Monte Carlo methods and stochastic control problems

November 9th, 2005

Ricky Rambharat

Deterministic stochastic optimal control

- Developed in Rogers (2005) preprint.
- Based on a dual result of American option pricing:

$$\begin{aligned} Y_0^* &= \sup_{\tau \in \mathcal{T}} E_0[Z_\tau] \\ &= \inf_{M \in \mathcal{M}_0} E_0 \left[\sup_{0 \le t \le T} (Z_t - M_t) \right], \end{aligned}$$

  where, $\mathcal{M}_0$ is the space of uniformly integrable martingales started at zero (see Rogers, 2002).
- Smallest supermartingale majorant to the payoff function (see Myneni, 1992).
- Results in an upper bound on the option price.

The Optimization Problem

Let X be a Markov process taking values in $\mathcal{X}$. The goal is to control X over choice of controls $a \in \mathcal{A}$, where $\mathcal{A}$ is the class of adapted processes with values in some set $\mathcal{U}$ of admissable controls.

The controlled transitions have density $\phi(x, y; a)$ w.r.t. some reference Markovian transition P^*.

The valuation function of the problem starting from state x at time j is,

$$V_j(x) \quad = \quad \text{sup}_{a \in \mathcal{A}} E\left[\sum_{r=j}^{T-1} f_r(X_r, a_r) + F(X_T) | X_j = x\right]$$

Change of measure

Define

$$\Lambda_t(a) \quad = \quad \prod_{r=0}^{t-1} \phi(X_r, X_{r+1}; a_r),$$

Recast the optimization problem as

$$V_0(X_0) \quad = \quad \text{sup}_{a \in \mathcal{A}} E^* \left[\sum_{j=0}^{T-1} \Lambda_j(a) f_j(X_j, a_j) + \Lambda_T(a) F(X_T) \right]$$

Result for stochastic control problem

First main result (Theorem 1 of Rogers, 2005)

$V_0(X_0) =$

$\min_{(h_j)} E^* \left[\sup_a \left\{ \sum_{j=0}^{T-1} \Lambda_j(a) \left\{ f_j(X_j, a_j) - \eta_{j+1} + E_j^*(\eta_{j+1}) \right\} + \Lambda_T(a) F(X_T) \right\} \right],$

where,

$\eta_{j+1} = h_{j+1}(X_{j+1}) \phi(X_j, X_{j+1}; a_j)$

- Subtracted martingale difference $\eta_{j+1} - E_j^*(\eta_{j+1})$.
- Pathwise maximization over the controls.
- Minimize over the choice of the martingale difference sequence.

Note: Rogers (2005) also gives a multiplicative version of this result – see Theorem 2 of his preprint.

Another characterization

This is the value function in a result stated in Theorem 3:

$X_{j+1} = \xi(j, X_j, a_j, \epsilon_{j+1}), j = 0, \ldots, T-1.$

Define,

$Ph_{j+1}(x, a) = E\left[h_{j+1}(\xi(j, x, a, \epsilon_{j+1}))\right]$

$V_0(X_0) =$
$\min_{(h_j)} E\left[\sup_a \left\{\sum_{j=0}^{T-1} \left(f_j(X_j, a_j) - h_{j+1}(X_{j+1}) + Ph_{j+1}(X_j, a_j)\right) + F(X_T)\right\}\right]$

Note: Rogers (2005) establishes a recursive version to the above result in order to execute efficient numerical computations.

Sketch of Algorithm

Suppose that $B = \sup_{a,x,x'} \phi(x, x'; a) < \infty$.

Let $\{V_j^{(0)}\}_{j=0}^T$ be a sequence of function from $\mathcal{X}$ to $\mathcal{X}$, with $V_T^{(0)} = F$.

Define recursively the functions $\{V_k^{(n)}\}_{k=0}^T$ for $n = 1, 2, \ldots$ by

$$V_k^{(n+1)}(x) =$$
$$E^* \left[\sup_a \left\{ \sum_{j=k}^{T-1} \Lambda_{k,j}(a) \left\{ f_j(X_j, a_j) - V_{j+1}^{(n)}(X_{j+1}) \phi(X_j, X_{j+1}; a_j) + P V_{j+1}^{(n)}(X_j, a_j) \right\} \right.\right.$$
$$\left.\left. + \Lambda_{k,T}(a) F(X_T) \right\} \middle| X_k = x \right],$$

for $x \in \mathcal{X}$ and $k = 0, \ldots, T$, where,

$\Lambda_{k,j}(a) = \prod_{r=k}^{j-1} \phi(X_r, X_{r+1}; a_r)$, and

$P\psi(x, a) = E^*[\phi(x, X_1; a)\psi(X_1) | X_0 = x]$.

Let $\Delta_k^{(n)} = \sup_x |V_k^{(n)}(x) - V_k^{(n-1)}(x)|$,

$k = 0, 1, \ldots, T$, $n \geq 1$, we get a bound

$\Delta_k^{(n)} \leq (1 + B) \sum_{r=k+1}^T \Delta_r^{(n-1)}$.

Main Steps of the Algorithm

- Propose an approximation (h_j) to the value.
- Evaluate $E[\text{sup}_a \ldots]$
- Improve on the approximation of (h_j)

Bellman recursions

$$\begin{aligned} V_{n-1}(x) &= \text{sup}_a E^*[f(x,a) + \phi(x, X_1; a) V_n(X_1) | X_{n-1} = x], (1 \leq n \leq t) \\ V_T(x) &= F(x) \end{aligned}$$

Issues

- How to place the points of $\mathcal{X} \in R^N$ at the start of the dynamic programming algorithm.
- Would hope to place points in regions where the optimally-controlled process is most likely to go – but we do not know where this will be.

Rogers' proposal

- Set $k = 0$.
- Set reference measure $P^{(k)}$ = $(P^* \text{for} k = 0)$.
- Propose approximations $h_n^{(k)}$ to $V_n^{(k)}$.
- Simulate N paths and optimize pathwise – at each time n we obtain an approximation $\hat{V}_n^{(k+1)}(X_n^{(j)})$ to to $V_n^{(k+1)}$ at each of the points $X_n^{(1)}, \dots, X_n^{(N)}$ visited by the simulated paths.
- Regress approximate value onto basis – find some linear combination of basis functions that matches $\hat{V}_n^{(k+1)}(X_n^{(j)})$ at the points $X_n^{(j)}$.
- Propose a $P^{(k+1)}$. Transitions from position x at time n will be determined by selecting a point $X_n^{(j)}$ from $\{X_n^1, ..X_n^N\}$ at random, points closer to x being selected with higher probability, and then jumping from the chosen point according to the transition law for the action a, which was optimal for the j-th path.
- Go to simulation step.

An example from the preprint

Consider a controlled Markov process on the unit circle $[0, 2\pi]$ whose dynamics are given by

$$X_{t+1} = X_t + \epsilon_{t+1} + a_t \quad (1)$$

- ϵ_t have density proportional to cos(x).
- The control a lies in $[0, 2\pi]$.
- Objective: $\sum_{t=0}^{T} \beta^t [cos(X_t) + cos(a_t)]$