Derivative Securities, Fall 2010
Mathematics in Finance Program
Courant Institute of Mathematical Sciences, NYU
Jonathan Goodman
`http://www.math.nyu.edu/faculty/goodman/teaching/DerivSec10/resources.html`

Week 11

1 Interest rate models and the Feynman Kac formula

Picking up at the end of the week 10 notes, I want to do the calculations of $E[Z]$ and $\text{var}(Z)$ in more detail. We find $E[Z]$ by integrating the formula below (19) there

$$E[Z] \;=\; \mu_Z(T) \;=\; -\int_0^T E\left[r_t\right]\, dt \;=\; -\overline{r}T + \frac{\overline{r}-r_0}{a}\left(1-e^{-aT}\right) \; . \tag{1}$$

When T is large, the exponential is small, which leads to the approximations

$$E[Z] \;\approx\; -\overline{r}T + \frac{\overline{r}-r_0}{a} \;\approx\; -\overline{r}T \; .$$

The variance of Z is the expected square of the random part of Z, which is

$$U \;=\; \sigma \int_{t=0}^{T} \int_{s=0}^{t} e^{-a(t-s)}\, dW_s\, dt \; .$$

We evaluate $E\left[U^2\right]$ by writing U as a single integral involving dW_s and then using the Ito isometry formula. For this, we change the order of integration to put the dt integral on the inside. And for that, define $H(s,t)=1$ if $s \leq t$ and $H(s,t)=0$ if $s > t$. Then

$$U \;=\; \sigma \int_{t=0}^{T} \int_{s=0}^{T} H(s,t) e^{-a(t-s)}\, dW_s\, dt \; .$$

It is easy to check that you can change the order of integration in situations like this. As long as $f(s,t)$ is a reasonable function,

$$\int_{t=0}^{T} \left(\int_{s=0}^{T} f(s,t)\, dW_s \right) dt \;=\; \int_{s=0}^{T} \left(\int_{t=0}^{T} f(s,t)\, dt \right) dW_s \; .$$

You can see this, for example, by approximating both integrals by Riemann sums, changing the order of summation, and taking the limits $\delta s \to 0$ and $\delta t \to 0$. In our case, we get

$$U \;=\; \sigma \int_{s=0}^{T} \left(\int_{t=s}^{T} e^{-a(t-s)}\, dt \right) dW_s \;=\; \frac{\sigma}{a} \int_{s=0}^{T} \left(1 - e^{-a(T-s)} \right) dW_s \; . \tag{2}$$

The Ito isometry formula then gives

$$E\left[U^2\right] \;=\; \frac{\sigma^2}{a^2}\int_0^T \left(1-2e^{-at}+e^{-2at}\right)\,dt$$

$$\text{var}(Z) \;=\; \sigma_Z^2(T) \;=\; \frac{\sigma^2}{a^2}\left\{T \;-\; \frac{2}{a}\left(1-e^{-aT}\right) \;+\; \frac{1}{2a}\left(1-e^{-2aT}\right)\right\}\;. \tag{3}$$

For large T, this is approximately

$$\text{var}(Z) \;\approx\; \frac{\sigma^2 T}{a^2}\;.$$

Thus, the large T approximation to $B(0,T)$ is

$$B(0,T) \;=\; e^{\mu_Z(T)+\sigma_Z(T)^2/2} \;\approx\; e^{-\left(\overline{r}\,-\,\sigma^2/2a^2\right)T}\;. \tag{4}$$

This implies that the long time approximation to the yield is $Y_T \approx \overline{r} - \sigma^2/a^2$.

Note that in the exact formula (1) $\mu_Z(T)$ depends linearly on r_0. In the exact formula (3), the variance is independent of r_0. Therefore, if we use the exact expressions in (4), the result is a formula of the form

$$B(r_0,T) \;=\; e^{-C(T)\,+\,D(T)r_0}\;. \tag{5}$$

For this reason,[1] the bond pricing formula is called an *affine* model. We have just seen that the Vasicek short rate model leads to an affine pricing formula. The formulas for $C(T)$ and $D(T)$ can be read off from (1) and (3). In particular

$$D(T) \;=\; -\,\frac{1-e^{-aT}}{a}\;. \tag{6}$$

The formula for $C(T)$ is more complicated, but is easy to figure out and program.

The backward equation gives a different way to find these and other solutions. The one we need is equation (19) from week 7. We have a quantity of the form (18) from week 7, with $x=r$, $f=B$ (the bond price), and $V=-r$. As I said then, the PDE (19) often is called the Feynman Kac formula, but really (18) is the Feynman Kac formula for the solution of the PDE (19). To be more precise, $f(x,t)$ is the expected value starting at time t and ending at time T. Let us write $B(r,t,T)$ for the price of a zero coupon bond starting at time t and maturing at time T. That price is unknown today (time $t=0$) but will be known at time t, and at that time will be a function of the short rate, r_t. Adapting (19) from week 7 with this notation, the bond price satisfies the PDE

$$\partial_t B(r,t,T) \;+\; \frac{\sigma^2}{2}\partial_r^2 B(r,t,T) \;+\; a\left(\overline{r}\;-\;r\right)\partial_r B(r,t,T) \;-\; rB(r,t,T)\;. \tag{7}$$

[1] A function of the form $f(x)=Ax$ is called *linear*. A function $f(x)=Ax+b$ is called *affine* if $b\neq 0$. This may contradict what you learned when you were little, that the equation of a line $y=ax+b$ is a linear equation.

Here, the maturity time T is just a parameter. The PDE is in the variables r and t. I write $B(r,t,T)$ to emphasize that the time variable of the PDE is not the maturity time of the bond, but the time when the bond is purchased.

Here is a way to check that the signs are right in (7). Suppose r is a known constant, which means that $\sigma = 0$ and $a = 0$. Then it becomes just $\partial_t B - rB = 0$. The solution is $B(t,T) = C_T e^{rt}$. Recall that T is just a parameter. The constant in the solution can depend on this parameter. We determine C_T from the final condition at time $t = T$, which is $B(T,T) = 1$. This gives $C_T e^{rT} = 1$, or $C_T = e^{-rT}$, which leads to the comfortingly familiar $B(t,T) = e^{-r(T-t)}$. The the equation also contains $LB = \frac{\sigma^2}{2}\partial_r^2 B(r,t,T) + a\,(\overline{r} \,-\, r)\,\partial_r B(r,t,T)$. This is the generator of the stochastic process (18) from week 10.

The PDE (7) helps us find the exponential affine solutions. This is done using the *ansatz*[2]

$$B(r,t,T) \;=\; e^{-C(t,T)\,+\,D(t,T)r}\;. \tag{8}$$

An ansatz is a suggested functional form of the solution. The ansatz method for solving equations is to guess (or be led to somehow[3]) the functional form of the solution and then calculate to see whether a solution of that form is possible. This relies on the *uniqueness* of the solution – there is only one solution. If you have a solution that satisfies all the conditions (the PDE, boundary conditions, final conditions, etc.), then you have *the* solution.

In the present case, we substitute the formula (8) into the PDE (7). We use the notation $\dot{C} = \partial_t C(t,T)$, and the same for D. We have

$$\partial_t B \;=\; -\dot{C}e^{-C(t,T)\,+\,D(t,T)r} \;+\; r\dot{D}e^{-C(t,T)\,+\,D(t,T)r} \;=\; \left(-\dot{C} + r\dot{D}\right)B\;,$$

and

$$\partial_r B \;=\; DB\;\;,\quad \partial_r^2 B = D^2 B\;.$$

When you substitute all these derivatives into (7), there is a common factor B in every term. I write the result after that factor has been removed:

$$-\dot{C} + r\dot{D} \;+\; \frac{\sigma^2}{2}D^2 \;+\; a\,(\overline{r} - r)\,D - r \;=\; 0\;.$$

Recall that in this formula, σ, a, and $\overline{r}$ are constants of the model, while r is a variable. If this is to hold for every r, the coefficients of r and the terms independent of r must vanish separately. If we rewrite the equation as

$$r\left(\dot{D} - aD - 1\right) \;+\; \left(-\dot{C} + \frac{\sigma^2}{2}D^2 + a\overline{r}D\right) \;=\; 0$$

[2]The German word *ansatz* (written *Ansatz* in German) does not seem to have any other meaning than its technical one, except that, like *exponential*, it is leaking back to the non-technical language.

[3]A mathematicians is like a donkey. A donkey does not move, even to find water, unless pulled. A mathematician would never find the solution of an equation without being pulled to it in some way - by analogy, by studying special cases, etc.

That leads to the two equations

$$\dot{D} - aD - 1 \;=\; 0 \;, \tag{9}$$

and

$$-\dot{C} + \frac{\sigma^2}{2} D^2 + a\overline{r} D \;=\; 0 \;. \tag{10}$$

The final condition for B, which is $B(r, T, T) = 1$, leads to (see (8)) final conditions $C(T, T) = 0$ and $D(T, T) = 0$.

The solutions of these equations is straightforward if a little tedious. Even if we were unable to solve them analytically, it would be easy to find numerical solutions to the ordinary differential equations with final conditions, the Scientific Computing class will tell you how. Starting with (9), our vast experience with ordinary differential equations (see any undergraduate ODE book) tells us to look for a solution of the form $D(t, T) = d_1(T)e^{at} + d_2(T)$. The ODE (9) tells us that $ad_2 = -1$, so $d_2 = \frac{-1}{a}$. The final condition then gives $0 = d_1(T)e^{rT} - \frac{1}{a}$, which gives $d_1(T) = \frac{1}{a}e^{-rT}$ and

$$D(t, T) \;=\; \frac{e^{-a(T-t)} \;-\; 1}{a} \;. \tag{11}$$

Setting $t = 0$, this agrees with the earlier (6), which is a check on some nontrivial algebra.

With (11) in hand, one can find solve (10) by squaring (11) and integrating each of the five terms (three from D^2 and two from D). However, the leading terms for large T and small t are from the $\frac{-1}{a}$ part of D. The $\frac{\sigma^2}{2}D^2$ turns this into $\frac{\sigma^2}{2a^2}$, and the $a\overline{r}D$ turns this into just $\overline{r}$. This gives the leading order long term behavior (4). The homework will examine the various shapes of the yield curve that are possible in the Vasicek model.

The CIR model, (21) from week 10, also has exponential affine solutions that may be found by the ansatz method. The backward equation is

$$\partial_t B \;+\; \frac{\sigma^2 r}{2} \;+\; a\,(\overline{r} - r)\,\partial_r B \;-\; rB \;=\; 0 \;. \tag{12}$$

The exponential affine ansatz (8) leads for this model to

$$-\dot{C} \;+\; r\dot{D} \;+\; \frac{\sigma^2 r}{2} D^2 \;+\; a\,(\overline{r} - r)\,D \;-\; r \;=\; 0 \;.$$

The differential equations for C and D now are

$$\dot{D} \;+\; \frac{\sigma^2}{2} D^2 \;-\; aD \;-\; 1 \;=\; 0 \;, \tag{13}$$

and

$$-\dot{C} \;+\; a\overline{r}D \;=\; 0 \;. \tag{14}$$

It is clear in retrospect that the factor $\sqrt{r}$ in the noise term of the CIR model (19) from week 10 probably was chosen to make explicit ansatz solutions such as these possible.

Even if we were not able to find explicit solution formulas for the backward equations (7) and (12), we would be able to find the solutions numerically using the forward Euler method. So you could use a model with a different power r_t^p instead of $r_t^{1/2}$. Instead of hours of painful algebra, you would have hours of painful programming, changing the coefficients in an old homework assignment code. Models with explicit solutions are useful, but models without explicit solutions also can be useful, and more realistic.

2 Affine exponential solutions and option pricing

The exponential affine formulas (8) make possible many things besides explicit formulas for bond prices. One is pricing interest rate contracts and options. A related application is "explicit" solutions to backward equations with exponential affine solutions using the Fourier transform. Both of these, like the ansatz calculations above, are simple in principle but time consuming in practice.

The most straightforward in principle is a simple loan option. A European style call on a loan, for example, would give the holder the right to receive K at time $T < T_1$ to pay one unit at time T_1. Let $B(T,T_1)$ be the price of the risk free zero coupon bond maturing at T_1, then if $B(T,T_1) < K$, we make a profit of $K - B(T,T_1)$ by buying the bond and exercising the option. Exercising the option gives us K. We use B of this to buy the bond and keep the remaining $K - B$ as profit. At time T, we get one unit from the bond and use this to repay the loan. The result is that the loan option has a payout at time T_1 equal to $V = (K - B(T,T_1))_+$. Because of our formula $B(T,T_1) = e^{-C(T,T_1)+rD(T,T_1)}$, we can get the price today of the payout at time T as

$$f(r,0) \;=\; E\left[\frac{(K - B(r_T,T,T_1))_+}{M_T}\right] \; . \tag{15}$$

In the Vasicek model, this can be evaluated in closed form by a method similar to that of problem 2 of assignment 10. From the solution of the Vasicek SDE (19) of week 10 we have that r_t is normal with mean $\overline{r} + e^{-aT}\left(r - \overline{r}\right)$. The random part of r_t is the Gaussian

$$X \;=\; \sigma \int_0^T e^{-a(T-t)}\, dW_t \; .$$

Similarly, M_T has the form $e^{-\mu(T)+U}$, where $\mu(t)$ is not random and has an explicit formula, and U is is another Gaussian given by (2). Now, because X and U are jointly normal, we can write

$$U \;=\; \alpha X \,+\, Y \; ,$$

where Y is normal, independent of X, with mean zero and variance $\text{var}(Y) = \sigma_{YY}$. To compute the coefficients, we need $\sigma_{XX} = \text{var}(X)$, $\sigma_{UU} = \text{var}(U)$, and $\sigma_{XU} = \text{cov}(X,U)$. First, $\sigma_{UX} = E[UX] = E[(\alpha X + Y)(X)] = \alpha\sigma_{XX}$, which gives $\alpha = \sigma_{XU}/\sigma_{XX}$. Then $\sigma_{UU} = E\left[(\alpha X + Y)^2\right] = \alpha^2\sigma_{XX} + \sigma_{YY}$, which gives $\sigma_{YY} = \sigma_{UU} - \alpha^2 = \sigma_{UU} - \sigma_{XU}^2/\sigma_{XX}$. With all this, (15) takes the form

$$f \;=\; E\left[\frac{\left(K - e^{-C+[X+\overline{r}+e^{-aT}(r-\overline{r})]D}\right)_+}{e^Y e^{-\mu(T)} e^{\alpha X}}\right] .$$

The reason for separating U into X and Y parts is now clear. Since Y is independent of X, we can take the expectation over Y out of the expectation over X, which gives

$$f \;=\; e^{\mu(T)} e^{\sigma_{YY}/2} E\left[\left(K^{-\alpha X} - e^{-C+e^{-aT}(r-\overline{r})D} e^{(D-\alpha)X}\right)_+\right] .$$

The last expectation is the kind of thing we can express in terms of two N functions as in the Black Scholes formula. It is not pretty, but it can be done. In some sense, this justifies the otherwise heuristic Black pricing model.

Other major interest rate contracts are *interest rate caps* and *swaptions*. A *floating rate* loan is one in which the borrower promises to pay $(r_t + s)dt$ times the notional in each time interval dt until maturity. At maturity, the borrower pays the notional and the loan is over. In practice, the time interval dt could be a month or a quarter, and r_t would be the LIBOR rate for that period stated at the beginning of the period. As usual, s is a fixed credit spread that depends on the borrower. A *swap* is the exchange of a floating rate loan for a stream of predetermined coupon payments. A *swaption* is the option to enter into a predetermined interest rate swap. Most swaptions are *Bermudian style*, which means that they can be exercised on any of a predetermined set of dates. This makes them intermediate between European (exercised on a single date) and American (exercised at any time), just as Bermuda is between Europe and America.

An interest rate cap is simply an agreement that the payments on a floating rate loan will not exceed a fixed cap rate K. The cap may be viewed and priced as a collection of *caplets*, one for each loan payment. A caplet is the right to pay Kdt instead of $(r_t + s)dt$ for time interval $(t, t+dt)$. Each caplet is a loan option of the kind discussed above. There also are interest rate *floors*, that protect the lender from interest rates going too low. Of course, a floor is a collection of floorlets (not a commonly used term), each of which is a loan option.

A final application of the exponential ansatz method is the "pricing" of contracts with exponential payouts. Suppose an option pays $e^{\lambda r_T}$ at time T. The ansatz method clearly can do this, it only needs different final conditions $D(T,T) = \lambda$ instead of $D(T,T) = 0$. Now, if you have an analytic formula for the resulting $f(r,t,T,\lambda)$, you can in particular apply it when $\lambda = ip$. Then if $V(r_T)$ is any payout function, you can express the resulting value function using the Fourier transform of $V(r)$ using $f(r,t,T,ip)$. In practice, this is done numerically using the FFT.

3 No arbitrage models

The simplest form of a no arbitrage model that makes sense is the *Hull and White* one factor model:

$$dr_t \;=\; a\left(\bar(r)(t) - r)t\right) dt \;+\; \sigma dW_t\;.$$

The function $\overline{r}(t)$ can be chosen so that the yield curve predicted by the model is equal to a desired (e.g. the current) yield curve.