Interest Rate Modelling and Derivative Pricing

Sebastian Schlenkrich

HU Berlin, Department of Mathematics

Summer term, 2023

Part V

Bermudan Swaption Pricing

Outline

Outline

Let's have another look at the cancellation option

Interbank swap deal example

Pays 3% on 100mm EUR

Start date: Oct 30, 2020

End date: Oct 30, 2040

(annually, 30/360 day count, modified following, Target calendar)

Pays 6-months Euribor floating rate on 100mm EUR

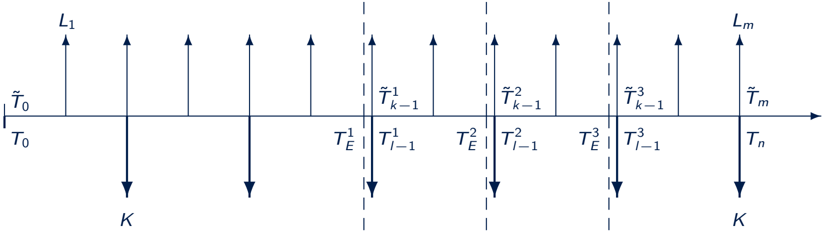
Start date: Oct 30, 2020

End date: Oct 30, 2040

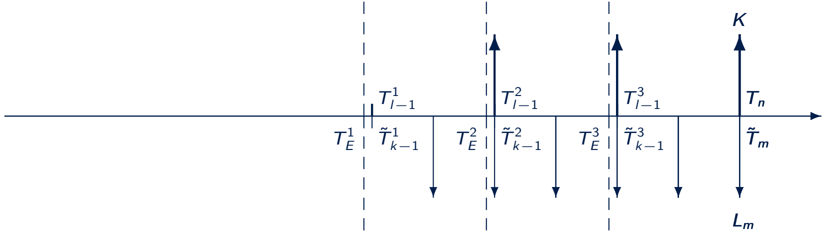
(semi-annually, act/360 day count, modified following, Target calendar)

Bank A may decide to early terminate deal in 10, 11, 12,..years.

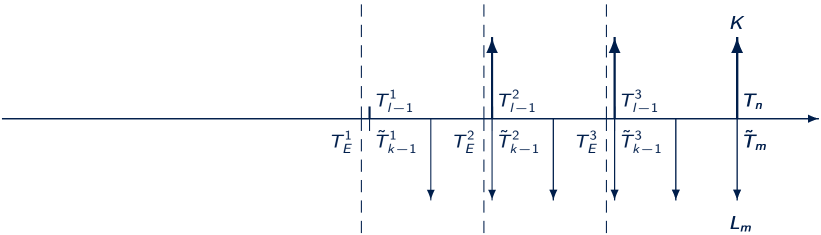
What does such a *Bermudan call right* mean?



[Bermudan cancellable swap] = [full swap] + [Bermudan option on opposite swap]



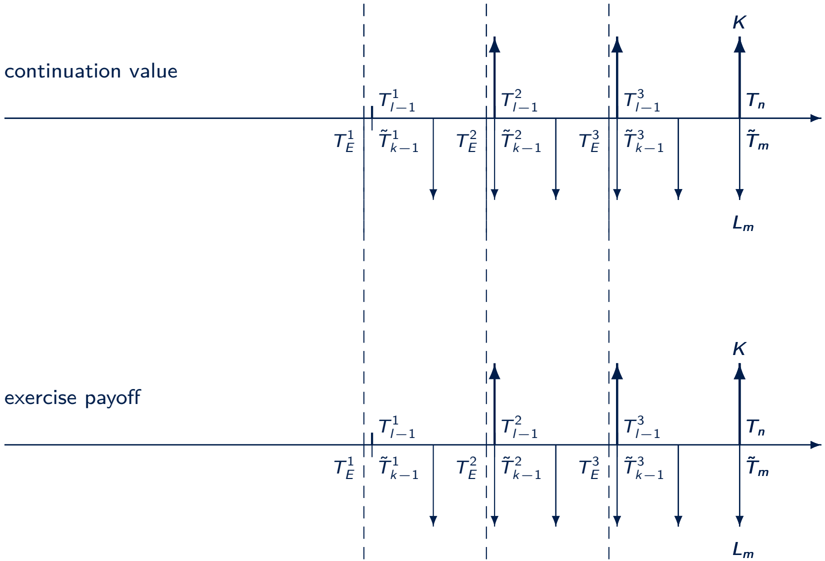
What is a Bermudan swaption?



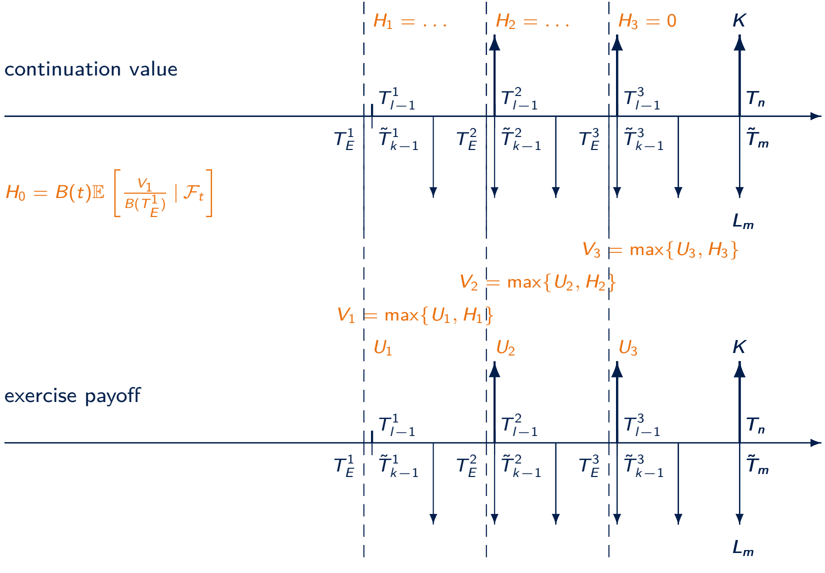
Bermudan swaption

A Bermudan swaption is an option to enter into a Vanilla swap with fixed rate K and final maturity T_n at various exercise dates $T^1_E, T^2_E, \ldots, T^{\bar{k}}_E$. If there is only one exercise date (i.e. $\bar{k} = 1$) then the Bermudan swaption equals a European swaption.

A Bermudan swaption can be priced via *backward induction*



A Bermudan swaption can be priced via *backward induction* - let's add some notation



First we specify the future payoff cash flows

- Choose a numeraire $B(t)$ and corresponding cond. expectations $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \mid \mathcal{F}_t]$.
- Underlying payoff U_k if option is exercised

$$
\begin{aligned}
U_k &= B(T_E^k) \sum_{T_i \geq T_E^k} \mathbb{E}_{T_E^k}\left[\frac{X_i(T_i)}{B(T_i)}\right] \\
&= B(T_E^k) \underbrace{\left[\sum_{T_i \geq T_E^k} K\tau_i P(T_E^k, T_i) - \sum_{\tilde{T}_j \geq T_E^k} L^\delta(T_E^k, \tilde{T}_{j-1}, \tilde{T}_{j-1} + \delta)\tilde{\tau}_j P(T_E^k, \tilde{T}_j)\right]}_{\text{future fixed leg minus future float leg}} \\
&= B(T_E^k) \left[\sum_{T_i \geq T_E^k} K\tau_i P(T_E^k, T_i) - \left[P(T_E^k, \tilde{T}_{j_k}) - P(T_E^k, \tilde{T}_m)\right]\right. \\
&\qquad \left. - \sum_{\tilde{T}_j \geq T_E^k} P(T_E^k, \tilde{T}_{j-1}) \left[D(\tilde{T}_{j-1}, \tilde{T}_j) - 1\right]\right].
\end{aligned}
$$

Then we specify the continuation value and optimal exercise (1/2)

- Continuation value $H_k(t)$ ($T_E^k \leq t \leq T_E^{k+1}$) represents the time-$t$ value of the remaining option if not exercised.
- Option becomes worthless if not exercised at last exercise date $T_E^{\bar{k}}$. Thus last continuation value $H_{\bar{k}}(T_E^{\bar{k}}) = 0$.
- Recall that Bermudan option gives the right but not the obligation to enter into underlying at exercise.
- Rational agent will choose the maximum of payoff and continuation at exercise, i.e.
$$V_k = \max\left\{U_k, H_k(T_E^k)\right\}.$$

Then we specify the continuation value and optimal exercise (2/2)

$$V_k = \max\left\{U_k, H_k(T_E^k)\right\}.$$

- V_k represents the Bermudan option value at exercise T_E^k. Thus we also must have for the continuation value

$$H_{k-1}(T_E^k) = V_k.$$

- Derivative pricing formula yields

$$\begin{aligned} H_{k-1}(T_E^{k-1}) &= B(T_E^{k-1}) \cdot \mathbb{E}_{T_E^{k-1}}\left[\frac{H_{k-1}(T_E^k)}{B(T_E^k)}\right] \\ &= B(T_E^{k-1}) \cdot \mathbb{E}_{T_E^{k-1}}\left[\frac{V_k}{B(T_E^k)}\right]. \end{aligned}$$

We summarize the Bermudan pricing algorithm

Backward induction for Bermudan options

Consider a Bermudan option with $\bar{k}$ exercise dates T_E^k ($k = 1, \ldots \bar{k}$) and underlying future payoffs with (time-T_E^k) prices U_k.

Denote $H_k(t)$ the option's continuation value for $T_E^k \leq t \leq T_E^{k+1}$ and set $H_{\bar{k}}\left(T_E^{\bar{k}}\right) = 0$. Also set $T_E^0 = t$ (i.e. pricing time today).

The option price can be derived via the recursion

$$
\begin{aligned}
H_k\left(T_E^k\right) &= B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{H_k(T_E^{k+1})}{B(T_E^{k+1})}\right] \\
&= B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{\max\left\{U_{k+1}, H_{k+1}(T_E^{k+1})\right\}}{B(T_E^{k+1})}\right].
\end{aligned}
$$

for $k = \bar{k} - 1, \ldots, 0$. The Bermudan option price is given by

$$
V^{\text{Berm}}(t) = H_0(t) = H_0(T_E^0).
$$

Some more comments regarding Bermudan pricing ...

- Recursion for Bermudan pricing can be formally derived via theory of optimal stopping and Hamilton-Jacobi-Bellman (HJB) equation.
- For more details, see Sec. 18.2.2 in Andersen/Piterbarg (2010).
- For a single exercise date $\bar{k} = 1$ we get

$$H_0(t) = B(t) \cdot \mathbb{E}_t \left[\frac{\max \{U_1, 0)\}}{B(T_E^1)} \right].$$

  This is the general pricing formula for a European swaption (if U_1 represents a Vanilla swap).
- In principle, recursion $H_k \left(T_E^k \right) = B(T_E^k) \cdot \mathbb{E}_{T_E^k} \left[\frac{\max\left\{ U_{k+1}, H_{k+1}(T_E^{k+1}) \right\}}{B(T_E^{k+1})} \right]$ holds for any payoffs U_k. However, computation

$$U_k = B(T_E^k) \sum_{T_i \geq T_E^k} \mathbb{E}_{T_E^k} \left[\frac{X_i(T_i)}{B(T_i)} \right]$$

  might pose additional challenges if cash flows $X_i(T_i)$ are more complex.

How do we price a Bermudan in practice?

- In principle, recursion algorithm for $H_k()$ is straight forward.
- Computational challenge is calculating conditional expectations

$$H_k\left(T_E^k\right) = B(T_E^k) \cdot \mathbb{E}_{T_E^k}\left[\frac{\max\left\{U_{k+1}, H_{k+1}(T_E^{k+1})\right\}}{B(T_E^{k+1})}\right].$$

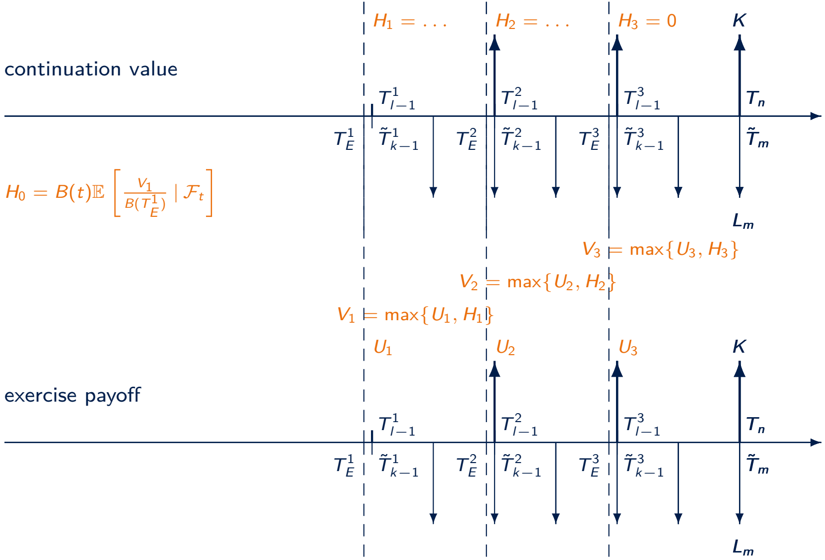
- Note, that this problem is an instance of the general option pricing problem

$$V(T_0) = B(T_0) \cdot \mathbb{E}\left[\frac{V(T_1)}{B(T_1)} \,|\, \mathcal{F}_{T_0}\right].$$

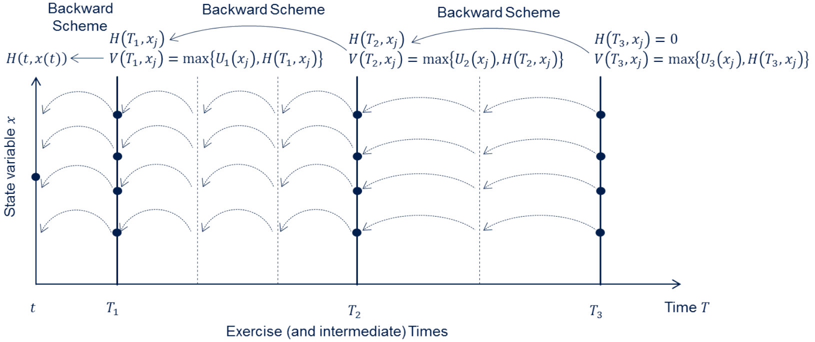
We can apply general option pricing methods to *roll-back* the Bermudan payoff.

Outline

Note that U_k, V_k and H_k depend on underlying state variable $x(T_E^k)$



Typically we need to discretise variables U_k, V_k and H_k on a grid of underlying state variables



Forthcomming, we discuss several methods to roll-back the payoffs.

Outline

Outline

Key idea is using the conditional density function in the Hull-White model

Recall that

$$V(T_0) = B(T_0) \cdot \mathbb{E}\left[\frac{V(T_1)}{B(T_1)} \,|\, \mathcal{F}_{T_0}\right].$$

We choose the T_1-maturing zero coupon bond $P(t, T_1)$ as numeraire. Then

$$\begin{aligned} V(T_0) &= P(T_0, T_1) \cdot \mathbb{E}^{T_1}\left[V(T_1) \,|\, \mathcal{F}_{T_0}\right] \\ &= P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx. \end{aligned}$$

State variable $x = x(T_1)$ is normally distributed with known mean and variance.

Hull-White model results yield density parameters of the state variable $x(T_1)$

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx.$$

State variable $x = x(T_1)$ is normally distributed with mean

$$\mu = \mathbb{E}^{T_1}\left[x(T_1) \,|\, \mathcal{F}_{T_0}\right] = G'(T_0, T_1)\left[x(T_0) + G(T_0, T_1)y(T_0)\right]$$

and variance

$$\sigma^2 = \mathrm{Var}\left[x(T_1) \,|\, \mathcal{F}_{T_0}\right] = y(T_1) - G'(T_0, T_1)^2 y(T_0).$$

Thus

$$p_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

and

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx.$$

Integral against normal density needs to be computed numerically by quadrature methods

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx.$$

- We can apply general purpose quadrature rules to function

$$f(x) = \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}.$$

  - Select a grid $[x_0, \ldots, x_N]$ and approximate e.g. via
  - Trapezoidal rule

$$\int_{-\infty}^{+\infty} f(x) \cdot dx \approx \sum_{i=1}^{N} \frac{1}{2}\left[f(x_{i-1}) + f(x_i)\right](x_i - x_{i-1})$$

  - or Simpson's rule with equidistant grid $(h = x_i - x_{i-1})$ and even sub-intervalls, then

$$\int_{-\infty}^{+\infty} f(x) \cdot dx \approx \frac{h}{3} \cdot \left[f(x_0) + 2\sum_{j=1}^{N/2-1} f(x_{2j}) + 4\sum_{j=1}^{N/2} f(x_{2j-1}) + f(x_N)\right].$$

There are some details that need to be considered - Select your integration domain carefully

- Infinite integral is approximated by definite integral

$$\int_{-\infty}^{+\infty} f(x) \cdot dx \approx \int_{x_0}^{x_N} f(x) \cdot dx \approx \cdots .$$

- Finite integration boundaries need to be chosen carefully by taking into account variance of $x(t)$.
- One approach is calculating variance to option expiry T_1, $\hat{\sigma}^2 = \text{Var}[x(T)] = y(T_1)$ and set

$$x_0 = -\lambda \cdot \hat{\sigma} \quad \text{and} \quad x_N = \lambda \cdot \hat{\sigma}.$$

- Note, that $\mathbb{E}^{T_1}[x(T_1)] = 0$, thus we do not apply a shift to the x-grid.

There are some details that need to be considered - Take care of the break-even point

- Note that convergence of quadrature rules depends on smoothness of integrand $f(x)$.
- Recall that

$$f(x) = V(x) \cdot p_{\mu,\sigma^2}(x) = \max\left\{U_{k+1}(x), H_{k+1}(x; T_E^{k+1})\right\} \cdot p_{\mu,\sigma^2}(x).$$

- Max-function is not smooth at $U_{k+1}(x) = H_{k+1}(x; T_E^{k+1})$.

Determine $x^\star$ (via interpolation of $H_{k+1}(\cdot)$ and numerical root search) such that

$$U_{k+1}(x^\star) = H_{k+1}(x^\star; T_E^{k+1})$$

and split integration

$$\int_{-\infty}^{+\infty} f(x) \cdot dx = \int_{-\infty}^{x^\star} f(x) \cdot dx + \int_{x^\star}^{+\infty} f(x) \cdot dx.$$

Can we exploit the structure of the integrand?

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} \frac{V(x; T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} dx.$$

- Integral against normal distribution density can be solved more efficiently:

1. Use Gauss–Hermite quadrature.
2. Interpolate only $V(x; T_1)$ by cubic spline and integrate exact.

Outline

Gauss–Hermite quadrature is an efficient integration method for smooth integrands

- Gauss–Hermite quadrature is a particular form of Gaussian quadrature.
- Choose a degree parameter d, and approximate

$$\int_{-\infty}^{+\infty} f(x) \cdot e^{-x^2} dx \approx \sum_{k=1}^{d} w_k \cdot f(x_k)$$

  with x_k $(i = 1, 2, ..., d)$ being the roots of the physicists' version of the Hermite polynomial $H_d(x)$ and w_k are weights with

$$w_k = \frac{2^{d-1} d! \sqrt{\pi}}{d^2 \left[H_{d-1}(x_k)\right]^2}.$$

- Roots and weights can be obtained, e.g. via stored values and look-up tables.

Variable transformation allows application of Gauss–Hermite quadrature to Hull-White model integration

We get

$$\int_{-\infty}^{+\infty} \frac{V(x;\,T_1)}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} V(\sqrt{2}\sigma x + \mu;\,T_1) \cdot e^{-x^2} dx$$

$$\approx \frac{1}{\sqrt{\pi}} \sum_{k=1}^{d} w_k \cdot V(\sqrt{2}\sigma x_k + \mu;\,T_1).$$

Some constraints need to be considered:

- Payoff $V(\cdot, T_1)$ is only available on the x-grid at T_1, thus $V(\cdot, T_1)$ needs to be interpolated.
- Gauss-Hermite quadrature does not take care of non-smooth payoff at break-even state $x^\star$, thus d needs to be sufficiently large to mitigate impact.

Outline

If we apply cubic spline interpolation anyway then we can also integrate exactly

Approximate $V(\cdot, T_1)$ via cubic spline on the grid $[x_0, \dots x_N]$ as

$$V(x, T_1) \approx C(x) = \sum_{i=0}^{N-1} \mathbb{1}_{\{x_i \le x < x_{i+1}\}} \sum_{k=0}^{d} c_{i,k} \cdot (x - x_i)^k \,.$$

Then

$$\begin{aligned}\int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx &\approx \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} \sum_{k=0}^{d} c_{i,k} \cdot (x - x_i)^k \cdot p_{\mu,\sigma^2}(x) \cdot dx \\ &= \sum_{i=0}^{N-1} \sum_{k=0}^{d} c_{i,k} \cdot \int_{x_i}^{x_{i+1}} (x - x_i)^k \cdot p_{\mu,\sigma^2}(x) \cdot dx.\end{aligned}$$

Thus, all we need is

$$I_{i,k} = \int_{x_i}^{x_{i+1}} (x - x_i)^k \cdot p_{\mu,\sigma^2}(x) \cdot dx.$$

We transform variables to make integration easier

First we apply the variable transformation $\bar{x} = (x - \mu)/\sigma$. This yields $p_{\mu,\sigma^2}(x) = p_{0,1}(\bar{x})/\sigma$ and

$$
\begin{aligned}
I_{i,k} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} (\sigma \bar{x} + \mu - x_i)^k \cdot p_{0,1}(\bar{x}) \cdot \frac{dx}{\sigma} \\
&= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^k \left(\bar{x} - \bar{x}_i\right)^k \cdot \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\bar{x}^2}{2}\right\}}_{\text{standard normal density}} d\bar{x}
\end{aligned}
$$

with the shifted grid points $\bar{x}_i = (x_i - \mu)/\sigma$.
Denote $\Phi(\cdot)$ the cumulated standard normal distribution function. Then

$$
\Phi'(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\bar{x}^2}{2}\right\} \quad \text{and} \quad \Phi''(x) = -x\Phi'(x).
$$

As a sub-step we aim at solving the integral

$$
\int_{\bar{x}_i}^{\bar{x}_{i+1}} \bar{x}^k \cdot \Phi'(\bar{x}) \cdot d\bar{x}.
$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple I

It turnes out that

$$\begin{aligned} F_0(\bar{x}) &= \int \Phi'(\bar{x}) d\bar{x} = \Phi(\bar{x}), \\ F_1(\bar{x}) &= \int \bar{x}\Phi'(\bar{x}) d\bar{x} = -\Phi'(\bar{x}), \\ F_2(\bar{x}) &= \int \bar{x}^2\Phi'(\bar{x}) d\bar{x} = \Phi(\bar{x}) - x \cdot \Phi'(\bar{x}), \\ F_3(\bar{x}) &= \int \bar{x}^3\Phi'(\bar{x}) d\bar{x} = -\left(\bar{x}^2 + 2\right) \cdot \Phi'(\bar{x}). \end{aligned}$$

This yields for $I_{i,0}$

$$I_{i,0} = \int_{\bar{x}_i}^{\bar{x}_{i+1}} \Phi'(\bar{x}) \cdot dx = F_0(\bar{x}_{i+1}) - F_0(\bar{x}_i)$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple II

and for $I_{i,1}$

$$\begin{aligned} I_{i,1} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma \left(\bar{x} - \bar{x}_i\right) \cdot \Phi'(\bar{x}) \cdot dx \\ &= \sigma \cdot \int_{\bar{x}_i}^{\bar{x}_{i+1}} \bar{x} \cdot \Phi'(\bar{x}) \cdot dx - \sigma \cdot \bar{x}_i \cdot I_{i,0} \\ &= \sigma \cdot [F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)] - \sigma \cdot \bar{x}_i \cdot I_{i,0}. \end{aligned}$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple III

We may proceed similarly for $I_{i,2}$

$$
\begin{aligned}
I_{i,2} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^2 \left(\bar{x} - \bar{x}_i\right)^2 \cdot \Phi'(\bar{x}) \cdot dx \\
&= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^2 \left[\bar{x}^2 - 2\bar{x}_i\bar{x} + \bar{x}_i^2\right] \cdot \Phi'(\bar{x}) \cdot dx \\
&= \sigma^2 \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] - 2\sigma^2\bar{x}_i \left[F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)\right] + \sigma^2\bar{x}_i^2 I_{i,0} \\
&= \sigma^2 \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] - 2\sigma\bar{x}_i \left[I_{i,1} + \sigma \cdot \bar{x}_i \cdot I_{i,0}\right] + \sigma^2\bar{x}_i^2 I_{i,0} \\
&= \sigma^2 \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] - 2\sigma\bar{x}_i I_{i,1} - \sigma^2\bar{x}_i^2 I_{i,0}
\end{aligned}
$$

We use cubic splines ($d = 3$) to keep formulas reasonably simple IV

and $I_{i,3}$

$$\begin{aligned}
I_{i,3} &= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^3 \left(\bar{x} - \bar{x}_i\right)^3 \cdot \Phi'(\bar{x}) \cdot dx \\
&= \int_{\bar{x}_i}^{\bar{x}_{i+1}} \sigma^3 \left[\bar{x}^3 - 3\bar{x}_i\bar{x}^2 + 3\bar{x}_i^2\bar{x} - \bar{x}_i^3\right] \cdot \Phi'(\bar{x}) \cdot dx \\
&= \sigma^3 \left[F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)\right] - 3\sigma^3\bar{x}_i \left[F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)\right] \\
&\quad + 3\sigma^3\bar{x}_i^2 \left[F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)\right] - \sigma^3\bar{x}_i^3 I_{i,0}.
\end{aligned}$$

Substituting terms as before yields

$$\begin{aligned}
I_{i,3} &= \sigma^3 \left[F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)\right] - 3\sigma\bar{x}_i \left[I_{i,2} + 2\sigma\bar{x}_i I_{i,1} + \sigma^2\bar{x}_i^2 I_{i,0}\right] \\
&\quad + 3\sigma^2\bar{x}_i^2 \left[I_{i,1} + \sigma \cdot \bar{x}_i \cdot I_{i,0}\right] - \sigma^3\bar{x}_i^3 I_{i,0} \\
&= \sigma^3 \left[F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)\right] - 3\sigma\bar{x}_i I_{i,2} - 3\sigma^2\bar{x}_i^2 I_{i,1} - \sigma^3\bar{x}_i^3 I_{i,0}.
\end{aligned}$$

Let's summarise the formulas...

We get

$$V(T_0) = P(x(T_0); T_0, T_1) \cdot \int_{-\infty}^{+\infty} V(x; T_1) \cdot p_{\mu,\sigma^2}(x) \cdot dx$$
$$\approx P(x(T_0); T_0, T_1) \cdot \sum_{i=0}^{N-1} \sum_{k=0}^{3} c_{i,k} \cdot I_{i,k}$$

with

$$I_{i,0} = F_0(\bar{x}_{i+1}) - F_0(\bar{x}_i)$$
$$I_{i,1} = \sigma \cdot [F_1(\bar{x}_{i+1}) - F_1(\bar{x}_i)] - \sigma \cdot \bar{x}_i \cdot I_{i,0}$$
$$I_{i,2} = \sigma^2 [F_2(\bar{x}_{i+1}) - F_2(\bar{x}_i)] - 2\sigma\bar{x}_i I_{i,1} - \sigma^2 \bar{x}_i^2 I_{i,0}$$
$$I_{i,3} = \sigma^3 [F_3(\bar{x}_{i+1}) - F_3(\bar{x}_i)] - 3\sigma\bar{x}_i I_{i,2} - 3\sigma^2 \bar{x}_i^2 I_{i,1} - \sigma^3 \bar{x}_i^3 I_{i,0}$$

and anti-derivative functions $F_k(x)$ as before.

Integrating a cubic spline versus a normal density function occurs in various contexts of pricing methods

- Method already yields good accuracy for smaller number of grid points.
- For larger number of grid points accuracy benefit compared to e.g. Simpson integration seems not too much.
- Either way, use special treatment of break-even point $x^\star$.
- Computational effort can become significant for larger number of grid points.
  - Bermudan pricing requires N^2 evaluations of $\Phi(\cdot)$ and $\Phi'(\cdot)$ per exercise.

Outline

PDE methods for finance and pricing are extensively studied in the literature

- We present the basic principles and some aspects relevant for Bermudan bond option pricing.
- Further reading:
  - L. Andersen and V. Piterbarg. *Interest rate modelling, volume I to III.* Atlantic Financial Press, 2010, Sec. 2.
  - D. Duffy. *Finite Difference Methods in Financial Engineering.* Wiley Finance, 2006

Outline

We can adapt the Black-Scholes equation to our Hull-White model setting

- Recall that state variable $x(t)$ follows the risk-neutral dynamics

$$dx(t) = \underbrace{[y(t) - a \cdot x(t)]}_{\mu(t,x(t))} dt + \sigma(t) \cdot dW(t).$$

- Consider an option with price $V = V(t, x(t))$, option expiry time T and payoff $V(T, x(T)) = g(x(T))$.
- Derivative pricing formula yields that discounted option price is a martingale, i.e.

$$d\left(\frac{V(t, x(t))}{B(t)}\right) = \sigma_V(t, x(t)) \cdot dW(t).$$

How can we use this to derive a PDE?

Apply Ito's Lemma to the discounted option price

We get

$$d\left(\frac{V(t,x(t))}{B(t)}\right) = \frac{dV(t,x(t))}{B(t)} + V(t)d\left(\frac{1}{B(t)}\right).$$

With $d\left(B(t)^{-1}\right) = -r(t)\cdot B(t)^{-1}\cdot dt$ follows

$$d\left(\frac{V(t,x(t))}{B(t)}\right) = \frac{1}{B(t)}\left[dV(t,x(t)) - r(t)\cdot V(t)\cdot dt\right].$$

Applying Ito's Lemma to option price $V(t,x(t))$ gives

$$\begin{aligned} dV(t,x(t)) &= V_t\cdot dt + V_x\cdot dx(t) + \frac{1}{2}V_{xx}\cdot[dx(t)]^2 \\ &= \left[V_t + V_x\cdot\mu(t,x(t)) + \frac{1}{2}V_{xx}\cdot\sigma(t)^2\right]dt + V_x\cdot\sigma(t)\cdot dW(t) \end{aligned}$$

with partial derivatives $V_t = \partial V(t,x(t))/\partial t$, $V_x = \partial V(t,x(t))/\partial x$ and $V_{xx} = \partial^2 V(t,x(t))/\partial x^2$.

Combining results yields dynamics of discounted option price

$$d\left(\frac{V\left(t,x(t)\right)}{B(t)}\right) = \frac{1}{B(t)}\underbrace{\left[V_t + V_x \cdot \mu\left(t,x(t)\right) + \frac{1}{2}V_{xx}\cdot\sigma(t)^2 - r(t)\cdot V\right]}_{\mu_V(t,x(t))}dt + \underbrace{\frac{V_x\cdot\sigma(t)}{B(t)}}_{\sigma_V(t,x(t))}\cdot dW(t).$$

Martingale property of $\frac{V(t,x(t))}{B(t)}$ requires that drift vanishes. That is

$$\mu_V(t,x(t)) = V_t + V_x\cdot\mu\left(t,x(t)\right) + \frac{1}{2}V_{xx}\cdot\sigma(t)^2 - r(t)\cdot V = 0.$$

Substituting $\mu\left(t,x(t)\right) = y(t) - a\cdot x(t)$ and $r(t) = f(0,t) + x(t)$ yields pricing PDE.

We get the parabolic pricing PDE with terminal condition

Theorem (Derivative pricing PDE in Hull-White model)

Consider our Hull-White model setup and a derivative security with price process $V(t, x(t))$ that pays at time T the payoff $V(T, x(T)) = g(x(T))$. Further assume $V(T, x(T))$ has finite variance and is attainable.
Then for $t < T$ the option price

$$V(t, x(t)) = B(t) \cdot \mathbb{E}^{\mathbb{Q}} \left[\frac{V(T, x(T))}{B(T)} \,|\, \mathcal{F}_t \right]$$

follows the PDE

$$V_t(t,x) + [y(t) - a \cdot x] \cdot V_x(t,x) + \frac{\sigma(t)^2}{2} \cdot V_{xx}(t,x) = [x + f(0,t)] \cdot V(t,x)$$

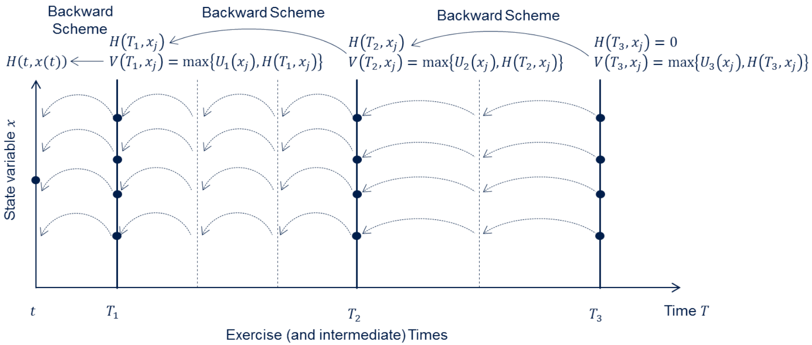
with terminal condition

$$V(T,x) = g(x).$$

Proof.

Follows from derivation above. □

How does this help for our Bermudan option pricing problem?



- We need option prices on a grid of state variables $[x_0, \ldots x_N]$

Solve Hull-White option pricing PDE backwards from exercise to exercise.

Pricing PDE is typically solved via finite difference scheme and time integration

- Use *method of lines (MOL)* to solve parabolic PDE:
  - First discretise state space.
  - Then integrate resulting system of ODEs with terminal condition in time-direction.
- We discuss basic (or general purpose) approach to solve PDE; for a detailed treatment see Andersen/Piterbarg (2010) or Duffy (2006).
- Some aspects may require special attention in the context of Hull-White model:
  - more problem-specific boundary discretisation,
  - non-equidistant grids with finer grid around break-even state $x^\star$,
  - upwinding schemes to deal with potentially dominant impact of convection term $[y(t) - a \cdot x] \cdot V_x(t, x)$ at the grid boundaries of x.

Outline

How do we discretise state space?

- PDE for $V(t, x(t))$ is defined on infinite domain $(-\infty, +\infty)$.
  - We don't get explicit boundary conditions from PDE derivation.
  - Thus, we require payoff-specific approximation.
  - Finally, we are interested in $V(0, 0)$.
- We use equidistant x-grid $x_0, \ldots, x_N$ with grid size h_x centered around zero and scaled via standard deviation of $x(T)$ at final maturity T,

$$x_0 = -\lambda \cdot \hat{\sigma} \quad \text{and} \quad x_N = \lambda \cdot \hat{\sigma}$$

  with $\hat{\sigma}^2 = \text{Var}[x(T)] = y(T)$ and $\lambda \approx 5$.
- Why not shift the grid by expectation $\mathbb{E}[x(T)]$ (as suggested in the literature)?
  - Pricing PDE is independent of pricing measure (used for derivation).
  - There is no *natural* measure under which $\mathbb{E}[x(T)]$ should be calculated.
  - In T-forward measure $\mathbb{E}^T[x(T)] = 0$ anyway.

Differential operators in state-dimension are discretised via central finite differences

For now leave time t continuous. We use notation $V(\cdot, x)$.
For inner grid points x_i with $i = 1, \ldots, N-1$

$$V_x(\cdot, x_i) = \frac{V(\cdot, x_{i+1}) - V(\cdot, x_{i-1})}{2h_x} + \mathcal{O}(h_x^2) \quad \text{and}$$

$$V_{xx}(\cdot, x_i) = \frac{V(\cdot, x_{i+1}) - 2V(\cdot, x_i) + V(\cdot, x_{i-1})}{h_x^2} + \mathcal{O}(h_x^2).$$

At the boundaries we impose condition

$$V_{xx}(\cdot, x_0) = \lambda_0 \cdot V_x(\cdot, x_0) \quad \text{and} \quad V_{xx}(\cdot, x_N) = \lambda_N \cdot V_x(\cdot, x_N).$$

This yields one-sided first order partial derivative approximations

$$V_x(\cdot, x_0) \approx \frac{2\left[V(\cdot, x_1) - V(\cdot, x_0)\right]}{(2 + \lambda_0 h_x)\, h_x} \quad \text{and} \quad V_x(\cdot, x_N) \approx \frac{2\left[V(\cdot, x_N) - V(\cdot, x_{N-1})\right]}{(2 - \lambda_N h_x)\, h_x}.$$

Some initial comments regarding choice of $\lambda_{0,N}$

- Often, $\lambda_{0,N} = 0$ (also suggested in the literature).

- With $\lambda_{0,N} = 0$ we have $V_{xx}(\cdot, x_0) = V_{xx}(\cdot, x_N) = 0$ and

$$V_x(\cdot, x_0) = \frac{V(\cdot, x_1) - V(\cdot, x_0)}{h_x} + \mathcal{O}(h_x^2) \quad \text{and}$$

$$V_x(\cdot, x_N) = \frac{V(\cdot, x_N) - V(\cdot, x_{N-1})}{h_x} + \mathcal{O}(h_x^2).$$

- However, for bond options the choice $V_{xx}(\cdot, x_0) = V_{xx}(\cdot, x_N) = 0$ might be a poor approximation.

- We will discuss an alternative choice for $\lambda_{0,N}$ later.

Now consider PDE for each grid point individually

Define the vector-valued function $v(t)$ via

$$v(t) = [v_0(t), \ldots, v_N(t)]^\top = [V(t, x_0), \ldots, V(t, x_N)]^\top \in \mathbb{R}^{N+1}.$$

Then state discretisation yields for inner points x_i with $i = 1, \ldots, N-1$,

$$v_i'(t) + [y(t) - a x_i] \frac{v_{i+1}(t) - v_{i-1}(t)}{2h_x} + \frac{\sigma(t)^2}{2} \frac{v_{i+1}(t) - 2v_i(t) + v_{i-1}(t)}{h_x^2} = [x_i + f(0,t)]\, v_i(t)$$

and for the boundaries

$$v_0'(t) + \left[y(t) - a x_0 + \lambda_0 \frac{\sigma(t)^2}{2} \right] \frac{2\,[v_1(t) - v_0(t)]}{(2 + \lambda_0 h_x)\, h_x} = [x_0 + f(0,t)]\, v_0(t),$$

$$v_N'(t) + \left[y(t) - a x_N + \lambda_N \frac{\sigma(t)^2}{2} \right] \frac{2\,[v_N(t) - v_{N-1}(t)]}{(2 - \lambda_N h_x)\, h_x} = [x_N + f(0,t)]\, v_N(t).$$

As before, we have the terminal condition

$$v_i(T) = g(x_i).$$

Parabolic PDE is transformed into linear system of ODEs with terminal condition.

It is more convenient to write system of ODEs in matrix-vector notation (1/2)

We get

$$v'(t) = M(t) \cdot v(t) = \begin{bmatrix} c_0 & u_0 & & \\ l_1 & \ddots & \ddots & \\ & \ddots & \ddots & u_{N-1} \\ & & l_N & c_N \end{bmatrix} \cdot v(t)$$

with time-dependent inner components c_i, l_i, u_i $(i = 1, \dots N-1)$,

$$c_i = \frac{\sigma(t)^2}{h_x^2} + x_i + f(0,t),$$

$$l_i = -\frac{\sigma(t)^2}{2h_x^2} + \frac{y(t) - ax_i}{2h_x},$$

$$u_i = -\frac{\sigma(t)^2}{2h_x^2} - \frac{y(t) - ax_i}{2h_x}.$$

It is more convenient to write system of ODEs in matrix-vector notation (2/2)

Boundary elements of $M(t)$ become

$$
\begin{aligned}
c_0 =& \frac{2\left[y(t) - ax_0 + \lambda_0 \frac{\sigma(t)^2}{2}\right]}{(2+\lambda_0 h_x)\, h_x} + x_0 + f(0,t), \\
c_N =& -\frac{2\left[y(t) - ax_N + \lambda_N \frac{\sigma(t)^2}{2}\right]}{(2-\lambda_N h_x)\, h_x} + x_0 + f(0,t), \\
u_0 =& -\frac{2\left[y(t) - ax_0 + \lambda_0 \frac{\sigma(t)^2}{2}\right]}{(2+\lambda_0 h_x)\, h_x}, \\
l_N =& \frac{2\left[y(t) - ax_N + \lambda_N \frac{\sigma(t)^2}{2}\right]}{(2-\lambda_N h_x)\, h_x}.
\end{aligned}
$$

Outline

Linear system of ODEs can be solved by standard methods

We have

$$v'(t) = f\left(t, v(t)\right) = M(t) \cdot v(t).$$

We demonstrate time discretisation based on θ*-method.* Consider equidistant time grid $t = t_0, \ldots, t_M = T$ with step size h_t and approximation

$$\frac{v(t_{j+1}) - v(t_j)}{h_t} \approx f\left(t_{j+1} - \theta h_t, (1-\theta)v(t_{j+1}) + \theta v(t_j)\right)$$

for $\theta \in [0, 1]$.

- In general, approximation yields method of order $\mathcal{O}(h_t)$.
- For $\theta = \frac{1}{2}$, approximation yields method of order $\mathcal{O}(h_t^2)$.

For our linear ODE we set $v^j = v(t_j)$, $M_\theta = M(t_{j+1} - \theta h_t)$ and get the scheme

$$\frac{v^{j+1} - v^j}{h_t} = M_\theta\left[(1-\theta)v^{j+1} + \theta v^j\right].$$

We get a recursion for the θ-method

Rearranging terms yields

$$[I + h_t \theta M_\theta] \, v^j = [I - h_t (1-\theta) M_\theta] \, v^{j+1}.$$

If $[I + h_t \theta M_\theta]$ is regular then we can solve for v^j via

$$v^j = [I + h_t \theta M_\theta]^{-1} [I - h_t (1-\theta) M_\theta] \, v^{j+1}.$$

Terminal condition is

$$v^M = [g(x_0), \dots, g(x_N)]^\top .$$

- Unless $\theta = 0$ (Explicit Euler scheme) we need to solve a linear equation system.
- Fortunately, matrix $[I + h_t \theta M_\theta]$ is tri-diagonal; solution requires $\mathcal{O}(M)$ operations.
- θ-method is A-stable for $\theta \geq \frac{1}{2}$.
- However, oscillations in solution may occur unless $\theta = 1$ (Implicit Euler scheme, L-stable).

Outline

Let's have another look at the boundary condition ...

We look at an example of a zero coupon bond option with payoff

$$V(x,T) = \left[P(x,T,T') - K\right]^+ .$$

For $x \ll 0$ option is far in-the-money and $V(x,t)$ can be approximated by intrinsic value $V(x,t) \approx \tilde{V}(x,t)$ with

$$\tilde{V}(x,t) = \left[P(x,t,T') - K\right]^+ = \left[\frac{P(0,T')}{P(0,t)} e^{-G(t,T)x - \frac{1}{2}G(t,T)^2 y(t)} - K\right]^+ .$$

This yields

$$\frac{\partial}{\partial x}\tilde{V}(x,t) = -G(t,T)\left[\tilde{V}(x,t) + K\right]$$

and

$$\frac{\partial^2}{\partial x^2}\tilde{V}(x,t) = \underbrace{-G(t,T)}_{\lambda}\frac{\partial}{\partial x}\tilde{V}(x,t).$$

Alternatively, for $x \gg 0$ option is far out-of-the-money and

$$\frac{\partial^2}{\partial x^2}\tilde{V}(x,t) = \frac{\partial}{\partial x}\tilde{V}(x,t) = 0.$$

We adapt approximation to our option pricing problem

- In principle, for a coupon bond underlying we could estimate $\lambda = \lambda(t)$ via option intrinsic value $\tilde{V}(x,t)$ and

$$\lambda(t) = \left[\frac{\partial^2}{\partial x^2}\tilde{V}(x,t)\right] / \frac{\partial}{\partial x}\tilde{V}(x,t) \quad \text{for} \quad \frac{\partial}{\partial x}\tilde{V}(x,t) \neq 0,$$

  otherwise $\lambda(t) = 0$.

- We take a more rough approach by approximating λ based only on previous solution

$$\begin{aligned}\lambda_{0,N} &= \left[\frac{\partial^2}{\partial x^2}V(x,t)\right] / \frac{\partial}{\partial x}V(x,t) \\ &\approx \left[\frac{\partial^2}{\partial x^2}V(x_{1,N-1},t+h_t)\right] / \frac{\partial}{\partial x}V(x_{1,N-1},t+h_t) \\ &\approx \frac{v_{0,N-2}^{j+1} - 2v_{1,N-1}^{j+1} + v_{2,N}^{j+1}}{h_x^2} / \frac{v_{2,N}^{j+1} - v_{0,N-2}^{j+1}}{2h_x}\end{aligned}$$

  for $v_{2,N}^{j+1} - v_{0,N-2}^{j+1}/(2h_x) \neq 0$, otherwise $\lambda_{0,N} = 0$.

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ I

Lemma

Assume $V = V(x)$ is twice continuously differentiable. Moreover, consider grid points x_{-1}, x_0, x_1 with equal spacing $h_x = x_1 - x_0 = x_0 - x_{-1}$. If there is a $\lambda_0 \in \mathbb{R}$ such that

$$V''(x_0) = \lambda_0 \cdot V'(x_0)$$

then

$$V'(x_0) = \frac{2\left[V(x_1) - V(x_0)\right]}{(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2).$$

Proof:

Denote $v_i = V(x_i)$. We have from standard Taylor approximation

$$V''(x_0) = \frac{v_{-1} - 2v_0 + v_1}{h_x^2} + \mathcal{O}(h_x^2) \quad \text{and} \quad V'(x_0) = \frac{v_1 - v_{-1}}{2h_x} + \mathcal{O}(h_x^2).$$

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ II

From $V''(x_0) = \lambda \cdot V'(x_0)$ follows

$$\frac{v_{-1} - 2v_0 + v_1}{h_x^2} + \mathcal{O}(h_x^2) = \lambda_0 \left[\frac{v_1 - v_{-1}}{2h_x} + \mathcal{O}(h_x^2) \right].$$

Multiplying with $2h_x^2$ gives the relation

$$2\left(v_{-1} - 2v_0 + v_1\right) + \mathcal{O}(h_x^4) = \lambda_0 h_x \left(v_1 - v_{-1}\right) + \mathcal{O}(h_x^4).$$

Reordering terms yields

$$\left(2 + \lambda_0 h_x\right) v_{-1} = 4v_0 + \left(\lambda_0 h_x - 2\right) v_1 + \mathcal{O}(h_x^4).$$

And solving for v_{-1} gives
$v_{-1} = \left[4v_0 + \left(\lambda_0 h_x - 2\right) v_1\right] / \left(2 + \lambda_0 h_x\right) + \mathcal{O}(h_x^4)$.

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ III

Now, we substitute v_{-1} in the approximation for $V'(x)$. This gives

$$\begin{aligned}
V'(x_0) &= \frac{v_1 - \left[\left[4v_0 + (\lambda_0 h_x - 2)\, v_1\right] / (2 + \lambda_0 h_x) + \mathcal{O}(h_x^4)\right]}{2h_x} + \mathcal{O}(h_x^2) \\
&= \frac{(2 + \lambda_0 h_x)\, v_1 - \left[4v_0 + (\lambda_0 h_x - 2)\, v_1\right]}{2\,(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2) + \mathcal{O}(h_x^3) \\
&= \frac{2v_1 - 4v_0 + 2v_1}{2\,(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2) \\
&= \frac{2\,(v_1 - v_0)}{(2 + \lambda_0 h_x)\, h_x} + \mathcal{O}(h_x^2).
\end{aligned}$$

- With constraint $V''(x_0) = \lambda \cdot V'(x_0)$ we can eliminate explicit dependence on second derivative $V''(x_0)$ and outer grid point $v_{-1} = V(x_{-1})$.

It turns out that accuracy of one-sided first order derivative approximation is of order $\mathcal{O}(h_x^2)$ IV

- Analogous result can be derived for upper boundery and down-ward approximation of first derivative.
- Resulting scheme is still second order accurate in state space direction.

Outline

We summarise the PDE pricing method

1. Discretise state space x on a grid $[x_0, \ldots, x_N]$ and specify time step size h_t and $\theta \in [0, 1]$.

2. Determine the terminal condition $v^{j+1} = \max\{U_{j+1}, H_{j+1}\}$ for the current valuation step.

3. Set up discretised linear operator M_θ of the resulting ODE system $\frac{d}{dt} v = M_\theta \cdot v$.

4. Incorporate appropriate product-specific boundary conditons.

5. Set up linear system $[I + h_t \theta M_\theta] v^j = [I - h_t (1 - \theta) M_\theta] v^{j+1}$.

6. Solve linear system for v^j by tri-diagonal matrix solver.

7. Repeat with step 3. until next exercise date or $t_j = 0$.

Outline

Monte Carlo methods are widely applied in various finance applications

- We demonstrate the basic principles for
  - path integration of Ito processes
  - exact simulation of Hull-White model paths
- There are many aspects that should also be considered, see e.g.
  - L. Andersen and V. Piterbarg. *Interest rate modelling, volume I to III.* Atlantic Financial Press, 2010, Sec. 3.
  - P. Glasserman. *Monte Carlo Methods in Financial Engineering.* Springer, 2003

Outline

Monte Carlo (MC) pricing is based on the Strong Law of Large Numbers

Theorem (Strong Law of Large Numbers)

Let $Y_1, Y_2, \ldots$ be a sequence of independent identically distributed (i.i.d.) random variables with finite expectation $\mu < \infty$. Then the sample mean $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ converges to μ a.s. That is

$$\lim_{n\to\infty} \bar{Y}_n = \mu \quad a.s.$$

- We aim at calculating $V(t) = N(t) \cdot \mathbb{E}^N\left[V(T)/N(T)\,|\,\mathcal{F}_t\right]$.
- For MC pricing simulate future discounted payoffs $\left\{\frac{V(T;\omega_i)}{N(T;\omega_i)}\right\}_{i=1,2,\ldots n}$.
- And estimate

$$V(t) = N(t) \cdot \frac{1}{n}\sum_{i=1}^{n} \frac{V(T;\omega_i)}{N(T;\omega_i)}.$$

Keep in mind that sample mean is still a random variable governed by central limit theorem (1/2)

Theorem (Central Limit Theorem)

Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables with finite expectation $\mu < \infty$ and standard deviation $\sigma < \infty$. Denote the sample mean $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Then

$$\frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1).$$

Moreover, for the variance estimator $s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}_n\right)^2$ we also have

$$\frac{\bar{Y}_n - \mu}{s_n/\sqrt{n}} \xrightarrow{d} N(0,1).$$

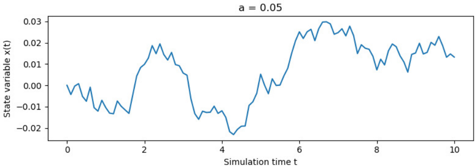
Keep in mind that sample mean is still a random variable governed by central limit theorem (2/2)

$$\frac{\bar{Y}_n - \mu}{s_n/\sqrt{n}} \xrightarrow{d} N(0,1).$$

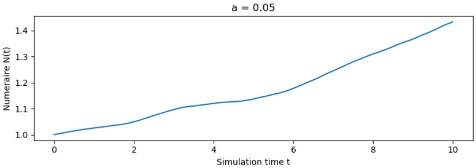
- Here, $N(0,1)$ is the standard normal distribution.
- $\xrightarrow{d}$ denotes convergence in distribution, i.e. $\lim_{n\to\infty} F_n(x) = F(x)$ for the corresponding cumulative distribution functions and all $x \in \mathbb{R}$ at which $F(x)$ is continuous.
- $s_n/\sqrt{n}$ is the standard error of the sample mean $\bar{Y}_n$.

How do we get our samples $V(T;\omega_i)/N(T;\omega_i)$?

1. Simulate state variables $x(t)$ on relevant dates t:



2. Simulate numeraire $N(t)$ on relevant dates t:



3. Calculate payoff $V(T, x(T))$ at observation/pay date T.

We need to simulate our state variables on the relevant observation dates

Consider the general dynamics for a process given as SDE

$$dX(t) = \mu(t, X(t)) \cdot dt + \sigma(t, X(t)) \cdot dW(t).$$

- Typically, we know initial value $X(t)$ ($t = 0$).
- We need $X(T)$ for some future time $T > t$.
- In Hull-White model and risk-neutral measure formulation we have

$$\mu(t, X(t)) = y(t) - a \cdot X(t), \quad \text{and}, \quad \sigma(t, X(t)) = \sigma(t).$$

There are several standard methods to solve above SDE. We will briefly discuss Euler method and Milstein method.

Euler method for SDEs is similar to Explicit Euler method for ODEs

- Specify a grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.

- Calculate sequence of state variables

$$X_{k+1} = X_k + \mu(t_k, X_k)\,(t_{k+1} - t_k) + \sigma(t_k, X_k)\,[W(t_{k+1}) - W(t_k)]\,.$$

- Drift $\mu(t_k, X_k)$ and volatility $\sigma(t_k, X_k)$ are evaluated at current time t_k and state X_k.

- Increment of Brownian motion $W(t_{k+1}) - W(t_k)$ is normally distributed, i.e.

$$W(t_{k+1}) - W(t_k) = Z_k \cdot \sqrt{t_{k+1} - t_k} \quad \text{with} \quad Z_k \sim N(0,1).$$

Milstein method refines the simulation of the diffusion term (1/2)

- Again, specify a grid of simulation times $t = t_0, t_1, \dots, t_M = T$.
- Calculate sequence of state variables

$$X_{k+1} = X_k + \mu(t_k, X_k)\,(t_{k+1} - t_k) + \sigma(t_k, X_k)\,[W(t_{k+1}) - W(t_k)]$$
$$+ \frac{1}{2}\sigma(t_k, X_k)\frac{\partial \sigma(t_k, X_k)}{\partial x}\left[(W(t_{k+1}) - W(t_k))^2 - (t_{k+1} - t_k)\right].$$

- Drift $\mu(t_k, X_k)$ and volatility $\sigma(t_k, X_k)$ are evaluated at current time t_k and state X_k.

Milstein method refines the simulation of the diffusion term (2/2)

- Requires calculation of derivative of volatility $\frac{\partial}{\partial x}\sigma(t_k, X_k)$ w.r.t. state variable.
- Increment of Brownian motion $W(t_{k+1}) - W(t_k)$ is normally distributed, i.e.

$$W(t_{k+1}) - W(t_k) = Z_k \cdot \sqrt{t_{k+1} - t_k} \quad \text{with} \quad Z_k \sim N(0,1).$$

- With $\Delta_k = t_{k+1} - t_k$ iteration becomes

$$\begin{aligned} X_{k+1} = X_k &+ \mu(t_k, X_k)\Delta_k + \sigma(t_k, X_k) Z_k \sqrt{\Delta_k} \\ &+ \frac{1}{2}\sigma(t_k, X_k)\frac{\partial\sigma(t_k, X_k)}{\partial x}\left(Z_k^2 - 1\right)\Delta_k. \end{aligned}$$

How can we measure convergence of the methods?

- We distinguish strong order of convergence and weak order of convergence.
- Consider a discrete SDE solution $\{X_k^h\}_{k=0}^M$ with $X_k^h \approx X(t + kh)$, $h = \frac{T-t}{M}$.

Definition (Strong order of convergence)

The discrete solution X_M^h at final maturity $T = t + hM$ converges to the exact solution $X(T)$ with strong orderβ if there exists a constant C such that

$$\mathbb{E}\left[\left|X_M^h - X(T)\right|\right] \leq C \cdot h^\beta.$$

- Strong order of convergence focuses on convergence on the individual paths.
- Euler method has strong order of convergence of $\frac{1}{2}$ (given sufficient conditions on $\mu(\cdot)$ and $\sigma(\cdot)$).
- Milstein method has strong order of convergence of 1 (given sufficient conditions on $\mu(\cdot)$ and $\sigma(\cdot)$).

For derivative pricing we are typically interested in weak order of convergence

We need some context for weak order of convergence

- A function $f : \mathbb{R} \to \mathbb{R}$ is polynomially bounded if $|f(x)| \le k(1+|x|)^q$ for constants k and q and all x.
- The set $\mathcal{C}_{\mathcal{P}}^{n}$ represents all functions that are n-times continuously differentiable and with 1st to nth derivative polynomially bounded.

Definition (Weak order of convergence)

The discrete solution X_M^h at final maturity $T = t + hM$ converges to the exact solution $X(T)$ with weak orderβ if there exists a constant C such that

$$\left|\mathbb{E}\left[f\left(X_M^h\right)\right] - \mathbb{E}\left[f\left(X(T)\right)\right]\right| \le C \cdot h^{\beta} \quad \forall f \in \mathcal{C}_{\mathcal{P}}^{2\beta+2}$$

for sufficiently small h.

- Think of f as a payoff function, then weak order of convergence is related to convergence in price.
- Euler method and Milstein method can be shown to have weak order 1 convergence (given sufficient conditions on μ and σ).

Some comments regarding weak order of convergence

Error estimate

$$\left|\mathbb{E}\left[f\left(X_M^h\right)\right] - \mathbb{E}\left[f\left(X(T)\right)\right]\right| \leq C \cdot h^\beta$$

requires considerable assumptions regarding smoothness of $\mu(\cdot)$, $\sigma(\cdot)$ and test functions $f(\cdot)$.

- In practice payoffs are typically non-smooth at the strike.
- This limits applicability of more advanced schemes with theoretical higher order of convergence.
- A fairly simple approach of a higher order scheme is based on Richardson extrapolation:
  - this method is also applied to ODEs,
  - see Glassermann (2000), Sec. 6.2.4 for details.
- Typically, numerical testing is required to assess convergence in practice.

The choice of pricing measure is crucial for numeraire simulation

Consider risk-neutral measure, then

$$N(T) = B(T) = \exp\left\{\int_0^T r(s)ds\right\} = \exp\left\{\int_0^T \left[f(0,s) + x(s)\right] ds\right\}$$
$$= P(0,T)^{-1} \exp\left\{\int_0^T x(s)ds\right\}.$$

Requires simulation or approximation of $\int_0^T x(s)ds$.

Suppose $x(t_k)$ is simulated on a time grid $\{t_k\}_{k=0}^M$ then we approximate integral via Trapezoidal rule

$$\int_0^T x(s)ds \approx \sum_{i=1}^M \frac{x(t_{k-1}) + x(t_k)}{2} (t_k - t_{k-1}).$$

Numeraire simulation is done in parallel to state simulation

$$N(t_k) = \frac{P(0,t_{k-1})}{P(0,t_k)} \cdot N(t_{k-1}) \cdot \exp\left\{\frac{x(t_{k-1}) + x(t_k)}{2} (t_k - t_{k-1})\right\}.$$

Alternatively, we can simulate in T-forward measure for a fixed future time T

Select a future time $\bar{T}$ sufficiently large. Then $N(0) = P(0, \bar{T})$.
At any pay time $T \leq \bar{T}$ numeraire is directly available via zero coupon bond formula

$$N(T) = P(x(T), T, \bar{T}) = \frac{P(0, \bar{T})}{P(0, T)} e^{-G(T,T')x(T) - \frac{1}{2}G(T,T')^2 y(T)}.$$

However, $\bar{T}$-forward measure simulation needs consistent model formulation or change of measure.
In particular

$$\underbrace{dW^{\bar{T}}(t)}_{\text{B.M. in } \bar{T}\text{-forward measure}} = \underbrace{\sigma_P(t, \bar{T})}_{\text{ZCB volatility}} \cdot dt + \underbrace{dW(t)}_{\text{B.M. in risk-neutral measure}}.$$

Another commonly used numeraire for simulation is the discretely compounded bank account

- Consider a grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.
- Assume we start with 1 EUR at $t = 0$, i.e. $N(0) = 1$.
- At each t_k we take numeraire $N(t_k)$ and buy zero coupon bond maturing at t_{k+1}. That is

$$N(t) = P(t, t_{k+1}) \cdot \frac{N(t_k)}{P(t_k, t_{k+1})} \quad \text{for} \quad t \in [t_k, t_{k+1}] .$$

Explicitly, define discretely compounded bank account as $\bar{B}(0) = 1$ and

$$\bar{B}(t) = P(t, t_{k+1}) \prod_{t_k < t} \frac{1}{P(t_k, t_{k+1})} .$$

We get

$$d\left(\frac{\bar{B}(t)}{P(t, t_{k+1})}\right) = \prod_{t_k < t} \frac{1}{P(t_k, t_{k+1})} \cdot d\left(\frac{P(t, t_{k+1})}{P(t, t_{k+1})}\right) = 0 \quad \text{for} \quad t \in [t_k, t_{k+1}] .$$

Simulating in $\bar{B}$-measure is equivalent to simulating in rolling t_{k+1}-forward measure.

Outline

Do we really need to solve the Hull-White SDE numerically?

Recall dynamics in T-forward measure

$$dx(t) = \left[y(t) - \sigma(t)^2 G(t,T) - a \cdot x(t)\right] \cdot dt + \sigma(t) \cdot dW^T(t).$$

That gives

$$x(T) = e^{-a(T-t)} \cdot \left[x(t) + \int_t^T e^{a(u-t)} \left(\left[y(u) - \sigma(u)^2 G(u,T)\right] du + \sigma(u) dW^T(u)\right)\right].$$

As a result $x(T) \sim N(\mu, \sigma^2)$ (conditional on t) with

$$\mu = \mathbb{E}^T \left[x(T) \,|\, \mathcal{F}_t\right] = G'(t,T)\left[x(t) + G(t,T)y(t)\right] \quad \text{and}$$

$$\sigma^2 = \text{Var}\left[x(T) \,|\, \mathcal{F}_t\right] = y(T) - G'(t,T)^2 y(t).$$

We can simulate exactly

$$x(T) = \mu + \sigma \cdot Z \quad \text{with} \quad Z \sim N(0,1).$$

Expectation calculation via $\mu = \mathbb{E}^T\left[x(T)\,|\,\mathcal{F}_t\right]$ requires carefull choice of numeraire

Consider grid of simulation times $t = t_0, t_1, \ldots, t_M = T$.
We simulate

$$x(t_{k+1}) = \mu_k + \sigma_k \cdot Z_k$$

with

$$\begin{aligned}
\mu_k &= G'(t_k, t_{k+1})\left[x(t_k) + G(t_k, t_{k+1})y(t_k)\right], \\
\sigma_k^2 &= y(t_{k+1}) - G'(t_k, t_{k+1})^2 y(t_k), \quad \text{and} \\
Z_k &\sim N(0,1).
\end{aligned}$$

Grid point t_{k+1} must coincide with forward measure for $\mathbb{E}^{t_{k+1}}[\cdot]$ for each individual step $k \to k+1$.
Numeraire must be discretely compounded bank account $\bar{B}(t)$ and

$$\bar{B}(t_{k+1}) = \frac{\bar{B}(t_k)}{P(x(t_k), t_k, t_{k+1})}.$$

Recursion for $x(t_{k+1})$ and $\bar{B}(t_{k+1})$ fully specifies path simulation for pricing.

Some comments regarding Hull-White MC simulation ...

- We could also simulate in risk-neutral measure or $\bar{T}$-forward measure.
  - This might be advantegous if also FX or equities are modelled/simulated.
  - Requires adjustment of conditional expectation μ_k and numeraire $N(t_k)$ calculation.
  - Variance σ_k^2 is invariant to change of meassure in Hull-White model.
- Repeat path generation for as many paths $1, \ldots, n$ as desired (or computationally feasible).
- For Bermudan pricing we need to simulate x and N (at least) at exercise dates $T_E^1, \ldots, T_E^{\bar{k}}$.
- For calculation of Z_k use
  - pseudo-random numbers or
  - Quasi-Monte Carlo sequences.

  as proxies forindependent $N(0,1)$ random variables accross time steps and paths.

We illustrate MC pricing by means of a coupon bond option example

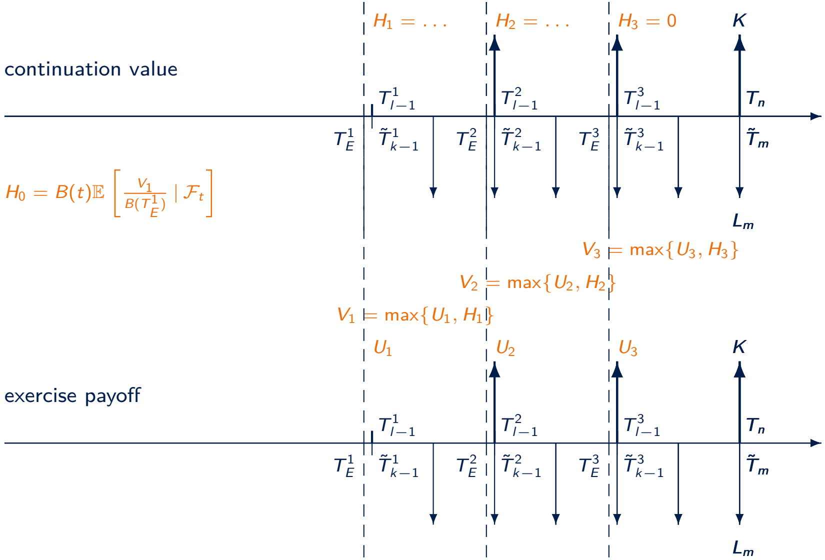
Consider coupon bond option expiring at T_E with coupons C_i paid at T_i ($i = 1, \ldots, u$, incl. strike and notional).

- Set $t_0 = 0$, $t_1 = T_E/2$ and $t_2 = T_E$ (two steps for illustrative purpose).
- Compute $2n$ independent $N(0,1)$ pseudo random numbers $Z^1, \ldots, Z^{2n}$.
- For all paths $j = 1, \ldots, n$ calculate:
  - μ_0^j, σ_0 and $\bar{B}^j(t_1)$; note μ_0^j and $\bar{B}^j(t_1)$ are equal for all paths j since $x(t_0) = 0$,
  - $x_1^j = \mu_0^j + \sigma_0 \cdot Z^j$,
  - μ_1^j, σ_1 and $\bar{B}^j(t_2)$; note now μ_1^j and $\bar{B}^j(t_2)$ depend on x_1^j,
  - $x_2^j = \mu_1^j + \sigma_1 \cdot Z^{n+j}$,
  - payoff $V^j(t_2) = \left[\sum_{i=1}^{u} C_i \cdot P(x_2^j, t_2, T_i)\right]^+$ at $t_2 = T_E$.
- Calculate option price (note $\bar{B}(0) = 1$)

$$V(0) = \bar{B}(0) \cdot \frac{1}{n} \sum_{j=1}^{n} \frac{V^j(t_2)}{\bar{B}^j(t_2)}.$$

Outline

Let's return to our Bermudan option pricing problem

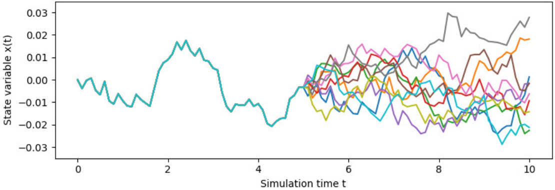


In this setting we need to calculate future conditional expectations

- Assume we already simulated paths for state variables x_k, underlyings U_k and numeraire B_k for all relevant dates t_k.
- We need continuation values H_k defined recursively via $H_{\bar{k}} = 0$ and

$$H_k = B_k \mathbb{E}_k \left[\frac{\max\left\{U_{k+1}, H_{k+1}\right\}}{B_{k+1}} \right].$$

- In principle, we could use nested Monte Carlo:



- In practice, nested Monte Carlo is typically computationally not feasible.

A key idea of American Monte Carlo is approximating conditional expectation via regression

Conditional expectation

$$H_k = \mathbb{E}_k \left[\frac{B_k}{B_{k+1}} \max \left\{ U_{k+1}, H_{k+1} \right\} \right]$$

is a function of the path $x(t)$ for $t \leq t_k$.

For non-path-dependent underlyings U_k, H_k can be written as function of $x_k = x(t_k)$, i.e.

$$H_k = H_k(x_k).$$

We aim at finding a regression operator

$$\mathcal{R}_k = \mathcal{R}_k [Y]$$

which we can use as proxy for H_k.

What do we mean by regression operator?

Denote $\zeta(\omega) = [\zeta_1(\omega), \ldots, \zeta_q(\omega)]^\top$ a set of basis functions (vector of random variables).

Let $Y = Y(\omega)$ be a target random variable.

Assume we have outcomes $\omega_1, \ldots, \omega_{\bar{n}}$ with control variables $\zeta(\omega_1), \ldots, \zeta(\omega_{\bar{n}})$ and observations $Y(\omega_1), \ldots, Y(\omega_{\bar{n}})$.

A regression operator $\mathcal{R}[Y]$ is defined via

$$\mathcal{R}[Y](\omega) = \zeta(\omega)^\top \beta$$

where the regression coefficients β solve linear least squares problem

$$\left\| \begin{bmatrix} \zeta(\omega_1)^\top \beta - Y(\omega_1) \\ \vdots \\ \zeta(\omega_{\bar{n}})^\top \beta - Y(\omega_{\bar{n}}) \end{bmatrix} \right\|^2 \to \min .$$

Linear least squares system can be solved e.g. via QR factorisation or SVD.

A basic pricing scheme is obtained by replacing conditional expectation of future payoff by regression operator

Approximate $\tilde{H}_k \approx H_k$ via $\tilde{H}_{\bar{k}} = H_{\bar{k}} = 0$ and

$$\tilde{H}_k = \mathcal{R}_k \left[\frac{B_k}{B_{k+1}} \max \left\{ U_{k+1}, \tilde{H}_{k+1} \right\} \right] \quad \text{for} \quad k = \bar{k} - 1, \ldots, 1.$$

- Critical piece of this methodology is (for each step k)
  - choice of regression variables $\zeta_1, \ldots, \zeta_q$ and
  - calibration of regression operator $\mathcal{R}_k$ with coefficients β.
- Regression variables $\zeta_1, \ldots, \zeta_q$ must be calculated based on information up to t_k.
  - They must not look into the future to avoid upward bias.
- Control variables $\zeta(\omega_1), \ldots, \zeta(\omega_{\bar{n}})$ and observations $Y(\omega_1), \ldots, Y(\omega_{\bar{n}})$ for calibration should be simulated on paths independent from pricing.
  - Using same paths for calibration and payoff simulation also incorporates information on the future.

What are typical basis functions?

State variable approach

Set $\zeta_i = x(t_k)^{i-1}$ for $i = 1, \ldots, q$. Typical choice is $q \approx 4$ (i.e. polynomials of order 3). For multi-dimensional models we would set $\zeta_i = \prod_{j=1}^{d} x_j(t_k)^{p_{i,j}}$ with $\sum_{j=1}^{d} p_{i,j} \leq r$.

- Very generic and easy to incorporate.

Explanatory variable approach

Identify variables $y_1, \ldots y_{\bar{d}}$ relevant for the underlying option. Set basis functions as monomials

$$\zeta_i = \prod_{j=1}^{\bar{d}} y_j(t_k)^{p_{i,j}} \quad \text{with} \quad \sum_{j=1}^{\bar{d}} p_{i,j} \leq r.$$

- Can be chosen option-specific and incorporate information prior to t_k.
- Typical choices are co-terminal swap rates or Libor rates (observed at t_k).

Regression of the full underlying can be a bit rough - we may restrict regression to exercise decision only

For a given path consider

$$H_k = \frac{B_k}{B_{k+1}} \max\left\{U_{k+1}, H_{k+1}\right\}$$
$$= \frac{B_k}{B_{k+1}} \left[\mathbb{1}_{\{U_{k+1}>H_{k+1}\}} U_{k+1} + \left(1 - \mathbb{1}_{\{U_{k+1}>H_{k+1}\}}\right) H_{k+1}\right].$$

Use regression to calculate $\mathbb{1}_{\{U_{k+1}>H_{k+1}\}}$.
Calculate $\mathcal{R}_{k+1} = \mathcal{R}_{k+1}\left[U_{k+1} - H_{k+1}\right]$, set $H_{\bar{k}} = 0$ and

$$H_k = \frac{B_k}{B_{k+1}} \left[\mathbb{1}_{\{\mathcal{R}_{k+1}>0\}} U_{k+1} + \left(1 - \mathbb{1}_{\{\mathcal{R}_{k+1}>0\}}\right) H_{k+1}\right] \quad \text{for} \quad k = \bar{k}-1, \ldots, 1.$$

- Think of $\mathbb{1}_{\{\mathcal{R}_{k+1}>0\}}$ as an exercise strategy (which might be sub-optimal).
- This approach is sometimes considered more accurate than regression on regression.
- For further reference, see also Longstaff/Schwartz (2001).

We summarise the American Monte Carlo method

1. Simulate n paths of state variables x_k^j, underlyings U_k^j and numeraires B_k^j $(j = 1, \ldots, n)$ for all relevant times t_k $(k = 1, \ldots \bar{k})$.
2. Set $H_{\bar{k}}^j = 0$.
3. For $k = \bar{k} - 1, \ldots 1$ iterate:
   3.1 Calculate control variables $\left\{\zeta_i^j = \zeta_i(\omega_j)\right\}_{i=1,\ldots,q}^{j=1,\ldots,\hat{n}}$ and regression variables $Y^j = U_k^j - H_k^j$ for the first $\hat{n}$ paths ($\hat{n} \approx \frac{1}{4} n$).
   3.2 Calibrate regression operator $\mathcal{R}_{k+1} = \mathcal{R}_{k+1}[Y]$ which gives coefficients β.
   3.3 Calculate control variables $\left\{\zeta_i^j = \zeta_i(\omega_j)\right\}_{i=1,\ldots,q}^{j=\hat{n}+1,\ldots n}$ for remaining paths and (for all paths)

$$H_k^j = \frac{B_k^j}{B_{k+1}^j} \left[\mathbb{1}_{\left\{\mathcal{R}_{k+1}(\omega_j) > 0\right\}} U_{k+1}^j + \left(1 - \mathbb{1}_{\left\{\mathcal{R}_{k+1}(\omega_j) > 0\right\}}\right) H_{k+1}^j \right].$$

4. Calculate discounted payoffs for the paths $j = \hat{n} + 1, \ldots n$ not used for regression

$$H_0^j = \frac{B_k^j}{B_{k+1}^j} \max\left\{ U_1^j, H_1^j \right\}.$$

5. Derive average $V(0) = \frac{1}{n - \hat{n}} \sum_{j=\hat{n}+1}^{n} H_0^j$.

Some comments regarding AMC for Bermudans in Hull-White model

- AMC implementations can be very bespoke and problem specific.
  - See literature for more details.
- More explanatory variables or too high polynomial degree for regression may deteriorate numerical solution.
  - This is particularly relevant for 1-factor models like Hull-White.
  - Single state variable or co-terminal swap rate should suffice.
- AMC with Hull-White for Bermudans is *not* the method of choice.
  - PDE and integration methods are directly applicable.
  - AMC is much slower and less accurate compared to PDE and integration.

AMC is the method of choice for high-dimensional models and/or path-dependent products.

Outline

References I

F. Ametrano and M. Bianchetti.
Everything you always wanted to know about Multiple Interest Rate Curve Bootstrapping but were afraid to ask (April 2, 2013).
Available at SSRN: http://ssrn.com/abstract=2219548 or http://dx.doi.org/10.2139/ssrn.2219548, 2013.

L. Andersen and V. Piterbarg.
Interest rate modelling, volume I to III.
Atlantic Financial Press, 2010.

D. Bang.
Local-stochastic volatility for vanilla modeling.
https://ssrn.com/abstract=3171877, 2018.

M. Beinker and H. Plank.
New volatility conventions in negative interest environment.
d-fine Whitepaper, available at www.d-fine.de, December 2012.

D. Brigo and F. Mercurio.
Interest Rate Models - Theory and Practice.
Springer-Verlag, 2007.

References II

D. Duffy.
Finite Difference Methods in Financial Engineering.
Wiley Finance, 2006.

M. Fujii, Y. Shimada, and A. Takahashi.
Collateral posting and choice of collateral currency - implications for derivative pricing and risk management (may 8, 2010).
Available at SSRN: https://ssrn.com/abstract=1601866, May 2010.

M. Giles and P. Glasserman.
Smoking adjoints: fast monte carlo greeks.
Risk, January 2006.

P. Glasserman.
Monte Carlo Methods in Financial Engineering.
Springer, 2003.

A. Griewank and A. Walther.
Evaluating derivatives: principles and techniques of algorithmic differentiation - 2nd ed.
SIAM, 2008.

References III

P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward.
Managing smile risk.
Wilmott magazine, September 2002.

P. Hagan and G. West.
Interpolation methods for curve construction.
Applied Mathematical Finance, 13(2):89–128, 2006.

M. Henrard.
Interest rate instruments and market conventions guide 2.0.
Open Gamma Quantitative Research, 2013.

M. Henrard.
A quant perspective on ibor fallback proposals.
https://ssrn.com/abstract=3226183, 2018.

M. Henrard.
A quant perspective on ibor fallback consultation results.
https://ssrn.com/abstract=3308766, 2019.

References IV

J. Hull and A. White.
Pricing interest-rate-derivative securities.
The Review of Financial Studies, 3:573–592, 1990.

Y. Iwashita.
Piecewise polynomial interpolations.
OpenGamma Quantitative Research, 2013.

A. Lyashenko and F. Mercurio.
Looking forward to backward-looking rates: A modeling framework for term rates replacing libor.
https://ssrn.com/abstract=3330240, 2019.

U. Naumann.
The Art of Differentiating Computer Programs: An Introduction to Algorithmic Differentiation.
SIAM, 2012.

V. Piterbarg.
Funding beyond discounting: collateral agreements and derivatives pricing.
Asia Risk, pages 97–102, February 2010.

References V

R. Rebonato.
Volatility and Correlation.
John Wiley & Sons, 2004.

S. Shreve.
Stochastic Calculus for Finance II - Continuous-Time Models.
Springer-Verlag, 2004.

Contact

Dr. Sebastian Schlenkrich
Office: RUD25, R 1.201
Mail: sebastian.schlenkrich@hu-berlin.de