Numerical Analysis / Probability Theory

Reducing variance in the numerical solution of BSDEs

Samu Alanko [a], Marco Avellaneda [b]

[a] *Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012, United States*
[b] *Finance Concepts SARL, 49-51 Avenue Victor-Hugo, 75116 Paris, France*

Received *****; accepted after revision *****

Presented by A. Bensoussan

Abstract

Numerical methods based on time discretization and estimation of conditional expectations for solving backward stochastic differential equations (BSDEs) have been the object of considerable research, particularly in view of the applications to finance. We introduce and implement a simple control variate technique to reduce the simulation error of the conditional expectation estimates in BSDE methods. These modifications increase the accuracy of the existing algorithms without additional computational cost. *To cite this article: S. Alanko, M. Avellaneda, C. R. Acad. Sci. Paris, Ser. I 340 (2005).*

Résumé

Réduction de variance pour la solution numérique des BSDEs. Les méthodes numériques basées sur discrétisation de pas de temps et estimation d'espérances conditionnelles pour la résolution d'équations différentielles stochastiques rétrogrades (BSDEs) ont fait l'objet d'études récentes, en particulier pour leurs applications en Finance. Nous proposons ici une technique basée sur variables de contrôle afin de réduire l'erreur dans la simulation des estimateurs d'espérance conditionnelles. Ces modifications peuvent être adaptées facilement aux algorithmes connus pour augmenter leur efficacité avec essentiellement le même temps de calcul. *Pour citer cet article : S. Alanko, M. Avellaneda, C. R. Acad. Sci. Paris, Ser. I 340 (2005).*

1. Introduction

Several approaches for solving backward stochastic differential equations (BSDEs) have been considered in the literature (see for instance [5,7] and references therein). One type of numerical scheme is based on

Email addresses: alanko@courant.nyu.edu (Samu Alanko), avellaneda@courant.nyu.edu (Marco Avellaneda).

time discretization and estimating conditional expectations (see e.g. [8,6,2,3,4]). Since these conditional expectations cannot be evaluated explicitly, methods such as Least Squares Monte Carlo or kernel regression are used. We suggest here a simple modification to these methods in order to reduce the simulation error of the conditional expectation estimates.

The main idea is captured in the following elementary observation. Let W_t be a standard Brownian motion and f a sufficiently smooth function satisfying e.g. a polynomial growth condition. Integration by parts with dominated convergence theorem then shows that

$$f'(x) = \lim_{\Delta t \to 0} \mathbb{E}\left[f(x + W_{\Delta t})\tfrac{W_{\Delta t}}{\Delta t}\right], \tag{1}$$

$$f''(x) = \lim_{\Delta t \to 0} \mathbb{E}\left[f(x + W_{\Delta t})\tfrac{W_{\Delta t}^2 - \Delta t}{\Delta t^2}\right]. \tag{2}$$

These formulas suggest an approach for estimating derivatives by Monte-Carlo. Unfortunately, replacing expectations by empirical averages leads to poor convergence when Δt is small. A simple Taylor expansion argument shows that the sample variances are $f(x)^2 \Delta t^{-1} + \mathcal{O}(1)$ in (1) and $2f(x)^2 \Delta t^{-2} + \mathcal{O}(\Delta t^{-1})$ in (2) which blow up as $\Delta t \to 0$, thus leading to a large standard error in the estimates. This problem can be avoided by using estimators based on the equivalent formulas

$$f'(x) = \lim_{\Delta t \to 0} \mathbb{E}\left[\Big(f(x + W_{\Delta t}) - f(x)\Big)\tfrac{W_{\Delta t}}{\Delta t}\right], \tag{1*}$$

$$f''(x) = \lim_{\Delta t \to 0} \mathbb{E}\left[\Big(f(x + W_{\Delta t}) - f(x) - f'(x)W_{\Delta t}\Big)\tfrac{W_{\Delta t}^2 - \Delta t}{\Delta t^2}\right] \tag{2*}$$

obtained by subtracting the first order Taylor terms from $f(x + W_{\Delta t})$ in order to make the numerator and the denominator of the same order in Δt while keeping the expectation unchanged. This leads to sample variances $2f'(x)^2 + \mathcal{O}(\Delta t)$ and $\frac{37}{2} f''(x)^2 + \mathcal{O}(\Delta t)$ which, to the leading order, do not depend on Δt and thus allow much smaller values of Δt to be used.

2. Application to the numerical solution of BSDEs

To keep the notation simple, we consider the one-dimensional case; the ideas extend readily to multiple dimensions. Suppose that we are given a fully non-linear parabolic PDE in $[0, T] \times \mathbb{R}$

$$u_t + \mathcal{L}u = f(t, x, u, u_x, u_{xx}), \quad u(x, T) = g(x) \tag{3}$$

where $\mathcal{L}$ is the generator of a diffusion $\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t$ with a fixed initial value $X_0 = x$. If the PDE has a sufficiently smooth solution, then it follows by Ito's Lemma that the four processes $Y_t = u(X_t, t), Z_t = u_x(X_t, t), \Gamma_t = u_{xx}(X_t, t), A_t = (u_{xt} + \mathcal{L}u_x)(X_t, t)$ satisfy the second order BSDE

$$\mathrm{d}Y_t = f(t, X_t, Y_t, Z_t, \Gamma_t)\,\mathrm{d}t + \sigma(t, X_t)Z_t\,\mathrm{d}W_t, \quad \mathrm{d}Z_t = A_t\,\mathrm{d}t + \sigma(t, X_t)\Gamma_t\,\mathrm{d}W_t \tag{4}$$

with the terminal condition $Y_T = g(X_T)$. Conversely, if we can solve for Y_0 via the BSDE approach, we can find the numerical solution of the PDE (3) at a given point $(x, 0)$.

We consider two well-known numerical schemes for solving second-order BSDEs:

– Cheridito et al [3]:

$$\begin{cases} Y_N &= g(X_N), \quad Z_N = g'(X_N) \\ Y_{n-1} &= \mathbb{E}_{n-1}[Y_n] - f(t_{n-1}, X_{n-1}, Y_{n-1}, Z_{n-1}, \Gamma_{n-1})\Delta t \\ Z_{n-1} &= \frac{1}{\sigma_{n-1}}\mathbb{E}_{n-1}\left[Y_n \frac{\Delta W_{n-1}}{\Delta t}\right] \\ \Gamma_{n-1} &= \frac{1}{\sigma_{n-1}}\mathbb{E}_{n-1}\left[Z_n \frac{\Delta W_{n-1}}{\Delta t}\right] \end{cases}, \quad 1 \le n \le N \tag{5}$$

– Fahim et al [4]:

$$\begin{cases} Y_N &= g(X_N), \\ Y_{n-1} &= \mathbb{E}_{n-1}\left[Y_n\right] - f(t_{n-1}, X_{n-1}, Y_{n-1}, Z_{n-1}, \Gamma_{n-1})\Delta t \\ Z_{n-1} &= \frac{1}{\sigma_{n-1}}\mathbb{E}_{n-1}\left[Y_n \frac{\Delta W_{n-1}}{\Delta t}\right] \\ \Gamma_{n-1} &= \frac{1}{\sigma_{n-1}^2}\mathbb{E}_{n-1}\left[Y_n \frac{(\Delta W_{n-1})^2 - \Delta t}{\Delta t^2}\right] \end{cases}, \quad 1 \le n \le N. \tag{6}$$

Subscripts indicate evaluations at time $t_n = n\Delta t$, $\mathbb{E}_n\left[\cdot\right] = \mathbb{E}\left[\,\cdot \mid X_n\right]$, $\Delta W_{n-1} = W_n - W_{n-1}$ and $\sigma_n = \sigma(t_n, X_n)$. Since Z_n and Γ_n are approximations to $u_x(X_n, t_n)$ and $u_{xx}(X_n, t_n)$, we see that the expressions of Z_{n-1} and Γ_{n-1} are essentially the formulas (1) and (2). These formulas produce poor results for small time steps (see [7]). We suggest the following modifications based on the formulas (1*) and (2*):

– Modified scheme for Cheridito et al:

$$\begin{cases} Z_{n-1} &= \frac{1}{\sigma_{n-1}}\mathbb{E}_{n-1}\left[(Y_n - \mathbb{E}_{n-1}\left[Y_n\right])\frac{\Delta W_{n-1}}{\Delta t}\right] \\ \Gamma_{n-1} &= \frac{1}{\sigma_{n-1}}\mathbb{E}_{n-1}\left[(Z_n - Z_{n-1})\frac{\Delta W_{n-1}}{\Delta t}\right] \end{cases}, \quad 1 \le n \le N \tag{7}$$

– Modified scheme for Fahim et al:

$$\begin{cases} Z_{n-1} &= \frac{1}{\sigma_{n-1}}\mathbb{E}_{n-1}\left[(Y_n - \mathbb{E}_{n-1}\left[Y_n\right])\frac{\Delta W_{n-1}}{\Delta t}\right] \\ \Gamma_{n-1} &= \frac{1}{\sigma_{n-1}^2}\mathbb{E}_{n-1}\left[(Y_n - \mathbb{E}_{n-1}\left[Y_n\right] - \sigma_{n-1}Z_{n-1}\Delta W_{n-1})\frac{(\Delta W_{n-1})^2 - \Delta t}{\Delta t^2}\right] \end{cases}, \quad 1 \le n \le N. \tag{8}$$

These differ from the original ones only in the way that we have subtracted approximations of the first order Taylor expansion terms in the expressions of Z_{n-1} and Γ_{n-1}. Note that the correction terms are already computed at each step so the modifications do not add any significant computational cost.

3. Numerical examples

Example 1. (Simple estimator) We estimate the expectations described in the first section using the average of N samples. The naive estimators of (1) and (2) with their leading order standard errors are

$$\widehat{f'(x)} = \frac{1}{N\sqrt{\Delta t}}\sum_{n=1}^N Z_n f(x + \sqrt{\Delta t}Z_n), \quad \text{std} \approx \frac{|f(x)|}{\sqrt{\Delta t}\sqrt{N}} \tag{9}$$

$$\widehat{f''(x)} = \frac{1}{N\Delta t}\sum_{n=1}^N (Z_n^2 - 1) f(x + \sqrt{\Delta t}Z_n), \quad \text{std} \approx \frac{\sqrt{2}|f(x)|}{\Delta t\sqrt{N}} \tag{10}$$

where Z_n are independent standard normal random variables. Respectively, for (1*) and (2*) we have

$$\widehat{f'(x)}^* = \frac{1}{N\sqrt{\Delta t}}\sum_{n=1}^N Z_n\big(f(x + \sqrt{\Delta t}Z_n) - f(x)\big), \quad \text{std} \approx \frac{\sqrt{2}|f'(x)|}{\sqrt{N}} \tag{11}$$

$$\widehat{f''(x)}^* = \frac{1}{N\Delta t}\sum_{n=1}^N (Z_n^2 - 1)\big(f(x + \sqrt{\Delta t}Z_n) - f(x) - \sqrt{\Delta t}Z_n f'(x)\big), \quad \text{std} \approx \sqrt{\frac{37}{2}}\frac{|f''(x)|}{\sqrt{N}}. \tag{12}$$

In particular, if N is held constant and the time step Δt gets smaller, the naive estimates diverge whereas the standard errors of the new estimates do not to depend on Δt. For example, if we chose $\Delta t = 0.01$ we would need 10000 times more samples to get the same standard error for the naive estimate of f'' than what we would get by just using the new estimate. See Table 1.

Example 2. (Non-linear PDE) We used the four different BSDE schemes to solve the PDE

$$u_t + \tfrac{1}{2}\Sigma^2(u_{xx})x^2 u_{xx} = 0, \quad u(x,T) = (x-90)^+ - (x-110)^+ \tag{13}$$

where $\Sigma(x) = \sigma_{\min}\mathbf{1}_{x<0}(x) + \sigma_{\max}\mathbf{1}_{x\ge 0}(x)$. The value $u(100,0)$ corresponds to the maximum value of the call spread $C_{90} - C_{110}$ in the Uncertain Volatility Model (see [1,7]) with spot price 100, zero interest rate, 1 year time to maturity and volatility band $[\sigma_{\min}, \sigma_{\max}] = [0.1, 0.2]$.

estimator	mean empirical	mean exact	std empirical	std exact	range (max−min)
naive $\widehat{f'(x)}$	−0.1941	−0.1960	0.0991	0.0980	0.7236
new $\widehat{f'(x)}^*$	−0.1958	−0.1960	0.0009	0.0009	0.0070
naive $\widehat{f''(x)}$	−0.9631	−0.9410	4.3782	4.3836	32.576
new $\widehat{f''(x)}^*$	−0.9395	−0.9410	0.0129	0.0128	0.0948

Table 1
We estimated the derivatives of $f(x) = \exp(-x^2)$ at the point $x = 0.2$ using both the naive estimate and the new estimate. Each estimate was computed 10,000 times using time step $\Delta t = 0.001$ and $N = 100{,}000$. [Nous comparons les estimations "naïves" avec les nouvelles estimations pour le calcul numérique des deux premières dérivées de la fonction $f(x) = \exp(-x^2)$ au point $x = 0.2$. Chaque calcul a été effectue 10000 fois, avec un pas de temps $\Delta t = 0.001$ et $N = 100000$.]

We generated 100,000 paths of the forward diffusion (geometric Brownian motion with zero drift and volatility 0.15). Conditional expectations were estimated using basis projections on 20 exponentials $e^{-x^2/100}$ centered equidistantly between the minimum and maximum values of the sample paths X_{t_n} at each $t_n = n\Delta t$. The experiments were repeated with different time steps $\Delta t = (10\cdot 2^i)^{-1}, i = 0, 1, \ldots, 6$.

The results are shown in Table 2. The original schemes only give sensible answers for large values of Δt and diverge as Δt gets smaller. This is to be expected since the variances of the sample points used in the estimation of the conditional expectations blow up as $\Delta t \to 0$. The modified versions behave much better for smaller time steps although eventually show divergence for small enough Δt. A larger number of sample paths would probably allow even smaller time steps to be used.

Scheme \ time step Δt	1/10	1/20	1/40	1/80	1/160	1/320	1/640
Cheridito et al	10.98	11.13	11.26	32.53	*	*	*
Fahim et al	10.98	11.18	11.68	18.48	41.12	*	*
Modified Cheridito et al	10.97	11.07	11.18	11.22	11.21	11.21	11.27
Modified Fahim et al	10.99	11.14	11.18	11.21	11.26	11.29	11.65

Table 2
The results given by different methods with different time steps. The correct price is 11.20 and the asterisk means the computation diverged. [Résultats du calcul numérique pour le Call-Spread avec les différentes méthodes et avec des pas de temps différents pour $u(100, 0)$. La valeur correcte est 11.20. Les astérisques indiquent quand le schéma BSDE diverge.]

References

[1] M. Avellaneda, A. Levy, A. Parás, Pricing and hedging derivative securities in markets with uncertain volatilities, Applied Mathematical Finance 2 (1995) 73-88.

[2] B. Bouchard, N. Touzi, Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations, Stochastic Processes and their applications 111 (2004) 175-206.

[3] P. Cheridito, H.M. Soner, N. Touzi, N. Victoir, Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs, Communications on Pure and Applied Mathematics 60 (2006) 1081-1110.

[4] A. Fahim, N. Touzi., X. Warin, A probabilistic numerical method for fully nonlinear parabolic PDEs, The Annals of Applied Probability 21 (2011) 1322-1364.

[5] E. Gobet, C. Labart, Solving BSDE with adaptive control variate, SIAM Journ. on Numer. Anal. 48 (2010) 257-277.

[6] E. Gobet, J.-P. Lemor, X. Warin, A regression-based Monte Carlo method to solve backward stochastic differential equations, The Annals of Applied Probability 15 (2005) 2172-2202.

[7] J. Guyon, P. Henry-Labordère, Uncertain Volatility Model: A Monte-Carlo approach, Journ. of Com. Fin. 14 (2011) 37-71

[8] J. Zhang, A numerical scheme for BSDEs, The Annals of Applied Probability 14 (2004) 459-488.