Continuous Time Finance Notes, Spring 2004 – Section 6, *March 3, 2004*
Notes by Robert V. Kohn, Courant Institute of Mathematical Sciences. For use in connection with the NYU course Continuous Time Finance.

The first bit of this section addresses the pricing of swaptions – something we didn't get around to before. The rest is a supplement to Hull's discussion of the trinomial tree version of Hull-White.

****

Pricing swaptions using Hull-White. We explained in Section 4 how a caplet is equivalent to a put option on a zero coupon bond. A similar argument shows that a floorlet is equivalent to a call option on a zero coupon bond. So we can easily derive formulas for the prices of caplets and floorlets using Black's formula (problem 1 of HW 3 covers the case of a caplet). Caps and floors are simply portfolios of caplets and floorlets, so we've priced them too. But what about swaptions?

The first observation is general: the task of pricing a swaption is identical to that of pricing a suitable option on a coupon bond. To be specific, let's suppose the underlying swap exchanges the floating rate for fixed rate k, the interest payments being at times $T_1, \ldots, T_N$ with a return of principal at T_N. (The holder of the swap receives the fixed rate and pays the floating rate.) Consider the associated swaption, which gives the holder the right to enter into this swap at time T_0. For simplicity assume the time intervals $T_j - T_{j-1}$ are all the same length Δt. The value of the underlying swap at time T_0 is then

$$P(T_0, T_N) + k\Delta t \sum_{j=1}^{N} P(T_0, T_j) - 1$$

times the notional principal. Indeed, the first term is the value at time T_0 of the principal payment at T_N; the second term is the value at time T_0 of the coupon payments; and the third term is the value at time T_0 of a (short position in a) bond which pays the floating rate. Therefore the payoff of the swaption at time T_0 is

$$(P(T_0, T_N) + k\Delta t \sum_{j=1}^{N} P(T_0, T_j) - 1)_+ \ .$$

This is identical to the payoff of a call option on a coupon bond (with interest rate k and payments at times T_j) with strike 1.

The next observation is special to Hull-White (well, it's a bit more general than that: the argument works for any one-factor short-rate model). We claim that a call or put on a coupon bond is equivalent to a suitable portfolio of calls or puts options on zero coupon bonds. To explain why, let's focus on the case of a call. Recall that $P(t, T) = A(t, T) \exp[-B(t, T)r(t)]$. The key point is that $P(t, T)$ is really a function of three variables:

t, T, and $r(t)$, and it is *monotone* in the third argument $r(t)$. So there is a unique "critical value" r_* for the short rate at time T_0 such that an option with payoff

$$X = \left(P(T_0, T_N) + k\Delta t \sum_{j=1}^{N} P(T_0, T_j) - K \right)_+$$

is in the money (at time T_0) precisely if $r(T_0) < r_*$. Moreover we can write X as a sum of call payoffs on zero-coupon bonds,

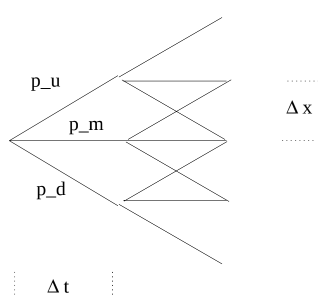
$$X = \left(P(T_0, T_N) - A(T_0, T_N)e^{-B(T_0,T_N)r_*}\right)_+ + k\Delta t \sum_{j=1}^{N} \left(P(T_0, T_j) - A(T_0, T_j)e^{-B(T_0,T_j)r_*}\right)_+$$

since each term is in the money at time T_0 exactly if $r(T_0) < r_*$. Option prices are additive – the value of a portfolio is the sum of the values of its component instruments – and we know how to price options on zero-coupon bonds. So we also know how to price swaptions.

****

The trinomial tree approximation to an additive random walk. You may have been exposed to trinomial trees as a scheme for pricing equity-based options. The setup is in some ways similar to the more familiar binomial trees, except we are not using the trinomial tree to *find* the risk-neutral dynamics; rather we are simply using it to *simulate* the risk-neutral dynamics – or, equivalently, to evaluate expected final-time payoffs by working backward through the tree. (One advantage of the trinomial tree: it amounts to an explicit finite-difference method for the backward Kolmogorov PDE. Second advantage: it lets us be sure there's always a node at a specified value of the stock price; this is important e.g. for valuing barrier options.)

To see the main issues, let's consider the trinomial tree approximation to the scaled Brownian motion process $dx = \sigma\, dw$ with $x(0) = 0$. The tree is determined once we choose Δx, Δt, and the probabilities p_u, p_m, p_d of going up, staying constant, or going down at each node (see the figure). To understand what choices are appropriate, recall that we identify



the continuous-time limit of a (binomial or trinomial) tree by applying the central limit theorem. Therefore the statistics of the continuous-time limit depend only on the means and variances of the one-period increments. Thus to get the continuous-time limit $dx = \sigma\, dw$ we require

$$p_u = p_d$$

so the one-period increment has mean value 0, and

$$p_u(\Delta x)^2 + p_d(\Delta x)^2 = \sigma^2 \Delta t$$

so the one-period increment has variance $\sigma^2 \delta t$. Of course we also require the probabilities all to be positive and

$$p_u + p_m + p_d = 1.$$

Since $p_u = p_d$ let's simplify the notation by calling it p. The condition for the variance requires

$$p = \frac{\sigma^2}{2} \frac{\Delta t}{(\Delta x)^2}$$

and positivity of $p_m = 1 - 2p$ requires

$$\sigma^2 \Delta t/(\Delta x)^2 < 1.$$

Notice that we have a range of possible choices: any p between 0 and 1/2 is OK.

To value an option we work backward in the tree. This is equivalent to finding the expected final-time payoff $E[f(x(T))]$ for the stochastic process associated with the tree. We claim this is equivalent to a standard (explicit) finite-difference approximation of the backward Kolmogorov equation – which in the present setting is $u_t + \frac{1}{2}\sigma^2 u_{xx} = 0$. Indeed, when we work backward in the tree, we determine $u(x,t)$ at time t and nodal value x by

$$u(x,t) = pu(x + \Delta x, t + \Delta t) + pu(x - \Delta x, t + \Delta t) + (1 - 2p)u(x, t + \Delta t),$$

which amounts after some reorganization to

$$\frac{u(x, t + \Delta t) - u(x,t)}{\Delta t} + \tfrac{1}{2}\sigma^2 \frac{u(x + \Delta x, t + \Delta t) + u(x - \Delta x, t + \Delta t) - 2u(x, t + \Delta t)}{(\Delta x)^2} = 0.$$

We required above that the probabilities in the tree be positive. This is natural, since they are probabilities. But is it also necessary? Yes indeed! This is the condition that the numerical scheme be *stable*. Indeed, if the probabilities are positive then it's easy to see that an option with small final-time payoff has small value. (More: we have the "maximum principle" $E[|f(x_T)|] \leq \max |f|$.) If the probabilities go negative then we lose this property, and working backward in the tree becomes wildly unstable.

Hull likes to take $p = 1/6$ (in other words $\sigma^2 \Delta t = (1/3)(\Delta x)^2$). The reason is that with this choice, the trinomial tree is an unexpectedly good approximation to the PDE (in this constant-volatility, zero-drift setting). To see why, let $u(x,t)$ be the solution of the PDE (with specified final-time value $u(x,T) = f$). By Taylor expansion about the point $(x, t+\Delta t)$ we have

$$u(x, t + \Delta t) - u(x,t) = u_t(x, t + \Delta t)\Delta t - \tfrac{1}{2}u_{tt}(x, t + \Delta t)(\Delta t)^2 + \ldots$$

and

$$u(x+\Delta x, t+\Delta t)+u(x-\Delta x, t+\Delta t)-2u_x(x, t+\Delta t) = u_{xx}(x, t+\Delta t)(\Delta x)^2+\frac{1}{12}u_{xxxx}(x, t+\Delta t)(\Delta x)^4+\ldots$$

Therefore the error terms in our trinomial tree (finite difference) numerical scheme are, at leading order,

$$-(1/2)u_{tt}\Delta t + \frac{\sigma^2}{24}u_{xxxx}(\Delta x)^2.$$

But differentiating the PDE we have $u_{tt} = (1/4)\sigma^4 u_{xxxx}$ so the preceding expression is equal to u_{xxxx} times

$$-\frac{\sigma^4}{8}\Delta t + \frac{\sigma^2}{24}(\Delta x)^2.$$

The special choice $\sigma^2\Delta t = (1/3)(\Delta x)^2$ makes this vanish. For this special choice of parameters, the leading order error is associated with the *next* even term of the Taylor expansion, i.e. it is of order $(\Delta x)^4 \sim (\Delta t)^2$ rather than of order $(\Delta x)^2 \sim \Delta t$.

In the trinomial tree version of Hull-White we have a nonzero drift. I doubt that $p = 1/6$ is any longer so special, because the preceding argument should be disturbed by the drift. But anyway $p = 1/6$ is no worse than any other choice.

****

Why, when doing Hull-White, must we do something different when x is very large or very small? We focus now on the trinomial tree approximation to

$$dx = -ax\,dt + \sigma\,dw$$

with initial data $x(0) = 0$. Suppose we use the standard branching, i.e. the process goes up, stays the same, or goes down with probabilities p_u, p_m, p_d respectively. The condition that the mean be right is now

$$p_u\Delta x - p_d\delta x = -ax\Delta t \tag{1}$$

and the condition that the variance be right is

$$p_u(\Delta x)^2 + p_d(\Delta x)^2 = a^2x^2(\Delta t)^2 + \sigma^2\Delta t. \tag{2}$$

These equations determine p_u and p_d; they in turn determine $p_m = 1 - p_u - p_d$.

Working backward in this tree gives, as before, a finite-difference numerical scheme for solving the backward Kolmogorov PDE, which is now $u_t - axu_x + \frac{1}{2}\sigma^2 u_{xx} = 0$. Indeed, when working backward we determine $u(x,t)$ by the formula

$$u(x,t) = p_u u(x+\Delta x, t+\Delta t) + p_m u(x, t+\Delta t) + p_d u(x-\Delta x, t+\Delta t).$$

Since $p_m = 1 - p_u - p_d$ this amounts to

$$[u(x, t+\Delta t)-u(x,t)]+p_u[u(x+\Delta x, t+\delta t)-u(x, t+\Delta t)]+p_d[u(x-\Delta x, t+\delta t)-u(x, t+\Delta t)] = 0.$$

Writing δx for the increment of x (taking values Δx with probability p_u, $-\Delta x$ with probability p_d, and 0 with probability p_m), and using the Taylor expansions

$$u(x+\delta x, t+\Delta t) = u(x) + u_x \delta x + \tfrac{1}{2} u_{xx} (\delta x)^2 + \ldots, \quad u(x, t+\Delta t) = u(x,t) + u_t \Delta t + \ldots$$

we see that

$$u_t \Delta t + u_x E[\delta x] + \tfrac{1}{2} u_{xx} E[(\delta x)^2] = \text{higher order terms.}$$

Thus by getting the mean and variance of δx right, we are designing a finite-difference scheme for the backward Kolmogorov PDE.

Let's make (1) and (2) more explicit. Following Hull, we fix $\Delta x = \sigma\sqrt{3\Delta t}$. After N timesteps there are $2N+1$ spatial nodes, at $x = 0, \pm\Delta x, \ldots, \pm N\Delta x$; therefore we may set $x = j\Delta x$ with $-N \le j \le N$. A bit of algebra gives

$$p_u = \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 - aj\Delta t}{2}, \quad p_d = \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 + aj\Delta t}{2}, \quad p_m = \frac{2}{3} - a^2 j^2 (\Delta t)^2.$$

These probabilities must be *positive*. Writing $\xi = aj\Delta t$, the conditions for positivity are

$$\frac{1}{3} + \xi^2 - \xi > 0, \quad \frac{1}{3} + \xi^2 + \xi > 0, \quad \xi^2 < \frac{2}{3}.$$

The first two inequalities impose no conditions (the polynomials $1/3 + \xi^2 \pm \xi = 0$ have no real roots) but the last one requires

$$aj\Delta t < \sqrt{2/3} \approx .816.$$

We are in trouble after N timesteps, where $aN\delta \approx .816$.

****

Why does it work to truncate the tree? Fixing this problem is surprisingly easy. We truncate the tree at suitably chosen values of j, and use a different trinomial branching scheme at the top and bottom spatial nodes (see the figure, which however truncates the tree much earlier than would normally be done). To see that this works, consider the situation at the top spatial node $j_{\max}$. The branching scheme there (Hull's figure 21.7c) gets the mean and variance right when

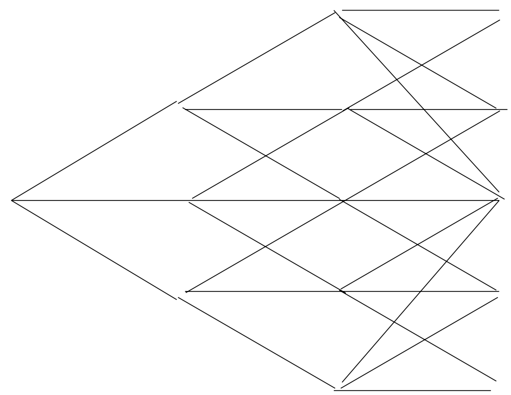
$$(-p_m - 2p_d)\Delta x = -ax\Delta t, \quad p_m (\Delta x)^2 + p_d (2\Delta x)^2 = a^2 x^2 (\Delta t)^2 + \sigma^2 \Delta t$$

which gives (using $\Delta x = \sigma\sqrt{3\Delta t}$ as before, and writing $x = j\Delta x$)

$$p_u = \frac{7}{6} + \frac{a^2 j^2 (\Delta t)^2 - 3aj\Delta t}{2}, \quad p_d = \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 - aj\Delta t}{2}, \quad p_m = -\frac{1}{3} - a^2 j^2 (\Delta t)^2 + 2aj\Delta t.$$

For these to be positive we need $\xi = aj\Delta t$ to satisfy

$$\frac{7}{3} + \xi^2 - 3\xi > 0, \quad \frac{1}{3} + \xi^2 - \xi > 0, \quad -\frac{1}{3} - \xi^2 + 2\xi > 0.$$



The first two conditions are always satisfied (the polynomials have no real roots), and the last is satisfied provided ξ lies between $1-\sqrt{2/3} = .184$ and $1+\sqrt{2/3} = 1.816$. It is natural to truncate the tree as early as possible (this minimizes the number of nodes). Therefore we should choose $j_{\max}$ so that

$$a j_{\max} \Delta t \text{ is slightly larger than } .184.$$

The situation at the bottom of the tree is symmetric, so we need not discuss it separately.

****

How do we calibrate the tree version of Hull-White to a given initial term structure? Recall that in continuous time, the solution of

$$dr = (\theta(t) - ar)\,dt + \sigma\,dw, \quad r(0) = r_0$$

was expressed as a sum of two terms: $r = \alpha(t) + x(t)$, where

$$dx = -ax\,dt + \sigma\,dw, \quad x(0) = 0$$

is independent of the initial term structure, and

$$\alpha(t) = f(0,t) + \frac{\sigma^2}{2a^2}\left(1 - e^{-at}\right)^2.$$

This decomposition suggests one way of proceeding: we may model $r(t)$ by the trinomial tree obtained by adding $\alpha(t_j)$ to each nodal value of x at time $t_j = j\Delta t$. This is reasonable – but at finite Δt it only gets the initial term structure *approximately* correct.

A more popular alternative is to choose the interest-rate tree so that it gives exactly the observed values for $P(0, j\Delta t)$ for each $j = 1, 2, 3, \ldots$. We do this by adding well-chosen values $\tilde{\alpha}_j$ to the values of x at times $j\delta t$, $j = 0, 2, 3, \ldots$. The values $\tilde{\alpha}_j$ can be chosen sequentially: choosing $\tilde{\alpha}_0$ to be the initial short rate r_0 assures that the tree predicts $P(0, \Delta t) = e^{-r_0 \Delta t}$ as desired; there is a unique value of $\tilde{\alpha}_1$ such that the tree correctly reproduces $P(0, 2\Delta t)$; etc. (This procedure is spelled out in more detail in Hull's book.)