t_k}^{(\pi)}) \in L^2$ and $\mathbb{E}\big[p(X_{t_k}^{(\pi)})[p(X_{t_k}^{(\pi)})]^{tr}\big]$ invertible
2. Generate L independent simulations $\{(W_{t_k}^{(\pi,l)})_{0\leq k\leq N}\}_{1\leq l\leq L}$ of $(W_{t_k})_{0\leq k\leq N}$ and Euler simulations $\{(X_{t_k}^{(\pi,l)})_{0\leq k\leq N}\}_{1\leq l\leq L}$
   Transform $v_{t_k}^l := \Big(p(X_{t_k}^{(\pi,l)}), \frac{\Delta W_{t_k}^l}{\sqrt{\Delta_k}} p(X_{t_k}^{(\pi,l)})\Big) \in \mathbb{R}^M \times \mathbb{R}^M$
3. Perform n Picard iterations for regression coefficients $\hat{\theta}_{t_k}^{(n,L)} = (\hat{\theta}_{Y,t_k}^{(n,L)}, \hat{\theta}_{Z,t_k}^{(n,L)}) \in \mathbb{R}^M \times \mathbb{R}^M$:

$$\hat{Y}_{t_j}^{(m,L,l)} := \hat{\rho}_j\left(\hat{\theta}_{Y,t_j}^{(m,L)} \cdot p(X_{t_j}^{(\pi,l)})\right), \; \hat{Z}_{t_j}^{(m,L,l)} := \hat{\rho}_j\left(\frac{\hat{\theta}_{Z,t_k}^{(m,L)}}{\sqrt{\Delta_k}} \cdot p(X_{t_j}^{(\pi,l)})\right)$$

$$\hat{\theta}_{t_k}^{(m+1,L)} := \arg\min_{\theta} \frac{1}{L}\sum_{l=1}^{L}\Big[\Big|\xi^{(\pi,l)} - \sum_{j=k}^{N-1} f(t_j, X_{t_j}^{(\pi,l)}, \hat{Y}_{t_j}^{(m,L,l)}, \hat{Z}_{t_j}^{(m,L,l)})\Delta_j - \theta \cdot v_{t_k}^l\Big|^2\Big]$$

4. Monte Carlo estimators:

$$\hat{Y}_{t_k}^{(n,L)} := \hat{\rho}_k\left(\hat{\theta}_{Y,t_k}^{(n,L)} \cdot p(X_{t_k}^{(\pi)})\right), \quad \hat{Z}_{t_k}^{(n,L)} := \hat{\rho}_k\left(\frac{\hat{\theta}_{Z,t_k}^{(n,L)}}{\sqrt{\Delta_k}} \cdot p(X_{t_k}^{(\pi)})\right)$$

Construction of truncation functions:

1. Take $\mathbb{R}$-valued functions $\rho(\cdot) := \max(1, C_0|p(\cdot)|)$, with C_0 appropriately chosen
2. Choose $\phi \in C_b^2(\mathbb{R})$ such that $\phi(x)\mathbf{1}_{[-3/2,3/2]}(x) = x$, $||\phi||_\infty \leq 2$, and $||\phi'||_\infty \leq 1$
3. $\hat{\rho}_k(x) := \rho\left(X_{t_k}^{(\pi)}\right) \cdot \phi\left(\frac{x}{\rho(X_{t_k}^{(\pi)})}\right)$

Properties:

- $\hat{\rho}_k(\hat{Y}_{t_k}^{(n)}) = \hat{Y}_{t_k}^{(n)}$ and $\hat{\rho}_k(\sqrt{\Delta_k}\hat{Z}_{t_k}^{(n)}) = \sqrt{\Delta_k}\hat{Z}_{t_k}^{(n)}$
- $|\hat{\rho}_k(x)| \leq 2\rho(X_{t_k}^{(\pi)}) \in L^2(X_{t_k}^{(\pi)}) \quad \forall x \in \mathbb{R},\ 0 \leq k \leq N-1$
- Lipschitz continuous with Lipschitz constant 1

Normal equations: Set $V_{t_k}^L := \frac{1}{L}\sum_{l=1}^L v_{t_k}^l [v_{t_k}^l]^{tr}$

$$\mathbb{E}\Big[v_{t_k}[v_{t_k}]^{tr}\Big]\hat{\theta}_{t_k}^{(n+1)} = \mathbb{E}\Bigg[v_{t_k}\Bigg(\xi^{(\pi)} - \sum_{j=k}^{N-1} f(t_j, X_{t_j}^{(\pi)}, \hat{Y}_{t_j}^{(n)}, \hat{Z}_{t_j}^{(n)})\Delta_j\Bigg)\Bigg]$$

$$V_{t_k}^L\ \hat{\theta}_{t_k}^{(n+1,L)} = \frac{1}{L}\sum_{l=1}^L v_{t_k}^l\Bigg(\xi^{(\pi,l)} - \sum_{j=k}^{N-1} f(t_j, X_{t_j}^{(\pi,l)}, \hat{Y}_{t_j}^{(n,L,l)}, \hat{Z}_{t_j}^{(n,L,l)})\Delta_j\Bigg)$$

Want invertible $V_{t_k}^L$:

$$\mathscr{A}_{t_0}^L := \left\{ \begin{array}{l} |V_{t_k}^L - \mathbb{E}\big[v_{t_k}[v_{t_k}]^{tr}\big]| \le |\pi|, \\ \forall\ q = 0, \ldots, D,\ \text{and}\ \ 0 \le k \le N-1 \end{array} \right\}$$

Task:

Find estimate for

$$\textbf{Err}(n,\pi,L) = \max_{0\le k\le N-1} \mathbb{E}|\hat{Y}^{(n)}_{t_k} - \hat{Y}^{(n,L)}_{t_k}|^2 + \sum_{k=0}^{N-1} \mathbb{E}|\hat{Z}^{(n)}_{t_k} - \hat{Z}^{(n,L)}_{t_k}|^2 \Delta_k$$

Use estimate to prove convergence and find rate of convergence w.r.t. L

Theorem

Assume:

- $(X^{(\pi)}_{t_k})_{0\le k\le N-1}$, $\left(p(X^{(\pi)}_{t_k})\right)_{0\le k\le N-1}$ have fourth moments
- $|\pi| < \min(1,\bar{V})$ for $\bar{V} := \min_{0\le k\le N-1} \lambda_{MIN}(\mathbb{E}\left[v_{t_k}[v_{t_k}]^{tr}\right])$

$$\textbf{Err}(n,\pi,L) \le C \sum_{k=0}^{N-1} \mathbb{E}\left[|\rho(P^{(\pi)}_{t_k})|^2 \mathbf{1}_{[\mathscr{A}^L_{t_0}]^C}\right]$$

$$+ \quad \left(\frac{CN}{(1-|\pi|)^2 L} + \frac{CN}{(\bar{V}-|\pi|)L}\right)\left(1 + e^{2KN}\sum_{m=1}^{n-1}(C|\pi|)^m\right)$$

First estimate of the error:

$$\textbf{Err}(n,\pi,L) \leq C \sum_{k=0}^{N-1} \mathbb{E}\left[|\rho(X_{t_k}^{(\pi)})|^2 \mathbf{1}_{[\mathscr{A}_{t_0}^L]^C}\right] + CN \max_{0\leq k\leq N-1} \mathbb{E}\left[|\mathbf{1}_{\mathscr{A}_{t_0}^L}(\hat{\theta}_{t_k}^{(n,L)} - \hat{\theta}_{t_k}^{(n)})|^2\right]$$

Decomposition:

$$\mathbf{1}_{\mathscr{A}_{t_0}^L}(\hat{\theta}_{t_k}^{(n,L)} - \hat{\theta}_{t_k}^{(n)}) = B_k^{(1,n)} + B_k^{(2,n)} + B_k^{(3,n)}$$

Set $\hat{Y}_{t_j}^{(n-1,l)} := \hat{\theta}_{Y,t_j}^{(n-1)} \cdot p(X_{t_j}^{(\pi,l)}),\ \hat{Z}_{t_j}^{(n-1,l)} := \frac{\hat{\theta}_{Z,t_j}^{(n-1)}}{\sqrt{\Delta_j}} \cdot p(X_{t_j}^{(\pi,l)})$

$$B_k^{(1,n)} := \Big([\mathbf{1}_{\mathscr{A}_{t_0}^L} V_{t_k}^L]^{-1} - \mathbb{E}\big[v_{t_k}[v_{t_k}]^{tr}\big]^{-1} \Big) \times \mathbb{E}\Big[v_{t_k} (\xi^{(\pi)} - \sum_{j=k}^{N-1} f\big(t_j, X_{t_j}^{(\pi)}, \hat{Y}_{t_j}^{(n-1)}, \hat{Z}_{t_j}^{(n-1)}\big) \Delta_j) \Big]$$

$$B_k^{(2,n)} := [\mathbf{1}_{\mathscr{A}_{t_0}^L} V_{t_k}^L]^{-1} \times \Bigg(\frac{1}{L} \sum_{l=1}^{L} v_{t_k}^l \Big(\xi^{(\pi,l)} - \sum_{j=k}^{N-1} f\big(t_j, X_{t_j}^{(\pi,l)}, \hat{Y}_{t_j}^{(n-1,l)}, \hat{Z}_{t_j}^{(n-1,l)}\big) \Delta_j \Big) - \mathbb{E}\Big[v_{t_k} \Big(\xi^{(\pi)} - \sum_{j=k}^{N-1} f\big(t_j, X_{t_j}^{(\pi)}, \hat{Y}_{t_j}^{(n-1)}, \hat{Z}_{t_j}^{(n-1)}\big) \Delta_j \Big) \Big] \Bigg)$$

$$B_k^{(3,n)} := \frac{[\mathbf{1}_{\mathscr{A}_{t_0}^L} V_{t_k}^L]^{-1}}{L} \sum_{l=1}^{L} v_{t_k}^l \Big(\sum_{j=k}^{N-1} \Big(f\big(t_j, X_{t_j}^{(\pi,l)}, \hat{Y}_{t_j}^{(n-1,l)}, \hat{Z}_{t_j}^{(n-1,l)}\big) - f\big(t_j, X_{t_j}^{(\pi,l)}, \hat{Y}_{t_j}^{(n-1,L,l)}, \hat{Z}_{t_j}^{(n-1,L,l)}\big) \Big) \Delta_j \Big)$$

k-uniform L^2-bounds for $B_k^{(q,n)}$

Define $\|\theta\|^2 = \max_{0\le k\le N-1} \mathbb{E}\left[|\theta_k|^2 \mathbf{1}_{\mathscr{A}_{t_0}^L}\right]$

Assume:

- $(X_{t_k}^{(\pi)})_{0\le k\le N-1}$, $\left(p(X_{t_k}^{(\pi)})\right)_{0\le k\le N-1}$ have fourth moments
- $|\pi| < \min(1,\bar{V})$ for $\bar{V} := \min_{0\le k\le N-1} \lambda_{MIN}(\mathbb{E}\left[v_{t_k}[v_{t_k}]^{tr}\right])$

(a) $\|(B_k^{(1,n)})\|^2 \le C(1-|\pi|)^{-2}L^{-1}$

(b) $\|(B_k^{(2,n)})\|^2 \le C(\lambda_{MIN}(\mathbb{E}\left[v_{t_k}[v_{t_k}]^{tr}\right]) - |\pi|)^{-1}L^{-1}$

(c) $\mathbb{E}\|(B_k^{(3,n)})\|_0^2 \le C|\pi|(\bar{V}-|\pi|)^{-1}\|(\hat{\theta}^{(n-1)} - \hat{\theta}^{(n-1,L)})\mathbf{1}_{\mathscr{A}_{t_0}^L}\|_0^2$,
where $\|\theta\|_0 := \max_{0\le k\le N-1}\left(e^{-2K(N-k)}|\theta_k|\right)$

Proof (c).

$$(\bar{V} - |\pi|)\|(B_k^{(3,n)})\|_0^2 \leq \max_{0\leq k\leq N-1} e^{-2K(N-k)} \times$$
$$\mathbb{E}\left[\mathbf{1}_{\mathscr{A}_{t_0}^L} \frac{1}{L}\sum_{l=1}^{L}\left|\sum_{j=k}^{N-1}(f_j^{(1,n-1,l)} - f_j^{(2,n-1,l)})\Delta_j\right|^2\right]$$
$$\leq C|\pi| \max_{0\leq k\leq N-1} e^{-2K(N-k)} \times$$
$$\mathbb{E}\Big[\Big|\sum_{j=k}^{N-1} e^{-2K(N-j)}\|(\hat{\theta}^{(n-1)} - \hat{\theta}^{(n-1,L)})\mathbf{1}_{\mathscr{A}_{t_0}^L}\|_0\Big|^2\Big]$$
$$\leq C|\pi|\|(\hat{\theta}^{(n-1)} - \hat{\theta}^{(n-1,L)})\mathbf{1}_{\mathscr{A}_{t_0}^L}\|_0^2 \qquad \square$$

2-dimensional Brownian motion $(W_t)_{0 \le t \le T}$

- $\mathrm{d}X_t = X_t \sigma \mathrm{d}W_t$, $x = 1$, $\sigma = (1, 0)$
- $-\mathrm{d}Y_t = P(Z_t)\mathrm{d}t - Z_t \mathrm{d}W_t$, $P(z) = |z_2|$, $\Phi(x) = (x_1 - K)_+$ with $K = 0.7$, $T = 1$

Statistical tool: average residual sum of squares (RSS) between the true and simulated Y-solution at each time point

$$RSS = \frac{1}{1000} \sum_{l=0}^{1000} \sum_{k=0}^{N-1} |\hat{Y}_{t_k}^{(n,L)}(\omega_l) - Y_{t_k}(\omega_l)|^2$$

Table containing mean and standard deviation of RSS over 50 runs of the algorithm for different sample size L and time-grid sizes N

Basis: "Ideal basis"

$$\phi^1(x) = x\Phi(d_+(x, T - t_k)) - K\Phi(d_-(x, T)), \quad \phi^2(x) = \Phi(d_-(x, t_k))$$
$$d_+(x,t) = \frac{\log(\frac{x}{K}) + t/2}{\sqrt{t}}, \quad d_-(x,t) = \frac{\log(\frac{x}{K}) - t/2}{\sqrt{t}}$$

Number of Picard iterations: $n = 6$

Grid size N	5	10	15
Mean	0.240	0.568	2.13
Std dev.	0.225	0.571	3.48

Table containing mean and standard deviation of RSS over 50 runs of the algorithm for different sample size L and time-grid sizes N

Basis: log-Hermite polynomials up to degree 8

Number of Picard iterations: $n = 7$

	Sample size L	10^3	10^4	10^5
$N = 5$	Mean	20.1	0.325	2.02×10^{-2}
	Std Dev.	57.5	1.29	1.05×10^{-2}
$N = 10$	Mean	38.8	2.58	3.01×10^{-2}
	Std Dev.	69.1	10.1	2.59×10^{-2}
$N = 15$	Mean	58.2	3.89	0.114
	Std Dev.	116.9	8.41	4.89×10^{-2}

Table containing mean and standard deviation of RSS over 50 runs of the algorithm for different sample size L

Basis: log-Hermite polynomials up to degree 20

Number of Picard iterations: $n = 7$

	Sample size L	10^3	10^4	10^5
$N = 10$	Mean	2.8×10^{12}	2.97×10^5	4.03×10^{-2}
	Std Dev.	1.38×10^{13}	1.35×10^6	2.42×10^{-2}

Thank you for your attention!

References

1. Becherer, D. From bounds on optimal growth towards a theory of good-deal hedging, Radon Series Comp. Appl. Math 8 (2009) 1-25
2. Bender, C, Denk, R, A forward scheme for backward SDEs (2007), Stoch. Proc. App. 117, 1793-1812
3. Gobet, E., Lemor, J., Warin, X., A Regression-Based Monte Carlo Method to Solve Backward Stochastic Differential Equations (2005), Ann. Appl. Probab., 2172-2202