Interest Rate Models

Ivan Matic

Department of Mathematics
Baruch College, CUNY
New York, NY 10010

Ivan.Matic@baruch.cuny.edu

Spring 2016

Outline

1. Bermudan swaptions
2. Valuation of American options
3. Hermite polynomials
4. Conditional expected values and their estimation

Early termination

- There are fixed income securities that offer a counterparty to exit during a specified period of time.
- These exit options can be
  (i) *Endogenous* – driven by interest rates only;
  (ii) *Exogenous* – driven by other risk factors such as credit or prepayment.
- The valuations of instruments have to account for the possibilities for early termination.
- We will focus only on those instruments whose treatment needs only the understanding of interest rate models.

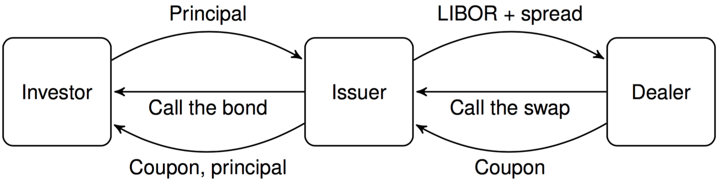
Early termination

- A *callable bond* allows the issuer to repay the principal and cancel all future coupon payments.
- A *putable bond* allows the bond holder to request the principal earlier and forfeit the future coupons.
- Typically these options are of American style, and they allow for multiple exercise dates.
- LIBOR market model is powerful enough to handle complex interest rate options.
- We will focus on Bermudan swaptions and their valuations using LMM style Monte Carlo simulations.

Bermudan swaptions

- A *callable swap* is a fixed for LIBOR swap that gives a periodic right to the **fixed payer** to cancel the remaining cash flows.
- *Call the swap* is the terminology that describes this action.
- The cancellation dates are pre-determined and cannot happen during the initial *lockout period*.
- Terminology example: *10 no call 2 swap*. This is a 10 year swap that can be called very 6 months, two business days preceding a fixed rate roll date with a 2 year lockout period.
- We will take a look at an example of how callable swaps are used.
- They are usually created on the back of bond issuance.

Bermudan swaptions



- The issuer sells a callable bond to the investor. The issuer has the right to call the bond.
- The issuer has access to a favorable LIBOR spread.
- The issuer finds a dealer willing to be a fixed payer on a callable swap. Callable swap makes the fixed coupons higher.
- If the dealer calls the swap, the issuer will call the bond.

Bermudan swaptions

- A *putable swap* is a fixed for LIBOR swap that gives a periodic right to the **fixed receiver** to cancel the remaining cash flows.
- The fixed rate for a putable swap is lower than the market rate for a regular swap (also known as the swap rate).
- Multiple exercise swap options are called Bermudan swaptions.
- Example: *A semi-annual 1 into 5 Bermudan receiver* is a semi-annual 5.50% 1 year Bermudan receiver swaption into a 5 year swap. The holder has the right to receive a 5.50% on a 5 year swap starting in 1 year from now; or a 4.5 year swap starting in 1.5 years, ..., or a 6 month swap starting in 5.5 years.
- The total maturity of transactions is 6 years.
- Bermudan swaptions trade over the counter.

Bermudan swaptions

- A callable swap is a basket of
  - Vanilla swap, and
  - Bermudan swaption.
- Example: 10 no call 1 callable swap paying 3%. The fixed payer would obtain the same payoff by entering into
  - Short position (fixed payer) on a 10 year swap paying 3% and
  - Long position in an 1 into 9 Bermudan receiver swaption struck at 3%.
- Calling the callable swap is equivalent to exercising the Bermudan receiver.

Principle of optimality and backward induction

- A big challenge in Monte Carlo based valuations come from the incompatibility of forward moving path generation and the backward moving optimal exercise decision process.
- We will discuss a version of the Longstaff-Schwartz methodology.
- Consider an American option with valuation time $T_0 = 0$ that can be exercised on a set of dates $T_1 < T_2 < \cdots < T_n$.
- If the options is exercised at time T_j, the option writer delivers the underlying instrument.
- Let $g_j(T_j)$ be the payoff at the exercise time T_j.

Principle of optimality and backward induction

- The act of exercising an American options is modeled by a stopping time τ.
- Assume that we are working under the terminal forward measure $\mathbb{Q}$.
- Today's value of the option is

$$V_0 = \max_{\tau} \mathbb{E}^{\mathbb{Q}} \left[P(0, \tau) g(\tau) \right],$$

  where the maximum is taken over all stopping times τ whose codomain is the set $\{T_1, T_2, \ldots, T_n\}$.
- We will use $\mathcal{T}\{T_1, T_2, \ldots, T_n\}$ for this set of stopping times.
- We want to determine the optimal stopping time τ^* (or the optimal exercise policy) for which

$$V_0 = \mathbb{E}^{\mathbb{Q}} \left[P(0, \tau^*) g(\tau^*) \right].$$

Principle of optimality and backward induction

- The value of the option also satisfies:

$$V_j = \max_{\tau = \mathcal{T}\left\{T_{j+1}, T_{j+2}, \ldots, T_n\right\}} \mathbb{E}_j \left[P(T_j, \tau) g(\tau)\right],$$

  here E_j stands for $\mathbb{E}^{\mathbb{Q}}\left[\cdot | \mathcal{F}_{T_j}\right]$.
- Backward induction starts from V_n.
- The valuation starts from the terminal condition

$$V_n = \mathbb{E}_n \left[g(T_n)\right].$$

Principle of optimality and backward induction

- At each point in time and each future scenario, we compare the following two quantities:
  - The remaining value of the option (continuation value): The value of the option that would not be exercised at this point;
  - The value of immediate exercise (intrinsic value).
- The higher of the two is the used in moving backward.
- The earliest exercise date at which the intrinsic value exceeds the continuation value is the optimal exercise date.
- The analysis is repeated for all future scenarios.

Principle of optimality and backward induction

- This algorithm is easy to implement for Hull-White model.
- In LMM model the valuations must be done using Monte-Carlo simulations.
- Naive Monte-Carlo is inefficient. Along every path and every future time when exercising is in question, we have to generate many new Monte-Carlo paths.
- It is difficult to estimate the conditional expected values given a future scenario.
- Thus the complexity becomes exponential.

Snell envelope

- We will now assume that the model is one-factor model driven by a single Brownian motion W.
- Multiple factors do not require new ideas, just more cumbersome notation.
- Let us introduce $X_j = \frac{W(T_j)}{\sqrt{T_j}}$.
- Let $X_{1:j}$ denote the vector $(X_1, X_2, \ldots, X_j)$.
- The payoff function $g(T_j)$ at time T_j will be denoted by g_j.
- g_j is a function of path, and when we want to emphasize this we will write $g_j(X_{1:j})$.

Snell envelope

- A Snell envelope is a process $S_j = S_j(X_{1:j})$ that represents the value of the option along the path. Its recursive definition starts with
$$S_n(X_{1:n}) = g_n(X_{1:n}).$$
- The terminal value of the option is simply its payoff.
- We denote by $C_j(X_{1:j})$ the continuation value of the option.
$$C_j(X_{1:j}) = \begin{cases} P_{j,j+1}\mathbb{E}_j[S_{j+1}], & \text{if } j \in \{0, 1, \ldots, n-1\}, \\ 0, & \text{otherwise.} \end{cases}$$
$P_{j,j+1}$ is the discount factor between T_j and T_{j+1}.
- We now move backwards for $j = n-1,\ n-2,\ \ldots,\ 1$.
$$S_j(X_{1:j}) = \max\left\{g_j(X_{1:j}), C_j(X_{1:j})\right\}.$$

Snell envelope

- S_0 is the value of the Snell envelope on its value date and is defined as
$$S_0 = P_{0,1}\mathbb{E}_0\left[S_1\right].$$
- The snell envelope satisfies $S_j \geq g_j$. It is the smallest supermartingale with this property.
- Recall that X_s is a supermartingale if $X(s) \geq \mathbb{E}\left[S(t)|\mathcal{F}_s\right]$ for $s \leq t$.
- We need to calculate the continuation values and payoff values.
- In order to make these calculations we have to be able to calculate the conditional expectations.
- We will do this using Hermite polynomials and Monte Carlo simulations.

Calculating conditional expectations

- Assume that f is a function that depends on the path $X_{1:j}$. We want to calculation $\mathbb{E}_{j-1}\left[f(X_{1:j})\right]$.
- Let us denote $\xi_j = \frac{1}{\sqrt{T_j - T_{j-1}}}\left(W(T_{j-1}) - W(T_j)\right)$. Then

$$
\begin{aligned}
X_j &= \frac{1}{\sqrt{T_j}} W(T_j) \\
&= \frac{1}{\sqrt{T_j}}\left(W(T_{j-1}) + \left(W(T_j) - W(T_{j-1})\right)\right) \\
&= \sqrt{\frac{T_{j-1}}{T_j}} X_{j-1} + \sqrt{1 - \frac{T_{j-1}}{T_j}} \cdot \xi_j \\
&= \sqrt{\alpha_j} X_{j-1} + \sqrt{1-\alpha_j}\,\xi_j, \text{ where } \alpha_j = \frac{T_{j-1}}{T_j}.
\end{aligned}
$$

Calculating conditional expectations

- Therefore

$$\mathbb{E}_{j-1}[f(X_{1:j})] \;=\; \mathbb{E}_{j-1}\left[f\left(X_{1:j-1}, \sqrt{\alpha_j}X_{j-1} + \sqrt{1-\alpha_j}\xi_j\right)\right].$$

- $X_{1:j-1}$ is deterministic and $\xi_j \sim N(0,1)$.
- The key is to approximate f by a polynomial.
- Let μ be the distribution function for the standard normal.
- $\langle f, g\rangle = \int_{-\infty}^{+\infty} f(x)g(x)\,d\mu(x)$ is a dot product, and the space $L^2(\mu)$ is a Hilbert space.
- From the basis $\{1, X, X^2, \dots\}$ we can create an orthonormal one using Gramm-Schmidt orthogonalization procedure.
- The obtained orthonormal basis consists of *normalized Hermite polynomials* h_0, h_1, ...

Hermite polynomials

- We will first guess an orthogonal basis for $L^2(\mu)$.
- Let us first define

$$\Psi(t, x) = e^{-t^2+2tx}.$$

- Keeping x as a constant, the function $\varphi(t) = \Psi(t, x)$ has a Maclaurin expansion

$$\Psi(t, x) = \sum_{n=0}^{\infty} H_n(x) \cdot \frac{t^n}{n!}.$$

- The coefficients $H_n(x)$ are polynomials in x.
- They are called Hermite polynomials and we will prove that they satisfy $\int_{-\infty}^{+\infty} H_n(x)H_m(x)e^{-x^2}\,dx = 0$ whenever $n \neq m$.

Hermite polynomials

- Notice that the definition of $H_n(x)$ is equivalent to

$$H_n(x) \;=\; \frac{d^n}{dt^n}\Psi(t,x)\bigg|_{t=0}.$$

- Observe that $\Psi(t,x) = e^{x^2} \cdot e^{-(t-x)^2}$. Hence

$$\begin{aligned} H_n(x) &= \frac{d^n}{dt^n}\left(e^{x^2}e^{-(t-x)^2}\right)\bigg|_{t=0} = e^{x^2}\,\frac{d^n}{dt^n}\left(e^{-(t-x)^2}\right)\bigg|_{t=0} \\ &= (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}. \end{aligned}$$

- The first few polynomials are: $H_0(x) = 1$, $H_1(x) = 2x$, $H_2(x) = 4x^2 - 2$, $H_3(x) = 8x^3 - 12x$.

Hermite polynomials

- Using the Leibnitz formula we obtain

$$
\begin{aligned}
H_n(x) &= (-1)^n e^{x^2} \frac{d^{n-1}}{dx^{n-1}} \left(-2xe^{-x^2}\right) \\
&= (-1)^{n-1} \cdot 2 \cdot e^{x^2} \frac{d^{n-1}}{dx^{n-1}} \left(xe^{-x^2}\right) \\
&= (-1)^{n-1} \cdot 2x \cdot e^{x^2} \frac{d^{n-1}}{dx^{n-1}} \left(e^{-x^2}\right) \\
&\quad +(-1)^{n-1} \cdot 2(n-1) \cdot e^{x^2} \frac{d^{n-2}}{dx^{n-2}} \left(e^{-x^2}\right) \\
&= 2xH_{n-1}(x) - 2(n-1)X_{n-2}(x).
\end{aligned}
$$

- The previous equation can also be written as

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x).$$

Hermite polynomials

- A similar calculation gives us

$$H_n'(x) = 2nH_{n-1}(x).$$

- Homework: Use induction to prove that for $n \neq m$

$$\int_{-\infty}^{+\infty} H_n(x)H_m(x)e^{-x^2}\,dx = 0.$$

- We now define normalized Hermite polynomials:

$$h_n(x) = \frac{1}{2^{\frac{n}{2}}\sqrt{n!}}H_n\left(\frac{x}{\sqrt{2}}\right).$$

Hermite polynomials

- Normalized Hermite polynomials satisfy:

$$\begin{aligned} \sqrt{n+1}h_{n+1}(x) &= xh_n(x) - \sqrt{n}h_{n-1}(x) \\ h_n'(x) &= \sqrt{n}h_{n-1}(x). \end{aligned}$$

- Homework: Prove that

$$\int_{-\infty}^{+\infty} h_m(x)h_n(x) \cdot \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = 1_{m=n}.$$

- Every function $f \in L^2(\mu)$ can be expressed as

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} a_n h_n(x), \quad \text{where} \\ a_n &= \int_{-\infty}^{+\infty} f(x)h_n(x)\, d\mu(x). \end{aligned}$$

Addition theorem

- Notice that

$$\begin{aligned}\Psi\left(t, \sqrt{\alpha}x + \sqrt{1-\alpha}y\right) &= e^{-t^2+2t\left(\sqrt{\alpha}x+\sqrt{1-\alpha}y\right)} \\ &= e^{-\alpha t^2+\sqrt{\alpha}x} \cdot e^{-(1-\alpha)t^2+\sqrt{1-\alpha}y} \\ &= \Psi\left(\sqrt{\alpha}t, x\right) \cdot \Psi\left(\sqrt{1-\alpha}t, y\right).\end{aligned}$$

- Comparing the Maclauring series and coefficient next to t^n we conclude that

$$H_n\left(\sqrt{\alpha}x + \sqrt{1-\alpha}y\right) = \sum_{k=0}^{n} \binom{n}{k} \cdot \alpha^{\frac{k}{2}} \cdot (1-\alpha)^{\frac{n-k}{2}} \cdot H_k(x) \cdot H_{n-k}(y).$$

Addition theorem

- Using the relationship between h_n and H_n we obtain

$$h_n\left(\sqrt{\alpha}x+\sqrt{1-\alpha}y\right)=\sum_{k=0}^{n}\sqrt{\binom{n}{k}}\cdot\alpha^{\frac{k}{2}}\cdot(1-\alpha)^{\frac{n-k}{2}}\cdot h_k(x)\cdot h_{n-k}(y).$$

- Homework: Prove that

$$\int_{-\infty}^{+\infty}h_n\left(\sqrt{\alpha}x+\sqrt{1-\alpha}y\right)\,d\mu(y)=\alpha^{\frac{n}{2}}h_n(x).$$

- Recall that $X_{1:j}=\sqrt{\alpha_j}X_{1:j-1}+\sqrt{1-\alpha_j}\xi_j$.
- We now want to calculate $\mathbb{E}_i\left[f\left(X_{1:j}\right)\right]$ for $i<j$.

Conditional expected value

- Assume first that $i = j - 1$.
- The function $f(X_{1:j})$ can be represented as the Fourier series

$$f(X_{1:j}) = \sum_{k=0}^{\infty} c_k(X_{1:j-1}) h_k(X_j),$$

where

$$c_k(X_{1:j-1}) = \int_{-\infty}^{+\infty} f(X_{1:j}) h_k(X_j)\, d\mu(X_j).$$

- In the previous formula, X_j is a real number and a variable of integration. The function f depends on $X_{1:j-1}$ and X_j. Therefore, the result of integration $c_k(X_{1:j-1})$ depends on $X_{1:j-1}$.

Conditional expected value

- We now have

$$\begin{aligned}\mathbb{E}_{j-1}\left[f(X_{1:j})\right] &= \sum_{k=0}^{\infty} c_k(X_{1:j-1})\mathbb{E}_{j-1}\left[h_k(\sqrt{\alpha_j}X_{j-1}+\sqrt{1-\alpha_j}\xi_j)\right] \\ &= \sum_{k=0}^{\infty} c_k(X_{1:j-1})\alpha_j^{\frac{k}{2}} h_k\left(X_{j-1}\right)\end{aligned}$$

- We can easily extend our method from $i = j-1$ to the general case $i < j$.

$$\mathbb{E}_i\left[f(X_{1:j})\right] = \mathbb{E}_i\left[\mathbb{E}_{i+1}\left[\cdots\mathbb{E}_{j-1}\left[f(X_{1:j})\right]\cdots\right]\right]$$

- The above formulas are exact. However, in practice we cannot usually find the sum of the infinite Fourier series and we cannot evaluate the integral that defines c_k.

Monte Carlo estimation

- Assume that we have generated N Monte Carlo paths $X^i_{1:n}$, $i = 1, 2, \ldots, N$.
- The first step is *truncation* of the series. For some finite κ we will consider the approximation:

$$f(X_{1:j}) \quad \approx \quad \sum_{k=0}^{\kappa} c_k(X_{1:j-1}) h_k(X_j),$$

- Therefore

$$\mathbb{E}_{j-1}\left[f(X_{1:j})\right] \quad \approx \quad \sum_{k=0}^{\kappa} \alpha_j^{\frac{k}{2}} c_k(X_{1:j-1}) h_k\left(X_{j-1}\right).$$

- Even low values for κ (such as $\kappa = 5$) lead to surprisingly good numerical results.

Monte Carlo estimation

- We will now calculate approximate values for the Fourier coefficients.
- The standard Monte Carlo estimation would approximate the integral in the following way

$$c_k(X_{1:j-1}) \quad = \quad \int_{-\infty}^{+\infty} f(X_{1:j}) h_k(X_j)\, d\mu(X_j) \approx \frac{1}{N} \sum_{i=1}^{N} f(X_{1:j}^i) h_k(X_j^i).$$

- However, the following trick leads to a much faster convergence:
- We will look for the vector $\hat{c}$ of coefficients $\hat{c}_k$ that minimize

$$\zeta(c_0, c_1, \ldots, c_\kappa) \quad = \quad \sum_{i=1}^{N} \left(f(X_{1:j}^i) - \sum_{k=0}^{\kappa} c_k h_k(X_j^i) \right)^2 .$$

Monte Carlo estimation

- The minimizing sequence $\hat{c}$ of

$$\zeta(c_0, c_1, \ldots, c_\kappa) \;=\; \sum_{i=1}^{N} \left(f(X_{1:j}^i) - \sum_{k=0}^{\kappa} c_k h_k(X_j^i) \right)^2 .$$

  satisfies $\nabla\zeta(\hat{c}) = 0$.
- The equation $\nabla\zeta(c) = 0$ is equivalent to $G\overrightarrow{c} = \overrightarrow{b}$, where G is the matrix and $\overrightarrow{b}$ the vector that satisfy:

$$G_{\ell k} = \frac{1}{N}\sum_{i=1}^{N} h_k(X_j^i)h_\ell(X_j^i), \qquad b_\ell = \frac{1}{N}\sum_{i=1}^{N} f(X_{1:j}^i)h_\ell(X_j^i).$$

Monte Carlo estimation

- As $N \to \infty$ the matrix G converges to the identity matrix.
- Therefore we obtain the following estimate for the expectation

$$\hat{E}\left[f(X_{1:j})\right] \approx \sum_{k=0}^{\kappa} \alpha_j^{\frac{k}{2}} \hat{c}_k h_k(X_{j-1}),$$

where

$$\hat{c} = G^{-1} \overrightarrow{b}.$$

Optimal exercise

- Assume now that we have N Monte-Carlo paths $X^i_{1:n}$ for $i \in \{1, 2, \ldots, N\}$.
- *Step 1.* We will first assume that the immediate payoff $g_j(X^i_{1:j})$ can be calculated without using Monte-Carlo methods.
- For example, the underlying asset for Bermudan swaption is a swap rate that can be priced using multi-curve built with data available at time T_j.
- Since all $g_j(X^i_{1:j})$ are known we can calculate Snell envelope.

$$C_n(X^i_{1:n}) = 0 \qquad \text{and} \qquad S_n(X^i_{1:n}) = g_n(X^i_{1:n}).$$

Optimal exercise

- Suppose that $C_j(X^i_{1:j})$ and $S_j(X^i_{1:j})$ are calculated for some j and all i. We can evaluate

$$
\begin{aligned}
C_{j-1}(X^i_{1:j-1}) &= P_{j-1,j}\hat{\mathbb{E}}_{j-1}\left[S_j(X_{1:j})\right] \\
S_{j-1}(X^i_{1:j-1}) &= \max\left\{C_{j-1}(X^i_{1:j-1}), g_{j-1}(X^i_{1:j-1})\right\}.
\end{aligned}
$$

- Assume now that the underlying instrument is a contingent claim. Its price is conditional expected value.
- We pre-compute the value of cash flow $p_j(X^i_{1:n})$ at time T_j for each Monte Carlo path and each j.
- Then we have

$$
g_j(X^i_{1:n}) = \mathbb{E}_j\left[p_j(X^i_{1:n})\right] = \mathbb{E}_j\left[\cdots\mathbb{E}_{n-1}\left[p_j(X^i_{1:n})\right]\cdots\right].
$$

Optimal exercise

- We can now summarize the procedure: Start with $j = n$ and calculate

$$\begin{aligned} g_\ell(X^i_{1:n}) &= p_\ell\left[X^i_{1:n}\right] \qquad \text{for all } \ell \in \{1, 2, \ldots, n\}; \\ C_n(X^i_{1:n}) &= 0; \\ S_n(X^i_{1:n}) &= g_n(X^i_{1:n}). \end{aligned}$$

- Once the values C_j and S_j are found, we calculate

$$\begin{aligned} p_\ell(X^i_{1:j-1}) &= \hat{\mathbb{E}}_{j-1}\left[p_\ell(X^i_{1:j})\right] \qquad \text{for all } \ell \in \{1, 2, \ldots, n\}; \\ C_{j-1}(X^i_{1:j-1}) &= P_{j-1,j}\hat{\mathbb{E}}_{j-1}\left[S_j(X^i_{1:j})\right]; \\ S_{j-1}(X^i_{1:j-1}) &= \max\left\{C_{j-1}(X^i_{1:j-1}), g_{j-1}(X^i_{1:j-1})\right\}. \end{aligned}$$

Optimal exercise

- It remains to calculate the optimal exercise time.
- Consider a particular Monte Carlo path $X^i = \left(X_0^i, X_1^i, \ldots, X_n^i\right)$.
- For each path we initialize the the value of the stopping time $\tau_*(X^i) = T_n$.
- As we go backward in time we update the value $\tau_*(X^i)$ to T_j if $g_j(X_{1:j}^i) \geq C_j(X_{1:j}^i)$.
- Formally, $\tau_*(X_{1:n}^i) = \min\left\{j : S_j(X_{1:j}^i) = g_j(X_{1:j}^i)\right\}$.
- The value of the option is the average over all Monte Carlo paths:

$$V_0 = \frac{1}{N}\sum_{i=1}^{N} P_{0,j_*} g_{j_*}(X_{1:j_*}^i),$$

where j_* is the index j for which $\tau_*(X_{1:n}^i) = T_j$.

References

- Andersen, L.; Peterbarg, V.: *Interest rate modeling*, Vol. 3, Atlantic Financial Press (2010)
- Glasserman, P.: *Monte Carlo methods in financial engineering*, Springer (2004).
- Lebedev, N.N.: *Special functions and their applications*, Dover (1972).
- Lesniewski, A.: http://lesniewski.us/presentations.html
- Longstaff, F., Schwartz, E.: *Valuing American options by simulation: a simple least squares approach*, Rev. Financial Stud., 14, 113 – 147 (2001)