UNIVERSITY OF COLOGNE
MATHEMATICAL INSTITUTE
KÖLN, GERMANY

MATHFINANCE CONFERENCE
DERIVATIVES AND RISK MANAGEMENT
15^{TH}-16^{TH} MARCH 2010
FRANKFURT, GERMANY

A Robust Regression Monte Carlo Method for Pricing High-Dimensional American-Style Options

Christian Jonen[1]

1 Introduction

Pricing options on more than one asset with an early-exercise feature is a challenging task due to the curse of dimensionality. We propose a new class of regression-based Monte Carlo methods for pricing high-dimensional American-style options. On the basis of the dynamic programming principle (DDP) in terms of the optimal stopping time, we fit the model function for the continuation value at every exercise date by robust regression rather than by ordinary least squares. By using robust regression, we are able to get a more accurate approximation of the continuation value due to detection of outliers. To prove convergence of our Robust Regression Monte Carlo (RRM) method, we use techniques of the statistical learning theory. Especially in the multidimensional case, taking bases with many functions are vital for an accurate option price estimation. To avoid that unfavorable increase in complexity, we propose a new technique for calculating the coefficients of the model function for the continuation value. Our approach runs with few basis functions, and it turns out that our suggested estimator is nearly unbiased. To improve the convergence of regression-based Monte Carlo estimators, we extend earlier results for variance reduction via importance sampling for American-style options. In comparison to existing Monte Carlo methods, we can improve convergence significantly by implementing our proposed approaches.

2 A Robust Regression Monte Carlo Method

To begin with, we introduce our RRM algorithm in subsection 2.1. We propose a way of reducing variance by importance sampling for American-style options in subsection 2.2. Finally, we illustrate our suggested approaches by showing some test results of a comprehensive option pricing study in subsection 2.3.

2.1 Algorithm

Based on some standard assumptions, the fair price of an American (or Bermudan) option at time t_0 is given by

$$\sup_{\tau \in \mathcal{T}_{0,L}} \mathrm{E}_Q[e^{-r\tau\triangle t} Z_\tau | \mathcal{F}_0], \tag{1}$$

where $\mathcal{T}_{0,L}$ is the set of all stopping times with values in $\{0, ..., L\}$; $(Z_l)_{0\leq l\leq L}$ is an adapted payoff process, and Q is the riskneutral probability measure, see [2]. We are able to solve the optimal stopping problem (1) by using the DDP in terms of the value process itself – that

---

[1] Corresponding author, E-mail: cjonen@math.uni-koeln.de

is the common way – or, alternatively, by using the DDP in terms of the optimal stopping time given by

$$\begin{aligned}\tau_L &= L\\ \tau_l &= \begin{cases} l & , Z_l \geq \mathrm{E}_Q[e^{-r(\tau_{l+1}-l)\triangle t} Z_{\tau_{l+1}} | \mathcal{F}_l] =: C_l \\ \tau_{l+1} & , \text{otherwise} \end{cases} \quad , l = L-1, ..., 0 \end{aligned} \tag{2}$$

– that is exactly the idea of the Least Squares Monte Carlo (LSM) method suggested in [6]. In so doing, we obtain the fair option price at time t_0 by

$$V_0 = \mathrm{E}_Q[e^{-r\tau_0 \triangle t} Z_{\tau_0} | \mathcal{F}_0]. \tag{3}$$

On the basis of the set of N simulated paths of our underlying model denoted by $\mathbb{S} = \{S_{nl}\}_{l=1,...,L}^{n=1,...,N}$, the idea of our RRM method is to approximate the continuation value C_l in (2) by specifying a loss function $\ell(\cdot)$ and solving the following regression problem in every time step:

$$\min_{\overline{x} \in \mathbb{R}^M} \frac{1}{J} \sum_{j=1}^{J} \ell\left(C_l^j - \overline{C}_l^j\right), \tag{4}$$

where J is the number of in-the-money paths, $C_l^j = e^{-r(\tau_{l+1}^j - l)\triangle t} Z_{\tau_{l+1}^j}^j$ are the regressands and $\overline{C}_l^j = \sum_{m=0}^{M-1} \overline{x}_m \phi_m(S_{jl})$ is the model function for C_l with a basis $\{\phi_m(\cdot)\}_{m=0}^{M-1}$, $j = 1, ..., J$; we denote by τ_l^j and Z_l^j the approximated optimal stopping time and the payoff of path j, $j = 1, ..., J$, at time t_l, $l = 1, ..., L$, respectively. Some loss functions leading to a robust regression problem are listed in table 1, where γ, γ_i, $i = 0, ..., 3$, are transition points defining outliers; we give some simple rules for determining these points. Note that the LSM algorithm can be seen as a special case of our proposed method by selecting $\ell(r) = 0.5r^2$. As we can see in figure 1, there are some points, namely the red points, which are really far away from the gray surface. To this end, we motivate our approach by giving these outliers caused by strongly fluctuated paths fewer weight than the other points. At time t_0 we estimate the continuation value by its empirical mean $\overline{C}_0 = \frac{1}{N}\sum_{n=1}^{N} e^{-r\tau_1^n \triangle t} Z_{\tau_1^n}^n$. To guarantee an efficient implementation, solving robust regression problems needs special care. For this purpose, we pay attention on Newton-based methods and propose a new approach for solving robust regression with very good numerical properties. By doing so, we show local quadratic convergence for our proposed method. In order to prove convergence of our RRM estimator, we use techniques of the statistical learning theory. Based on an appropriate error decomposition, we show the consistency of our RRM estimator and formulate error bounds.

Table 1: Objective functions $\ell(\cdot)$ for robust regression.

	$\ell(\cdot)$
Huber	$\begin{cases} 0.5r^2 & , \lvert r\rvert \leq \gamma \\ \gamma\lvert r\rvert - 0.5\gamma^2 & , \lvert r\rvert > \gamma \end{cases}$
Jonen	$\begin{cases} \gamma_0 & , r \leq \gamma_0 \\ \gamma_1\lvert r\rvert - 0.5\gamma_1^2 & , \gamma_0 < r \leq \gamma_1 \\ 0.5r^2 & , \gamma_1 < r < \gamma_2 \\ \gamma_2\lvert r\rvert - 0.5\gamma_2^2 & , \gamma_2 \leq r < \gamma_3 \\ \gamma_3 & , r \geq \gamma_3 \end{cases}$

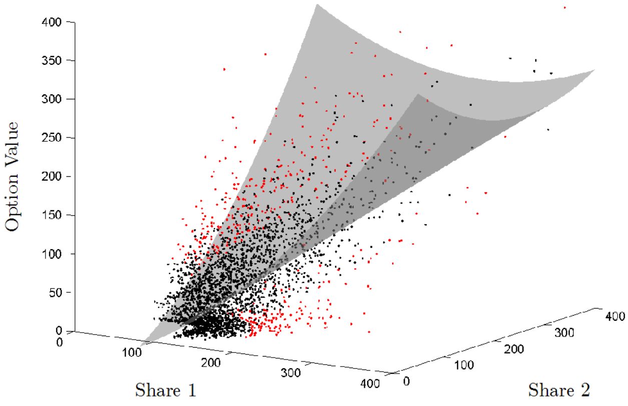

Figure 1: Approximation of the continuation value by least squares for a Maximum call option on two assets.

2.2 Variance Reduction via Importance Sampling

In order to improve the convergence behavior of our RRM estimator, we concentrate on variance reduction by a change of drift in Brownian motion of the underlying model. For this purpose, we assume that our underlying is driven by an It$\widehat{o}$ process under Q. By exploring Girsanov's theorem we are able to reformulate our option price (3) at t_0 as

$$V_0 = \mathrm{E}_Q[e^{-r\tau_0 \triangle t} Z_{\tau_0} | \mathcal{F}_0] \;\; = \;\; \mathrm{E}_{\widetilde{Q}}[e^{-r\tau_0 \triangle t} Z_{\tau_0} exp\{-\theta^T W_{\tau_0} + 0.5\|\theta\|^2 \tau_0\} | \mathcal{F}_0] \tag{5}$$

where W_t and $\widetilde{W_t} = W_t - \theta dt$ are standard Brownian motions under Q and $\widetilde{Q}$, respectively. Assuming that our optimal stopping rule τ_1 is known, we get an unbiased estimator of our continuation value at time date t_0 by approximating the quantity on the right-hand side of (5) by taking the average of the independent and identically distributed samples $e^{-r\triangle t \tau_1} Z_{\tau_1} exp\{-\theta^T W_{\tau_1} + 0.5\|\theta\|^2 \tau_1\}$. To find the drift θ minimizing variance of our new estimator, we derive the following optimization problem:

$$\min_{\theta \in \mathbb{R}^D} E_Q \left[e^{-r\tau_1 \triangle t} Z_{\tau_1} exp\{-\theta^T W_{\tau_1} + 0.5\|\theta\|^2 \tau_1\} \right] \tag{6}$$

It can be shown that the minimum of (6) exists and is unique. Solving our stochastic optimization problem (6) by stochastic approximation suggests itself. Since a sub-linearity condition is not fulfilled in our mathematical context, we use a truncated version of Robbins-Monro algorithm suggested in [3]. Alternatively, by generating independent and identically distributed samples, we are able to replace the expectation of task (6) by its empirical mean. We show that an optimal solution $\widehat{\theta}_N$ of the resulting deterministic optimization task provides an approximation of the exact optimal solution of problem (6). Following, the drift $\widehat{\theta}_N$ can be calculated by optimization methods, for instance by the (Quasi-) Newton-Raphson method. Compared to existing variance reduction techniques via importance sampling, our approach is based on a known stopping rule.

2.3 Experiments

The structure of the continuation value is often difficult to estimate, especially for options with a non-smooth payoff. That is why we propose a new technique called Split for calculating coefficients to get a better approximation locally. The idea is to split the domain of asset values into several subdomains and to estimate the continuation value for each subdomain separately. In order to illustrate the performance of our proposed approaches, table 2 and both figures, 2 and 3, show test results of a comprehensive study of different type of multi-asset options with an early-exercise feature. Table 2 indicates that our RRM method in combination with Split and our suggested loss function nearly eliminate bias for pricing a Bermudan call option on two assets. As we can see in figure 2, we are able to obtain a remarkable convergence improvement by using regression-based Monte Carlo methods in combination with Split and IS2 or IS3 for pricing a Bermudan max put option on five assets;

Table 2: Bias calculated by the LSM and RRM methods for a Bermudan max call option on two assets in multi-asset Black Scholes model.

S_0	LSM	LSM+Split	RRM+Split
90	-0.0294 (0.0343)	-0.0092 (0.0340)	-0.0027 (0.0353)
100	-0.0402 (0.0392)	-0.0102 (0.0371)	0.0014 (0.0366)
110	-0.0969 (0.0456)	-0.0432 (0.0428)	0.0007 (0.0429)

Notes. Characteristics of the option are: $r = 0.05, T = 3, \delta_d = 0.1, \sigma_d = 0.2$ *for all d, strike=100, nine exercise opportunities. The number of paths is 120000 and antithetic variables (AV) are used for variance reduction. Each bias is determined by fifty runs with different seeds in the random number generator and the following benchmark values:* 8.0931, 13.9018 *and* 21.3439 *for* $S_0 = 90$, $S_0 = 100$ *and* $S_0 = 110$, *respectively. Benchmark values result from the three-dimensional Binomial tree proposed in [7] with 9000 time steps. Values in parantheses are standard errors.*

we denote by IS1,IS2 and IS3 our importance sampling approach adapted from a truncated version of Robbins-Monro algorithm, Newton-Raphson and Quasi-Newton-Raphson method, respectively. Notify that variance reduction by IS2 and IS3 is superior to variance reduction by IS1 and AV. Furthermore, finding the drift minimizing variance seems to be more efficient by generating samples and solving the resulting deterministic optimization task, rather than using stochastic approximation. Even though the LSM algorithm with IS2 shows a nice convergence behavior, we obtain a slight acceleration by using our RRM method with IS3 and the Huber function, see figure 3. All illustrated experiments are based on the basis $\{1, S_1, ..., S_D, S_1^2, ..., S_D^2\}$ increasing linearly with number of assets D. We need some additional operations by using our RRM method – due to the iterative solver for robust regression – and by using our importance sampling approaches – due to searching the optimal drift. For this reason, figure 2 and 3 show the CPU time versus the root mean square error (RMSE). However, a comparison of RMSE against number of paths shows that our methods still perform better. A logical next step is to determine upper bounds; our comprehensive option pricing study involves for instance the Andersen-Broadie method proposed in [1]. Test results indicate that we are able to get tight confidence intervals by running our approaches in combination with their method. Anyway, the case studies illustrated here show that our approaches are catalysts for convergence.

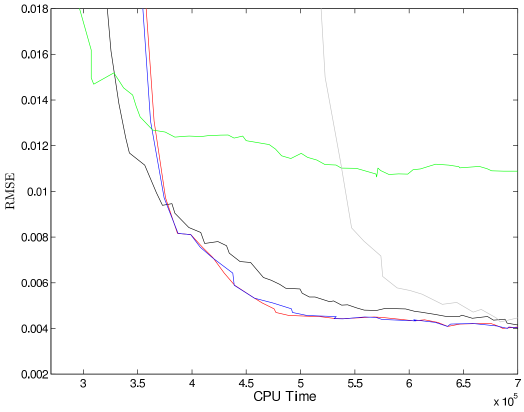

Figure 2: RMSE against CPU time for the LSM method in combination with AV, IS1, IS2, IS3 and Split for a Bermudan max put option on five assets.

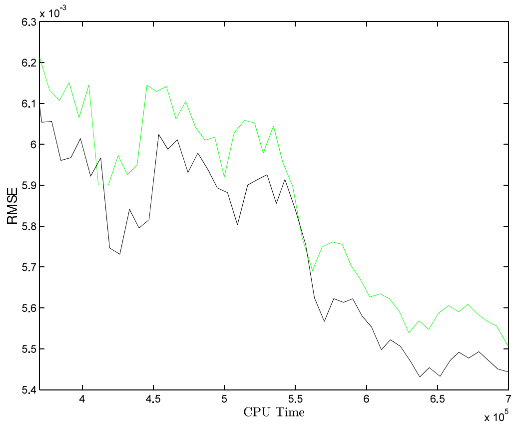

Figure 3: RMSE against CPU time for the LSM and RRM method with IS2 for a Bermudan max put option on four assets.

Notes. First, we approximate the optimal stopping rule by determining coefficients of the model function for C_l with 130000 paths; in addition, in figure 2 we use antithetic variables for reducing variance. Based on this approximated stopping rule we calculate option prices for an increasing number of paths as follows: In figure 2 the green and black lines show the RMSE by using the LSM method with and without Split, respectively, in combination with antithetic variables. The grey, red and blue lines show the RMSE by using the LSM method with IS1, IS2 and IS3, respectively, in combination with Split. In figure 3 the green and black lines show the RMSE by using the LSM and RRM methods, respectively, in combination with Split and IS2. Parameters of the option are: $T = 1$, $r = 0.045$, strike= 100, $S_0^d = 100$ for all d,

$$\sigma = \begin{pmatrix} 0.25 \\ 0.35 \\ 0.2 \\ 0.25 \\ 0.2 \end{pmatrix}, \ \delta = \begin{pmatrix} 0.05 \\ 0.07 \\ 0.04 \\ 0.06 \\ 0.04 \end{pmatrix}, \ P = \begin{pmatrix} 1.00 & -0.65 & 0.25 & 0.20 & 0.25 \\ -0.65 & 1.00 & 0.50 & 0.10 & 0.25 \\ 0.25 & 0.50 & 1.00 & 0.37 & 0.25 \\ 0.20 & 0.10 & 0.37 & 1.00 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 1.00 \end{pmatrix},$$

where P is the matrix of correlations. The approximated benchmark values are 1.378 for the option on five assets and 1.841 for the option on four assets, see [5].

3 Conclusion

We have presented the key ideas of our RRM method generalizing the LSM method. Due to taking outliers into account, we are able to reduce bias contrary to the LSM algorithm. Our convergence proof as well as our proposed robust regression solver allow for switching between least squares and robust regression at different time dates. We have proposed a loss function adapting to unknown distributions; other loss functions might be considered. A further improvement in detecting outliers might be to work with natural bounds for options. Our suggested splitting technique allows for running regression-based methods with few basis functions without convergence losses. We have applied variance reduction by a change of drift in Brownian motion successfully. Nevertheless, we will test other stochastic approximation approaches rather than a truncated version of Robbins-Monro algorithm for searching the optimal drift. To sum up, our proposed approaches accelerate convergence significantly. Although we have shown in [4] that an efficient random number generator in combination with a variance reduction technique may be superior to quasi-random numbers, we will extend our comprehensive option pricing study by using different types of random numbers with low discrepancy. A further possible area of future research might be to apply our ideas to nonparametric regression.

References

[1] L. Andersen and M. Broadie. Primal-dual simulation algorithm for pricing multidimensional american options. *Management Science*, 52:1222–1234, 2004.

[2] A. Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematicae*, 2:139–158, 1984.

[3] A.-J. Gao H.-F. Chen and L. Guo. Convergence and robustness of the robbins-monro algorithm truncated at randomly varying bounds. *Stochastic Processes and their Applications*, 27:217–231, 1988.

[4] C. Jonen. An efficient implementation of a least squares monte carlo method for valuing american-style options. *International Journal of Computer Mathematics*, 86:1024–1039, 2009.

[5] C.C.W. Leentvaar and C.W. Oosterlee. Multi-asset option pricing using a parallel fourier-based technique. *Journal of Computational Finance*, 12, 2008.

[6] F.A. Longstaff and E.S. Schwartz. Valuing american options by simulation: A simple least-squares approach. *Review of Financial Studies*, 14:113–147, 2001.

[7] M. Rubinstein. Return to oz. *Risk*, 11:67–71, 1994.