Journal of Statistical Planning and Inference 142 (2012) 2289–2307

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Pricing of American options in discrete time using least squares estimates with complexity penalties ☆

Michael Kohler [a], Adam Krzyżak [b,*]

[a] Department of Mathematics, Technische Universität Darmstadt, Schlossgartenstr. 7, D-64289 Darmstadt, Germany

[b] Department of Computer Science and Software Engineering, Concordia University, 1455 De Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8

ARTICLE INFO

Article history:
Received 16 April 2009
Received in revised form
11 February 2012
Accepted 14 February 2012
Available online 25 February 2012

Keywords:
American options
Neural networks
Nonparametric regression
Optimal stopping
Orthogonal series estimates
Rate of convergence
Regression based Monte Carlo methods
Smoothing splines
Splines

ABSTRACT

Pricing of American options in discrete time is considered, where the option is allowed to be based on several underlying stocks. It is assumed that the price processes of the underlying stocks are given by Markov processes. We use the Monte Carlo approach to generate artificial sample paths of these price processes, and then we use nonparametric regression estimates to estimate from this data so-called continuation values, which are defined as mean values of the American option for given values of the underlying stocks at time t subject to the constraint that the option is not exercised at time t. As nonparametric regression estimates we use least squares estimates with complexity penalties, which include as special cases least squares spline estimates, least squares neural networks, smoothing splines and orthogonal series estimates. General results concerning rate of convergence are presented and applied to derive results for the special cases mentioned above. Furthermore the pricing of American options is illustrated by simulated data.

1. Introduction

Monte Carlo methods for pricing of American options are very attractive in comparison to other methods in case of options which are based on several (correlated) stocks (so-called basket options), because in this case in the standard approach, in which the whole problem is reformulated as a free boundary problem for partial differential equations (cf., e.g., Chapter 8 in Elliott and Kopp, 1999), the numerical solution of this free boundary problem gets very complicated. And alternative methods based on binomial trees (cf., e.g., Chapter 1 in Elliott and Kopp, 1999) are in practice not able to model the correlation structure of more than two stocks correctly.

In this paper we consider American options in discrete time (sometimes also called Bermudan options). The price of such an option can be represented in a risk neutral market as a solution of an optimal stopping problem

$$V_0 = \sup_{\tau \in \mathcal{T}(0,\dots,T)} \mathbf{E}\{f_\tau(X_\tau)\}, \tag{1}$$

where f_t is the (discounted) payoff function, $X_0, X_1, \dots, X_T$ is the underlying stochastic process describing e.g., the prices of the underlying assets and the financial environment (like interest rates, etc.) and $\mathcal{T}(0,\dots,T)$ is the class of all

☆ Research sponsored by NSERC Grant RGPIN 270-2010.

* Corresponding author. Tel.: +1 514 848 2424x3007; fax: +1 514 848 2830.
E-mail addresses: kohler@mathematik.tu-darmstadt.de (M. Kohler), krzyzak@cs.concordia.ca (A. Krzyżak).

0378-3758/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2012.02.031

$\{0,\ldots,T\}$-valued stopping times, i.e., $\tau \in \mathcal{T}(0,\ldots,T)$ is a measurable function of $X_0,\ldots, X_T$ satisfying

$$\{\tau = \alpha\} \in \mathcal{F}(X_0,\ldots,X_\alpha) \quad \text{for all } \alpha \in \{0,\ldots,T\}.$$

In the sequel we assume that $X_0, X_1,\ldots, X_T$ is a $\mathbb{R}^d$-valued Markov process recording all necessary information about financial variables including prices of the underlying assets as well as additional risk factors driving stochastic volatility or stochastic interest rates. Neither the Markov property nor the form of the payoff as a function of the state X_t is restrictive and can always be achieved by including supplementary variables.

The computation of (1) can be done by the determination of an optimal stopping rule $\tau^* \in \mathcal{T}(0,\ldots,T)$ satisfying

$$V_0 = \mathbf{E}\,\{f_{\tau^*}(X_{\tau^*})\}. \tag{2}$$

Let

$$q_t(x) = \sup_{\tau \in \mathcal{T}(t+1,\ldots,T)} \mathbf{E}\,\{f_\tau(X_\tau)\,|\,X_t = x\} \tag{3}$$

be the so-called continuation value describing the value of the option at time t given $X_t = x$ and subject to the constraint of holding the option at time t rather than exercising it. Here $\mathcal{T}(t+1,\ldots,T)$ is the class of all $\{t+1,\ldots,T\}$-valued stopping times. Set $q_T(x) = 0$. It can be shown that

$$\tau^* = \inf\,\{s \ge 0 : q_s(X_s) \le f_s(X_s)\} \tag{4}$$

satisfies (2), i.e., τ^* is an optimal stopping time (cf., e.g., Chow et al., 1971 or Shiryayev, 1978). Therefore it suffices to compute the continuation values (3) in order to solve the optimal stopping problem (1).

One way to compute the continuation values is to use a regression representation like

$$q_t(x) = \mathbf{E}\{\max\,\{f_{t+1}(X_{t+1}), q_{t+1}(X_{t+1})\}\,|\,X_t = x\} \quad (t = 0,1,\ldots,T-1) \tag{5}$$

(cf., Tsitsiklis and Van Roy, 1999, further regression representations can be found in Longstaff and Schwartz, 2001; Egloff, 2005). Typically, the underlying distributions are rather complicated, therefore it is not possible to compute the conditional expectation in (5) directly.

The basic idea of regression-based Monte Carlo methods for pricing American options is to apply recursively regression estimates to artificially created samples of

$$(X_t, \max\{f_{t+1}(X_{t+1}), \hat{q}_{t+1}(X_{t+1})\})$$

(so-called Monte Carlo samples) to construct estimates $\hat{q}_t$ of q_t. In connection with linear regression this was proposed in Tsitsiklis and Van Roy (1999), and, based on a different regression estimation than (5), in Longstaff and Schwartz (2001). Nonparametric least squares regression estimates have been applied and investigated in this context in Egloff (2005), Egloff et al. (2007) and Kohler et al. (2010), smoothing spline regression estimates have been analyzed in Kohler (2008b), recursive kernel regression estimates have been considered in Barty et al. (2008), and local polynomial kernel estimates have been studied in Belomestny (2011).

In this paper we develop a general theory which covers the estimates of most of the above papers as well as additional ones (e.g., orthogonal series estimates). Our main theoretical results provide a unifying tool which enables to derive rate of convergence results for many estimates at once. Furthermore we illustrate the estimates by applying them to an option based on the average of three correlated stocks. The simulations show that the nonparametric estimates studied in this paper produce better results than the existing parametric ones.

1.1. Notation

The sets of natural numbers, natural numbers including zero, integers, real numbers and non-negative real numbers are denoted by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{R}_+$, respectively. The least integer greater than or equal to a real number x will be denoted by $\lceil x \rceil$. $\log(x)$ denotes the natural logarithm of $x > 0$. For a function $f : \mathbb{R}^d \to \mathbb{R}$ the partial derivative with respect to the j-th component will be denoted by

$$\frac{\partial f}{\partial x_j}.$$

We say that $a_n = O_{\mathbf{P}}(b_n)$ if $\limsup_{n\to\infty} \mathbf{P}(a_n > c \cdot b_n) = 0$ for some finite constant c.

1.2. Outline of the paper

The precise definition of the estimates and the main theoretical results concerning the rate of convergence of the estimates are given in Sections 2 and 3, respectively. The application of the estimates to simulated data will be described in Section 4, and the proofs will be given in Section 5.