FORC conference

Implementing Derivative Valuation Models

24 February 2006

FAST AMERICAN MONTE CARLO

Claudio Moni

OVERVIEW

- Products
- Dynamic programming principle
- Longstaff Schwartz approach
- Regression noise
- Variance reduction in pricing
- Variance reduction in regression
- Numerical results
- Extensions

PRODUCTS

- American/Bermudan options on a single stock.

- Bermudan swaptions: given a set of swaps $(Z_j)_{j=0,\ldots,L}, Z_L = 0$, the holder of the Bermudan B_i has the right to exchange his contract with Z_j at each time T_j, $j \geq i$.

- Cancellable swaps: let SW be a swap, which one party has the right to cancel at a pre-specified set of dates $\{T_0,\ldots,T_L\}$. Let $SW_{i,j}$ denote the sub-swap containing all cash-flows with anchor date $T_i \leq t < T_j$, and $C_{i,j}$ the associated cancellable swap. Then we can define a set of offsetting swaps $Z_j = -SW_{j,L}$. Let us denote by B_j the associated Bermudan swaption, we have
$$C_{j,L} = SW_{j,L} + B_j.$$

DYNAMIC PROGRAMMING PRINCIPLE

- We consider a probability space (Ω, A, Q), equipped with a discrete time filtration $(F_j)_{j=0,...,L}$. Given an adapted payoff process $(Z_j)_{j=0,...,L}$ with $Z_j \in L^2(\Omega, A, Q)$, $Z_L = 0$, we are interested in computing

$$\sup_{\tau \in \Upsilon_{0,L}} E(Z_\tau),$$

where $\Upsilon_{j,L}$ denotes the set of all stopping times with values in $\{j,...,L\}$.

- We have

$$\sup_{\tau \in \Upsilon_{0,L}} E(Z_\tau) = U_0,$$

where U_j is recursively defined by

$$U_L = Z_L,$$
$$U_j = \max(Z_j, E(U_{j+1} \mid F_j)), j = 0,...,L-1.$$

- This dynamic programming solution can be rewritten in terms of optimal stopping times τ_j as follows

$$\tau_L = L,$$
$$\tau_j = j1_{\{Z_j \geq E(Z_{\tau_{j+1}}|F_j)\}} + \tau_{j+1}1_{\{Z_j < E(Z_{\tau_{j+1}}|F_j)\}}, \; j = 0,\dots,L-1,$$
$$U_0 = E(Z_{\tau_0}).$$

- We assume that there is a (D-dimensional) (F_j)-Markov chain $(\mathbf{X}_j)_{j=0,\dots,L}$ such that, for $j=0,\dots,L,$

$$Z_j = f(j, \mathbf{X}_j)$$

for some Borel function $f(j,\cdot)$. We then have $U_j = V(j, \mathbf{X}_j)$ for some function $V(j,\cdot)$, and $E(Z_{\tau_{j+1}} \mid F_j) = E(Z_{\tau_{j+1}} \mid \mathbf{X}_j)$. We also assume that $\mathbf{X}_0$ is deterministic.

LONGSTAFF SCHWARTZ APPROACH

- Pricing by Monte Carlo simulation ($\leftarrow$multi-factor models)

- The LS approach to compute U_0 consists of two steps. The first step is to replace each conditional expectation $E(U_{j+1} \mid F_j)$ by its L^2 projection on a subspace generated by a finite set of F_j-measurable variables $\{e_k(\mathbf{X}_j)\}_{k=1,\ldots,M}$, $\| e_k(\mathbf{X}_j) \|_2 > 0$, plus a constant. We denote by P_j^M the projection operator. We recall that

$$P_j^M(Z) = \boldsymbol{\alpha}_J \cdot \mathbf{e}(\mathbf{X}_j) + E(Z) - \boldsymbol{\alpha}_J \cdot E(\mathbf{e}(\mathbf{X}_j)),$$

where

$$\boldsymbol{\alpha}_J = \mathbf{A}_j^{-1} Cov(\mathbf{e}(\mathbf{X}_j), U),$$

$\mathbf{A}_j$ being the covariance matrix of $\{e_k(\mathbf{X}_j)\}_{k=1,\ldots,M}$.

- The second step consists in running a preliminary simulation, computing empirical projections, and using the empirical projection coefficients in place of the exact ones in the actual pricing simulation:

$$P_j^{M(N)}(Z) = \boldsymbol{\alpha}_j^{(N)} \cdot \mathbf{e}(\mathbf{X}_j) + E^{(N)}(U) - \boldsymbol{\alpha}_j^{(N)} \cdot E^{(N)}(\mathbf{e}(\mathbf{X}_j)),$$

$$\boldsymbol{\alpha}_j^{(N)} = \mathbf{A}_j^{(N)-1} Cov^{(N)}(\mathbf{e}(\mathbf{X}_j), U),$$

where $\mathbf{A}_j^{(N)}$ is the empirical covariance matrix of $\{e_k(\mathbf{X}_j)\}_{k=1,\dots,M}$, obtained from an N-dimensional sample, and $Cov^{(N)}(\mathbf{e}(\mathbf{X}_j), U)_k$ is the empirical covariance of $e_k(\mathbf{X}_j)$ and U obtained from the same sample.

CONVERGENCE:

- Let us introduce the stopping times τ_j^M and $\tau_j^{M(N)}$ defined as follows:

$$\tau_L^M = L,$$

$$\tau_j^M = j1_{\{Z_j \geq P_j^M(Z_{\tau_{j+1}^M})\}} + \tau_{j+1}^M 1_{\{Z_j < P_j^M(Z_{\tau_{j+1}^M})\}},\ j = 0,\dots,L-1,$$

and

$$\tau_L^{M(N)} = L,$$

$$\tau_j^{M(N)} = j1_{\{Z_j \geq P_j^{M(N)}(Z_{\tau_{j+1}^{M(N)}})\}} + \tau_{j+1}^{M(N)} 1_{\{Z_j < P_j^{M(N)}(Z_{\tau_{j+1}^{M(N)}})\}},\ j = 0,\dots,L-1,$$

as well as the approximated option values defined as

$$U_j^M = E(Z_{\tau_j^M}),$$

$$U_j^{M(N)} = E(Z_{\tau_j^{M(N)}}).$$

Theorem (Clement, Lamberton, Protter, 2002)

- Assume that for $j = 1,....,L-1$ the sequence $e_k(\mathbf{X}_j)$ is total in $L^2(\sigma(\mathbf{X}_j))$. Then, for $j = 0,....,L$, U_j^M converges to U_j in L^2.

- Assume that for $j = 0,....,L-1$ $Q(\boldsymbol{\alpha}_j^M \cdot \mathbf{e}(\mathbf{X}_j) = Z_j) = 0$. Then, for $j = 0,....,L$, $U_j^{M(N)}$ converges to U_j^M almost surely.

Combining the two, we obtain the L^2 convergence of $U_j^{M(N)}$ to U_0.

REGRESSION NOISE

- What metric shall we use? Expected value of $U_0^{M(N)} - U_0^M$? L^2 norm of $P_j^{M(N)}(U_{j+1}) - P_j^M(U_{j+1})$? The latter, for practical reasons.

- One-period problem: Let U be a measurable variable, E an M-dimensional subspace of $L^2(\Omega, A, Q)$, P and P^N respectively the exact and the N-sample empirical projection from $L^2(\Omega, A, Q)$ into E.

- We measure the empirical regression error as the variance (L^2 norm) V^N of $P^{(N)}(U) - P(U)$. Note that V^N is a random variable as it depends on the simulation samples. Let us denote by σ^2 the variance of $U - P(U)$. Then

$$E(V^{(N)}) \cdot N \underset{N\to\infty}{\to} \sigma^2 M.$$

- Multi-period setting, the variable to be regressed is itself, in general, the result of previous regressions. Assuming that the number of basis function M is the same for all regressions, the orthogonal variance σ^2 depends then on M, as U and $P(U)$ do.

- However, we expect this dependence to be very mild, in our framework. The increment $U - P(U)$ essentially depends on the underlying variables' variation between the two dates *j* and *j+1*:

$$U = Z_{\tau_{j+1}^M},$$

$$U - P(U) = Z_{\tau_{j+1}^M} - P_j^M(Z_{\tau_{j+1}^M}),$$

$$\sigma^2 = \sigma^2_{O,M} + \sigma^2_M \approx \sigma^2_O,$$

$$\sigma^2_O = Var(Z_{\tau_{j+1}} - E(Z_{\tau_{j+1}} \mid F_j)),$$

$$\sigma^2_{O,M} = Var(Z_{\tau^M_{j+1}} - E(Z_{\tau^M_{j+1}} \mid F_j)) \underset{M\to\infty}{\to} \sigma^2_O$$

$$\sigma^2_M = Var(E(Z_{\tau_{j+1}} \mid F_j) - P^M_j(Z_{\tau^M_{j+1}})) \underset{M\to\infty}{\to} 0.$$

- Therefore, we neglect this dependency: the expected empirical regression error at each date is well approximated by

$$\frac{\sigma^2 \cdot M}{N},$$

where the variance term σ^2 is model-, product- and date-specific, and independent of M and N.

- Note that this result is different from that of Glasserman and Yu who find that the empirical regression error grows exponentially in M or faster.

VARIANCE REDUCTION IN MC SIMULATION

- The MC error is proportional to the variance of the simulated payoff. We write
$$X = (X - C) + C,$$
where C is another payoff whose PV can be computed efficiently without recurring to simulation, and such that the variance of the residual payoff *X*-*C* is smaller than that of the *X*.

- Typically, the payoff *C* is chosen to be that of another option, or that of static positions in some underlying assets.

- An alternative approach is to use dynamic control variates: the payoff *C* is given by the final value of a self financing dynamically rebalanced portfolio. While running the Monte Carlo simulation, we can use approximate greeks, and still obtain a substantial variance reduction.

- Clewlow and Cavernhill propose to use analytical deltas computed from simple models, within Monte Carlo simulations of more complex ones.

- Bouchaud, Potters and Sestovic propose to regress optimal (variance minimizing) deltas.

- We propose, within the LS-AMC framework, to compute approximate greeks from the previously computed option value approximations. Simple and flexible approach, as few extra calculations (and no extra simulations) are required if all explanatory variables are martingales or securities.

- We a little abuse of notation, we denote by $\mathbf{X}_j$ the (*D'*-dimensional) set of explanatory variables used at time *j*, $j = 0, \ldots, L$, and by

$\{e_k(\mathbf{X}_j)\}_{k=1,\ldots,M}$, the set of basis function that generate the projection space.

- To simplify the implementation, we assume from now on that all explanatory variables are securities. As a result, we can use them directly as hedging instruments.

- Let us denote by N_j the numeraire at time j, we set

$$\delta_{j,d} = \frac{\partial P_j(U_{j+1})}{\partial X_{j,d}}, j = 1,\ldots,L-1,$$

$$H_j = \sum_{d=1,\ldots,D} \delta_{j-1,d} \cdot \left(\frac{X_{j,d}}{N_j} - \frac{X_{j-1,d}}{N_{j-1}}\right), j = 2,\ldots,L,$$

$$C = \sum_{j=2,\ldots,L} H_j.$$

Note that

$$E(C) = E(H_j) = 0.$$

- Although the use of securities as explanatory variables may look as a restriction, in practice it turns out not to be the case.

VARIANCE REDUCTION INSIDE REGRESSIONS

- Starting from the theoretical result above, we look at possible ways to reduce the empirical regression error by reducing σ^2. Two strategies are proposed.

- The first is to adjust the variable to be projected, replacing the original payoff by a less volatile equivalent hedged payoff. Due to the linearity of $P(\cdot)$,
$$P(U) = P(U - C) + P(C)$$
We then look for a payoff *C* for which we can compute *P(C)* exactly, and such that the orthogonal variance of *U*-*C* is smaller than that of *U*:
$$Var(Z - C - P(Z - C)) \ll Var(Z - P(Z)).$$

- Our candidate is the approximate local hedging portfolio: we compute the standard projection ($\boldsymbol{\alpha}_j^{M(N)}$) first, then compute the approximate deltas with respect to the explanatory variables,

$$h_d = \sum_{k=1,\dots,M} \boldsymbol{\alpha}_{j,k}^{M(N)} \frac{\partial}{\partial X_d} \mathbf{e}(\mathbf{X}_j), d = 1,\dots,D',$$

and finally set

$$C = \sum_{d=1,\dots,D'} h_d \left(\frac{X_{j+1,d}}{N_{j+1}} - \frac{X_{j,d}}{N_j}\right).$$

As the explanatory variables are securities, $P(C) = 0$.

- The second proposed strategy is to enlarge the projection space first, reducing the orthogonal variance σ^2, in such a way that a second projection from this enlarged space into the original one can be computed exactly. From a mathematical point of view, we replace the empirical projection $P^{(N)}$ with a hybrid projection

$$\widetilde{P}^{(N)} = P \circ P_Y^{(N)},$$

where $P_Y^{(N)}$ denotes the empirical projection on a subspace Y of $L^2(\Omega, A, Q)$ containing E, and P the projection from form Y into E. Convergence can be proved.

- The enlargement of the projection space comes at a cost, namely $\frac{M_Y}{M_{\mathbf{E}}}$.

- We propose to use some "deltas" as orthogonal extensions of E, precisely to project on

$$Y = < \{e_k(\mathbf{X}_j)\}_{k=0,\dots,M}, \{\frac{X_{j+1,d}}{N_{j+1}} - \frac{X_{j,d}}{N_j}\}_{d=1,\dots,D'}\} > .$$

As the explanatory variables are securities,

$$P(\frac{X_{j+1,d}}{N_{j+1}} - \frac{X_{j,d}}{N_j}) = P(E(\frac{X_{j+1,d}}{N_{j+1}} - \frac{X_{j,d}}{N_j}) \mid F_j)) = 0.$$

EMPIRICAL RESULTS

- For the sake of providing a simple example, we apply the variance reduction techniques to price a Bermudan option in the Black & Scholes model. Let the risk neutral dynamics of the underlying asset S be given by

$$dS_t = \sigma S_t dW_t,$$

$$S_0 = 1,$$

with $\sigma = 20\%$, and let us assume zero rates (and dividends).

- The Bermudan option has ten years maturity, and can be exercised every year into an atm call option on S whose notional changes with time as in the following table:

Expiry (years)	1	2	3	4	5	6	7	8	9	10
Notional	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1

The value of this option (computed by PDE) is *14.82%*.

- As far as the LS-AMC pricing, we use the underlying asset S as explanatory variable, and regress the Bermudan as N-degree polynomial function of S. The simulation (and regression) dates are the ten exercise dates, plus two extra dates, one week and two months from today.

- The insertion of the first date allows us to "hedge" the Bermudan from the very beginning of the simulation (actually, from the second week). The insertion of the second date makes the first regressions hopeful.

- Variance reduction is defined as the ratio between the variance of the hedged Bermudan (Bermudan + Hedge Portfolio) and that of the Bermudan alone.

- From the following results it appears that we can easily speed up the simulation by more than ten times.

- Note that the variance reduction ratio is stable with respect to the number of simulation samples.

Samples	500	5,000	50,000	500,000
Pre-samples	15000	15,000	15,000	15,000
Polynomial degree	4	4	4	4
Bermudan	15.935%	14.733%	14.654%	14.713%
Berm std	0.950%	0.278%	0.088%	0.028%
Hegde portfolio	-0.828%	-0.032%	0.040%	0.008%
H.p. std	0.875%	0.262%	0.083%	0.026%
Hedged berm	15.107%	14.701%	14.695%	14.721%
H.b. std	0.265%	0.080%	0.025%	0.008%
Variance reduction	7.787%	8.333%	7.989%	7.966%

- This implies that if we use variance reduction level as an indicator of the quality of the approximations, we can estimate the latter more quickly than by looking at the PV of the Bermudan.

- Different polynomial degrees:

Samples	500,000	500,000	500,000	500,000	500,000	500,000	500,000	500,000	500,000
Pre-samples	15,000	15,000	15,000	15,000	15,000	15,000	15,000	15,000	15,000
Polynomial degree	2	3	4	5	6	7	8	9	10
Hedged berm	14.512%	14.715%	14.721%	14.711%	14.740%	14.757%	14.755%	14.757%	14.764%
H.b. std	0.009%	0.008%	0.008%	0.008%	0.008%	0.009%	0.014%	0.012%	0.009%
Variance reduction	10.12%	8.06%	7.97%	8.17%	8.37%	11.41%	25.49%	19.05%	11.19%

- The use of a higher degree polynomial should make the approximations more accurate, increasing the value of the option and producing better greeks and therefore a better variance reduction ratio. But also introduces more noise.

- In this case, it is better to use two different approximations, a high-order one for determining the exercise strategy (pricing), and a lower-order one for variance reduction purposes (hedging).

Pre-samples	Hedged berm	H.b. std	Variance reduction
1,000	14.648%	0.013%	22.15%
2,000	14.787%	0.010%	12.21%
3,000	14.789%	0.008%	8.90%
4,000	14.803%	0.008%	9.38%
5,000	14.802%	0.008%	8.84%
6,000	14.804%	0.008%	8.19%
7,000	14.805%	0.008%	8.44%
8,000	14.799%	0.008%	8.45%
9,000	14.803%	0.008%	8.37%
10,000	14.778%	0.008%	8.23%
12,000	14.741%	0.008%	7.99%
15,000	14.761%	0.008%	9.01%
Samples = 500,000; regression order for pricing = 8; regression order for hedging = 4.			

- Finally, we show how the iterated regression in section (the first of the two proposed strategies) can be used to improve the regression accuracy: we perform a 2nd order regression as a preliminary step to both the 8th order pricing regression, and to the 4th order hedging one. Note that, as a result, the whole procedure becomes more stable, and better price and variance reduction levels are reached.

Pre-samples	Hedged berm	H.b. std	Variance reduction
1,000	14.804%	0.008%	9.56%
2,000	14.803%	0.008%	8.23%
3,000	14.809%	0.008%	7.89%
4,000	14.811%	0.008%	8.01%
5,000	14.813%	0.008%	7.95%
6,000	14.812%	0.008%	7.97%
7,000	14.811%	0.008%	7.97%
8,000	14.806%	0.008%	7.88%
9,000	14.807%	0.007%	7.86%
10,000	14.812%	0.008%	7.90%
12,000	14.812%	0.008%	7.87%
15,000	14.812%	0.008%	7.89%
Samples = 500,000; first regression order = 2; regression order for pricing = 8; regression order for hedging = 4.			

EXTENSIONS

- Non-American claims. Even though approximating functions are not needed to price such claims, they can still be estimated in LS regression for variance reduction purposes.

- Greeks: differentiate inside the integral and hedge dynamically the new payoff.

BIBLIOGRAPHY

Clement, E., Lamberton, D., Protter, P. (2002) An analysis of a least squares regression method for American option pricing. *Finance and Stochastics* 6, pp.449-471.

Clewlow, L., Carverhill, A.(1994) On the simulation of contingent claims. *Journal of Derivatives* 2, pp.66-74.

Glasserman, P. (2004) Monte Carlo Methods in Financial Engineering. *Springer-Verlag*, Berlin.

Glasserman, P., Yu, B. (2004) Number of Paths Versus Number of Basis Functions in American Option Pricing. *Annals of Applied Probability*, 14, 4, pp.2090-2119.

Longstaff, F.A, Schwartz, E.S. (2001) Valuing American options by simulation: a simple least-squares approach. *Rev. Financial Stud.* 14, pp.113-148.

Moni, C. (2005) Fast American Monte Carlo. *SSRN* Working paper.

Piterbarg, V. (2003) A Practitioner's Guide to Pricing and Hedging Callable Libor Exotics in Forward Libor Models. *SSRN* Working paper.

Potters, M., Bouchaud, J., Sestovic, D. (2000) Hedged Monte-Carlo: low variance derivative pricing with objective probabilities. *SSRN* Working paper.