

# On the stability the least squares Monte Carlo

Oleksii Mostovyi

Received: 20 January 2011 / Accepted: 11 October 2011 / Published online: 20 November 2011  
© Springer-Verlag 2011

**Abstract** Consider least squares Monte Carlo (LSM) algorithm, which is proposed by Longstaff and Schwartz (Rev Financial Studies 14:113–147, 2001) for pricing American style securities. This algorithm is based on the projection of the value of continuation onto a certain set of basis functions via the least squares problem. We analyze the stability of the algorithm when the number of exercise dates increases and prove that, if the underlying process for the stock price is continuous, then the regression problem is ill-conditioned for small values of the time parameter.

**Keywords** Option pricing · Optimal stopping · American option · Least squares Monte Carlo · Monte Carlo methods · Stability · Ill-Conditioning

## 1 Introduction

The problem of pricing American and Bermudan style options is of fundamental importance in option pricing theory. In the continuous time settings McKean [7] proposed an algorithm for pricing an American put option with an infinite maturity via Ordinary Differential Equations and Partial Differential Equations. Further developments of this technique and applications to other American style securities are considered in Peskir and Shiryaev [9]. Another approach is pricing via Monte-Carlo simulations that is described by Glasserman [2]. One of the most difficult tasks in the theory of pricing of American and Bermudan options is the determination of an optimal stopping rule and the valuing of the option under such a rule. Longstaff and Schwartz [5] proposed an algorithm for pricing American and Bermudan style options

---

O. Mostovyi (✉)  
Department of Mathematical Sciences, Carnegie Mellon University,  
Pittsburgh, PA 15213, USA  
e-mail: omostovy@andrew.cmu.edu

via Monte Carlo simulations, least squares Monte Carlo or LSM. This technique is especially useful when we deal with multi-factor processes. In this case the methods based on binomial, trinomial trees, or partial differential equations become slow and thus inefficient due to the high dimensionality of the problem.

As in the majority of the numerical algorithms the starting point of LSM for American options is a substitution of the continuous time interval with a discrete set of exercise dates. Practically, by doing this we substitute the American option with a Bermudan one. Then for each exercise time (except the first and the last one) we project the value of continuation onto a set of basis functions via linear regression. Clement, Lamberton and Protter [1] investigated the convergence of the algorithm with the growth of the number of the basis functions and the Monte Carlo simulations. Under fairly general conditions they proved the almost sure convergence of the complete algorithm. Also, they obtained the rate of convergence when the number of Monte Carlo simulations increases and showed that the normalized error of the algorithm is asymptotically Gaussian. However, they considered a fixed partition of the time interval and thus, essentially, they discussed the properties of the Bermudan, not American option. Glasserman and Yu [3] investigated the behavior of LSM with the simultaneous grows of the number of the basis functions and the number of the Monte-Carlo simulations and estimated the rate of convergence in some more specific settings. Moreno and Navas [8] considered the LSM for different basis functions, namely, power series, Laguerre, Legendre, Chebyshev polynomials, and deduced that the algorithm converges at least for American put options when the underlying problem has a small number of factors. Stentoft [10] obtained the rate of convergence of the algorithm in the two period multidimensional case.

In the present work we consider the stability of LSM algorithm, when the number of exercise dates increases in such a way that there are exercise dates close to an initial time, which we assume to be equal to zero without loss of generality. We prove that the algorithm is unstable when the time parameter is close to zero, because the underlying regression problem is ill-conditioned.

The remainder of this work is organized as follows. In Sect. 2, we describe the algorithm. In Sect. 3, we prove the main result, which is formulated in Proposition 1, instability of the algorithm for the small values of the time parameter due to the ill-conditioning of the corresponding matrix in the regression problem. In addition we present the results of the numerical simulations that illustrate the assertions of Proposition 1. In Sect. 4, we give the concluding remarks.

## 2 Description of the algorithm

Assume that the stock price process  $X = \{X_t\}_{t \in [0, T]}$  is given by the strong solution to the following Stochastic Differential Equation:

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t dW_t, \\ X_0 &= x_0. \end{aligned} \tag{1}$$

Here  $x_0 \in \mathbb{R}_{++}$  is a constant,  $W = \{W_t\}_{t \in [0, T]}$  is a Brownian Motion (possibly multidimensional) on a filtered probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$ , where  $\mathbb{P}$  is the risk-neutral probability measure,  $\mu$  and  $\sigma$  are the progressively measurable functionals, such that the strong solution to Eq. (1) exists and unique on the time interval  $[0, T]$ , see Karatzas and Shreve [6], Chapter 5, for the discussion of this topic. The time horizon is  $T$ , which we assume to be a finite constant. Usually equations of such a form are used to describe the evolution of the stock prices in practice. Let the payoff of an American option at the time of exercise  $\tau$  is given by  $\varphi(X_\tau)$ , where  $\varphi$  is the corresponding payoff function. Then the value of the option is determined by the formula:

$$v(X_0) \triangleq \sup_{\tau \in \mathcal{A}} \mathbb{E}[e^{-\int_0^\tau r_u du} \varphi(X_\tau)].$$

Here  $r = \{r_t\}_{t \in [0, T]}$  is an interest rate process, which we assume to be deterministic for simplicity;  $\mathcal{A}$  is the set of the stopping times with respect to the filtration  $\{\mathcal{F}_t\}_{t \in [0, T]}$ .

To approximate  $v(X_0)$  numerically let us conduct  $N$  Monte-Carlo simulations of the process  $X$ . First, we need to divide the time interval  $[0, T]$  into  $M$  subintervals  $[t_m, t_{m+1}]$  of the length  $\Delta t \triangleq \frac{T}{M}$ , where  $t_m \triangleq \frac{T}{M}m, m = 1, \dots, M$ . Thus at every moment  $t_m$  we obtain  $N$  realizations of the process  $X_{t_m}^n, n = 1, \dots, N$ . Second, for each simulation we compute the value of the option at time  $t_M = T$  (under the assumption that the option was not exercised before  $T$ ):

$$C_{t_M}^n \triangleq \varphi(X_{t_M}^n), \quad n = 1, \dots, N.$$

Discounting these values we get a cash flow vector

$$b_{M-1} \triangleq d(C_{t_M}^1, \dots, C_{t_M}^N)^T$$

where  $d \triangleq \exp(-\int_{t_{M-1}}^{t_M} r_u du)$  is the discount factor.

To obtain the value of the option at  $t_{M-1}$ ,  $C_{t_{M-1}}$  (under the assumption the option was not exercised before  $t_{M-1}$ ), we chose a hypothesis of linear regression and project the cash flow vector  $b_{M-1}$ , for example, on a constant,  $X_{t_{M-1}}$ , and  $X_{t_{M-1}}^2$ . According to [5] this is one of the simplest yet successful regression models. According to [8] a good alternative choice of basis functions can be Hermite, Laguerre, Legendre, or Chebyshev polynomials. If we use  $1, X_{t_{M-1}}$ , and  $X_{t_{M-1}}^2$  as the basis, the estimate of the conditional expectation becomes

$$\mathbb{E}[C_{t_{M-1}} | \mathcal{F}_{t_{M-1}}] = \alpha + \beta X_{t_{M-1}} + \gamma X_{t_{M-1}}^2, \tag{2}$$

where  $\alpha, \beta, \gamma$  are some constants. Then along every path we compare values of immediate exercise,  $\varphi(X_{t_{M-1}}^n)$ , with values of continuation that are obtained by substitution of  $X_{t_{M-1}}^n$  into Eq. (2). The bigger of two gives  $C_{t_{M-1}}^n, n = 1, \dots, N$ . If value of immediate exercise is bigger we set  $C_{t_M}^n = 0$ .

Similarly we obtain  $\mathbb{E}[C_{t_m} | \mathcal{F}_{t_m}]$  for each  $m \in \{M - 1, \dots, 1\}$  via solving linear regression problems  $A(t_m)x(t_m) = b(t_m)$ , where components of the matrix  $A(t_m)$

depend on the regression hypothesis and the outcomes of the simulations,  $x(t_m)$  is an unknown vector of coefficients, and the vector  $b(t_m)$  is given by equation

$$b_{t_m} = d(C_{t_{m+1}}^1, \dots, C_{t_{m+1}}^N), \tag{3}$$

where  $d = \exp(-\int_{t_m}^{t_{m+1}} r_u du)$  is the discount factor.

Finally we discount the cash flow up to the moment of time  $t_0 = 0$  and compare it with the value of immediate exercise at time  $t_0$ ,  $\varphi(X_{t_0})$ . The bigger is the value of the option.

### 3 Ill-Conditioning for small $t$

Let  $\mathcal{P} = \{t_m = \frac{T}{M}m : m = 1, \dots, M - 1\}$  be a partition of the interval  $[T/M, T(M - 1)/M]$ . In order to compute the estimates of the value of the option at each  $t \in \mathcal{P}$  (under the assumption that it was not exercised before) we solve the linear regression problem

$$A(t)x(t) = b(t), \tag{4}$$

where  $x(t)$  is an unknown vector of coefficients, vector  $b(t)$  is determined by Eq. (3), and the matrix  $A(t)$  depends on the regression hypothesis and the outcome of the Monte Carlo simulations. Assume that we have chosen  $K$  continuous functions  $f_1, \dots, f_K$  as the hypothesis. Examples of such functions are power series, Laguerre, Legendre, Hermite polynomials, etc. In this case  $A(t)$  has the following form

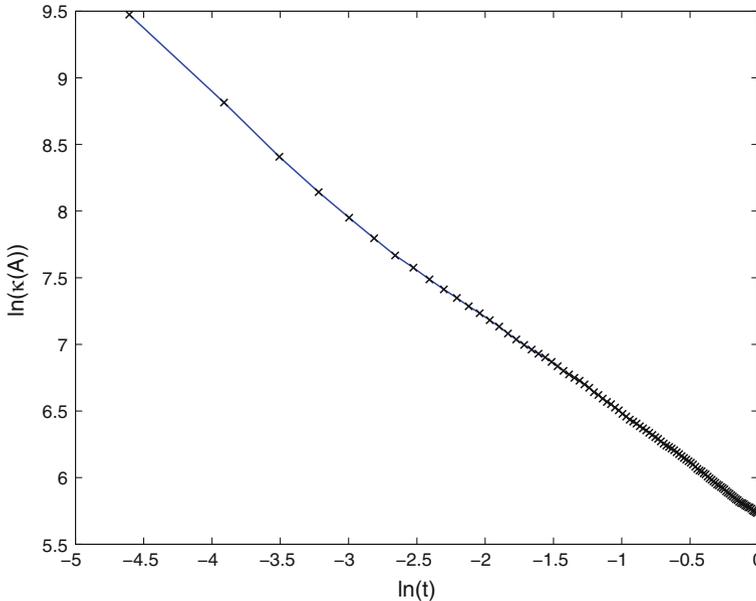
$$A(t) \triangleq \begin{pmatrix} f_1(X_t^1) & \dots & f_K(X_t^1) \\ \dots & \dots & \dots \\ f_1(X_t^N) & \dots & f_K(X_t^N) \end{pmatrix}, \tag{5}$$

where  $N$  is the number of the simulations. We show that, if the underlying process  $X$  is almost surely continuous, then for small  $t$  the problem (4) is ill-conditioned. For a matrix  $A$  in the  $l_2$  norm the *condition number* is defined as

$$\kappa(A) \triangleq \|A\| \cdot \|A^{-1}\| = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}, \tag{6}$$

where  $\sigma_{max}(A)$  and  $\sigma_{min}(A)$  are maximal and minimal singular values of the matrix  $A$  respectively. A problem with a low condition number is called *well-conditioned*, while a problem with a high condition number is called *ill-conditioned*.

Usually, problem (4) is solved via one of the following methods: Householder triangularization, Gram–Schmidt orthogonalization, singular value decomposition, or normal equations. Let  $\kappa$  be the condition number of matrix  $A$ . If  $(A^T A)^{-1}$  exists then the exact solution to the least-squares problem is given by the vector  $x = (A^T A)^{-1} A^T b$ , i.e. it is a product of the left-inverse of the matrix  $A$  and the vector  $b$ . One can see



**Fig. 1**  $\ln(\kappa(A))$  as a function of  $\ln(t)$ ,  $t \in (0.009, 1]$ ,  $X$ -lognormal process,  $\mu \equiv 0$ ,  $\sigma \equiv 0.15$ , 30,000 paths, 100 time steps, Milstein discretization scheme

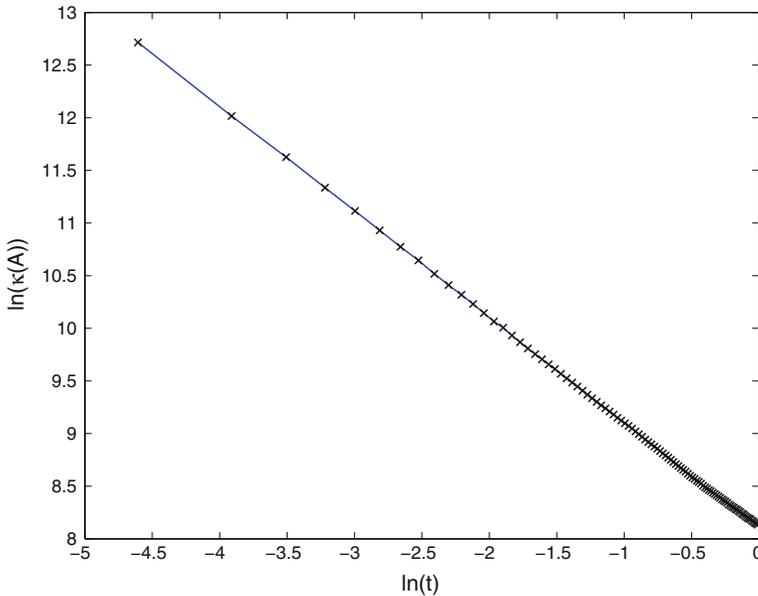
that the solution to the problem (4), obtained via normal equations, is governed by  $\kappa^2$ , whereas the solution obtained via SVD, Householder or Gram–Schmidt is governed by  $\kappa$ . Consequently, normal equations are the least stable with respect to the grows of the condition number. Nevertheless, the analytical solution to the least-squares problem is defined in terms of the normal equations.

Let  $\kappa(t)$  denote the condition number of the matrix  $A(t)$  for  $t \in [0, T]$ . We show below that  $\lim_{t \downarrow 0} \kappa(t) = \infty$   $\mathbb{P}$ -almost surely. Therefore, no matter what algorithm one uses (Householder, Gram–Schmidt, SVD or normal equations), for small values of time the underlying regression problem is ill-conditioned, and thus the algorithm is unstable.

We prove ill-conditioning for an arbitrary number of basis functions in the following proposition. In addition, we illustrate the phenomenon with the results of the numerical simulations in the case when regression is done on three basis functions  $1, x$ , and  $x^2$  (Fig. 1).

**Proposition 1** *Assume that the process  $X$  is given by Eq. (1),  $f_k, k = 1, \dots, K$ , are continuous functions, such that  $\sum_{k=1}^K f_k^2(X_0) > 0$ . Let for each  $t \in [0, T]$  the matrix  $A(t)$  is defined by Eq. (5) and  $\kappa(t)$  is the condition number of  $A(t)$ . Then*

$$\mathbb{P} \left[ \lim_{t \rightarrow 0^+} \kappa(t) = \infty \right] = 1. \tag{7}$$



**Fig. 2**  $\ln(\kappa(A))$  as a function of  $\ln(t)$ ,  $t \in (0.009, 1]$ ,  $X$ -normal process,  $\mu \equiv 0$ ,  $\sigma \equiv 0.03$ , 30,000 paths, 100 time steps, Euler discretization scheme

*Proof* Consider Eq. (5). It follows from Eq. (1) that all rows of  $A(0)$  are identical. Consequently the rank of  $A(0)$  equals to 1. Let us look at the following matrix:

$$\frac{(A^T A)(0)}{N} = \begin{pmatrix} f_1^2(X_0) & f_1(X_0)f_2(X_0) & \cdots & f_1(X_0)f_K(X_0) \\ f_2(X_0)f_1(X_0) & f_2^2(X_0) & \cdots & f_2(X_0)f_K(X_0) \\ \cdots & \cdots & \cdots & \cdots \\ f_K(X_0)f_1(X_0) & f_K(X_0)f_2(X_0) & \cdots & f_K^2(X_0) \end{pmatrix}. \quad (8)$$

The matrix  $(A^T A)(0)/N$  has two eigenvalues:  $\sum_{k=1}^K f_k^2(X_0)$  is the first one (with multiplicity 1), 0 is the second one (with multiplicity  $K - 1$ ). Thus  $A(0)$  has singular values  $\sqrt{N \sum_{k=1}^K f_k^2(X_0)}$  and 0, consequently  $\kappa(0) = \infty$  (Fig. 2).

Note that the matrix  $(A^T A)(t)$  is real and symmetric, thus the eigenvalues of  $(A^T A)(t)$  are real for all  $t$ 's. Since eigenvalues are continuous functions of the components of the matrix and the components in turn are a.s. continuous processes, we deduce that for small  $t$ 's the first eigenvalue is in the neighborhood of  $\sum_{k=1}^K f_k^2(X_0)N$ , which is greater than zero by the assumption of proposition, whereas all other eigenvalues are in the neighborhood of 0. The conclusion,  $\mathbb{P}[\lim_{t \rightarrow 0^+} \kappa(t) = \infty] = 1$  follows from continuity of the underlying process  $X$  and the basis functions  $f_k$ 's,  $1 \leq k \leq K$ .  $\square$

Intuitively Proposition 1 shows that for small values of  $t$  the condition number is large.

## 4 Concluding remarks

We proved that for a continuous underlying stock price process, LSM algorithm for pricing American options is unstable when time parameter is small. An interesting question is to obtain an exact bound on applicability of this algorithm. A possible criterion of applicability is the condition number of matrix  $A(t)$ ,  $\kappa(t)$ . For example, if  $\kappa(t)$  exceeds a certain value, one can treat (4) as a rank deficient least squares problem (see [4] for details), or switch to another method: backward induction or the method introduced by McKean [7] of option pricing via Ordinary Differential Equations or Partial Differential Equations considered on a smaller domain. For certain problems it is possible to obtain desired accuracy using relatively small number of time intervals, then one does not have to solve the regression problem for small  $t$ 's, and consequently the algorithm can be stable. Also, if the underlying process  $X$  is discontinuous with high probability the algorithm can be stable even for small values of the time parameter.

## References

1. Clement, E., Lamberton, D., Protter, P.: An analysis of a least squares regression method for American option pricing. *Finance Stoch.* **6**, 449–471 (2002)
2. Glasserman, P.: *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York (2004)
3. Glasserman, P., Yu, B.: Number of paths versus number of basis functions in American option pricing. *Ann. Appl. Probab.* **14**(4), 2090–2119 (2004)
4. Golub, G.H., Van Loan, C.F.: *Matrix Computations*, 3rd edn. Johns Hopkins University Press, Baltimore and London (1996)
5. Longstaff, F., Schwartz, E.: Valuating American option by simulation: a simple least-squares approach. *Rev. Financial Studies* **14**, 113–147 (2001)
6. Karatzas, I., Shreve, S.: *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer-Verlag, New York (1998)
7. McKean, H.P. Jr.: Appendix: a free boundary problem for the heat equation arising from a problem in mathematical economics. *Ind. Manag. Rev.* **6**, 32–39 (1965)
8. Moreno, M., Navas, J.F.: On the Robustness of Least-Squares Monte Carlo (LSM) for Pricing American Options. Mimeo, Universitat Pompeu Fabra, Barcelona (2001)
9. Peskir, G., Shiryaev, A.: *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics. ETH Zurich, Birkhauser, Basel (2006)
10. Stentoft, L.: Convergence of the least squares Monte Carlo approach to American option valuation. *Manag. Sci.* **50**(9), 1193–1203 (2004)