1. Itô's Lemma

Itô's lemma is an indispensible tool for working with continuous time random processes. This note informally 'derives' it using Taylor series approximations. First, what does Itô's lemma say?

Suppose that some variable y is a function $f(s,t)$ of s and t, and that s follows a continuous random process that can be described by

$$ds(t) = \alpha\, dt + \sigma\, dz(t)$$

Here α is the instantaneous expected rate of change in s and σ is its instantaneous standard deviation. Then y will also follow a random process — one induced by the randomness of s. Itô's lemma relates the characteristics of the y process to the s process. Specifically, y will follow

$$dy(t) = (\alpha f_s + \tfrac{1}{2}\sigma^2 f_{ss} + f_t)\, dt + \sigma f_s\, dz(t)$$

where subscripts denote partial derivatives. What stands out is that the expected rate of change in y is not simply the sum of its rate of change due to the passage of time, f_t, and the expected rate of change in s times y's sensitivity to s, αf_s, but also has a term involving the volatility of s and the second derivative of f.

To see why this is so, suppose that we are initially at some s, t and that a short interval of time Δt passes. During this time there will be some associated Δz. Expanding f in a Taylor series around its starting value,

$$\begin{aligned} f(s+\alpha\Delta t+\sigma\Delta z, t+\Delta t) = \\ f(s,t) + [\alpha\Delta t + \sigma\Delta z]\, f_s + \Delta t\, f_t + \tfrac{1}{2}[\alpha^2\Delta t^2 + 2\alpha\sigma\Delta t\Delta z \\ + \sigma^2\Delta z^2]\, f_{ss} + [\alpha\Delta t^2 + \sigma\Delta z\Delta t]\, f_{st} + \tfrac{1}{2}\Delta t^2\, f_{tt} \\ + \text{ third and higher order terms} \end{aligned}$$

Now subtract $f(s,t)$ from both sides to get an expression for Δf and take its expected value. Since z is following a standard Brownian motion, Δz is normally distributed with expected value 0 and variance Δt. I.e., $\mathrm{E}[\Delta z^2] = \Delta t$.

$$\begin{aligned} \mathrm{E}[\Delta f] &= [\tfrac{1}{2}\sigma^2 f_{ss} + \alpha f_s + f_t]\,\Delta t \\ &\quad + \text{ second and higher order terms in } \Delta t \end{aligned}$$

For small Δt we can ignore the higher order terms, giving us the expected rate of change in f corresponding to Itô's lemma above. Finally, subtract this expected Δf from Δf itself to get the random part of the change in f:

$$\Delta f - \mathrm{E}[\Delta f] = \sigma f_s\, \Delta z + \text{ second and higher order terms in } \Delta t \text{ and } \Delta z$$

This gives us the last part of the Itô's lemma expression.

The main point of all this is that the Jensen's inequality effect, embodied in the second derivative term, does not fade away as Δt becomes small. This is a consequence of $z(t)$ almost everywhere following an extremely "jiggly" path. The parameters α and σ can be functions of t and s. A formally correct proof can be found in Malliaris and Brock (1982) and the works that they cite.

An example

In a world with a constant nominal interest rate r, a bond portfolio with value of \$1 at time 0 and continuously reinvested coupon payments will be worth $B(t) = e^{rt}$ at time t. Suppose that the price level evolves randomly according to the stochastic process

$$dP = \pi P\, dt + \sigma P\, dz$$

where π is the expected inflation rate and σ is its proportional standard deviation per unit time. The real value of the bond portfolio at time t will be

$$b(t) = \frac{B(t)}{P(t)} = \frac{e^{rt}}{P(t)}$$

What is the expected real return on the bonds?

Applying Itô's lemma to b, with

$$\begin{array}{lcll} b_t & = & re^{rt}/P & = & rb \\ b_P & = & -B/P^2 & = & -b/P \\ b_{PP} & = & 2B/P^3 & = & 2b/P^2 \end{array}$$

we get db:

$$db(t) = (\sigma^2 - \pi + r)\, b\, dt + \sigma b\, dz$$

Thus the expected real rate of return to holding nominal bonds in this world of uncertain inflation is $r - \pi + \sigma^2$.

The n–dimensional case

Here we simply state the extension of Itô's lemma to the case of several variables. Suppose $y = f(s_1, s_2, ..., s_n, t)$ is a function of n random state variables and time. Let the vector s follow a joint random process described by

$$ds(t) = \alpha\, dt + \sigma\, dz(t)$$

where α is now a n-dimensional column vector, σ is a matrix with n rows and m columns, and $z(t)$ is a m-dimensional column vector of independent Brownian motions. Then

$$dy(t) \quad = \quad (\, f_t + \alpha' f_s + \tfrac{1}{2}\textstyle\sum_i \sum_j f_{ij}[\sigma\sigma']_{ij})\, dt + f_s'\sigma\, dz(t)$$

in which f_{ij} denotes $\partial^2 f/\partial s_i \partial s_j$ and f_s denotes the column vector of partial derivatives $\partial f/\partial s_i$.

Exercise

Let $P(t)$ be the price of a \$1 maturity value pure discount bond at time t, maturing at time T. Let $r(t)$ denote the continuously compounded yield to maturity at time t on discount bonds maturing at calendar date T. Suppose that r is know to follow the stochastic process

$$dr(t) = \alpha\, dt + \sigma r^{1/2} dz(t)$$

with σ a constant.

1. What is stochastic process is followed by the price $P(t)$?
2. What is the instantaneous expected yield at time t on holding the bond?
3. Does the answer to 2) make sense in the case where σ and α are both 0?
4. If σ equals 0, what is the time path followed by the riskless instantaneous interest rate?

2. Valuation by Arbitrage

Common to much of continuous time asset valuation theory is the result that the price of a security is the solution to some sort of partial differential equation (pde). This note derives the valuation pde from the notion that in equilibrium there will be no riskless arbitrage opportunities. We consider the case of just one 'state' variable, or dimension of relevant uncertainty for the securities involved. This does not require that all else in the economy be non-random. It only means that we *assume* that the prices of the securities we are examining are *not* contingent on those other factors.

Let the aspect of the world that is uncertain be some state variable s, with $s(t)$ denoting its level at time t. Assume its evolution over time follows an Itô process, i.e., can be meaningfully described by

$$ds = \alpha\, dt + \sigma\, dz \tag{1}$$

where dz is the increment in a standard Wiener process and α and σ may be functions of s and t. Let there be two tradeable securities whose values at time t are functions of t and $s(t)$. We use $A(s,t)$ and $B(s,t)$ to denote their contingent prices. Applying Itô's Lemma, these prices evolve according to

$$\begin{aligned} dA &= (\tfrac{1}{2}\sigma^2 A_{ss} + \alpha A_s + A_t)\, dt + \sigma A_s\, dz \\ dB &= (\tfrac{1}{2}\sigma^2 B_{ss} + \alpha B_s + B_t)\, dt + \sigma B_s\, dz \end{aligned} \tag{2}$$

In addition, cash may be risklessly borrowed or lent at an instantaneous floating interest rate r. This rate may also be a function of (s,t).

Now consider a portfolio consisting of 1 unit of A and $-A_s/B_s$ units of B. Suppose it is acquired completely by borrowing, with the proceeds of the short sale of B available to reduce the amount owed. One's net borrowing is thus $(A - A_s B/B_s)$. The value of this position, call it P, evolves as follows:

$$\begin{aligned} dP &= dA - \frac{A_s}{B_s} dB - (\text{net borrowings})\, r\, dt \quad = \\ &\left(\tfrac{1}{2}\sigma^2(A_{ss} - \frac{A_s}{B_s}B_{ss}) + \alpha(A_s - \frac{A_s}{B_s}B_s) + (A_t - \frac{A_s}{B_s}B_t) - r(A - \frac{A_s}{B_s}B)\right) dt \end{aligned} \tag{3}$$

Note that the position is riskless — the dz terms cancelled out as a result of the ratio of B to A chosen — and was costless to acquire. If dP was anything other than 0 then the position (or its exact opposite) would offer a sure profit, something for nothing. As long as there was at least one individual trading for whom more was better, such as situation could not persist. Thus $dP = 0$ is a requirement of market equilibrium.

Imposing this and rearranging equation (3) gives us

$$\frac{\frac{1}{2}\sigma^2 A_{ss} + \alpha A_s + A_t - rA}{A_s} \quad = \quad \frac{\frac{1}{2}\sigma^2 B_{ss} + \alpha B_s + B_t - rB}{B_s} \tag{4}$$

The numerator of each side is the expected return from holding the asset over and above the riskless return opportunity cost of holding it (the 'excess expected return'). The denominator is the sensitivity of the asset's value to fluctuations in the state, or number of units of 's–risk' one bears by holding it. Absence of arbitrage implies that this ratio is the same for each asset. This common value will be denoted by $\lambda(s,t)$ and called the market price of s–risk.

Since A could have been paired with a different asset, the main point is that there is single λ common to *all* assets whose prices are functions of s and t. Equating the left side of (4) to λ and rearranging gives the fundamental valuation pde

$$\frac{1}{2}\sigma^2 A_{ss} + (\alpha - \lambda)A_s + A_t - rA \quad = \quad 0 \tag{5}$$

If one is willing to assume a specific functional form for the risk price $\lambda(s,t)$, this pde can in principle be solved for the equilibrium state contingent price of the security. But the notion of no arbitrage by itself does not say what this risk price should be. It only says that the same aversion (or attraction) to s–risk will be embodied in all securities.

Rearranging (5) aids in interpreting it. Writing it as

$$\frac{\frac{1}{2}\sigma^2 A_{ss} + \alpha A_s + A_t}{A} \quad = \quad r + \lambda \frac{A_s}{A} \tag{6}$$

puts it in a form reminiscent of the Capital Asset Pricing Model. The left side is the expected rate of return to holding A. It equals the riskless rate plus an amount proportional to A's proportional sensitivity to s. This proportional sensitivity will have the same value as the local covariance of A with s divided by the local variance of s. If s was the value of the 'market portfolio' then λ would be its excess expected return. This is not the CAPM, however, since A and s are perfectly correlated locally.

Another interpretation of (5) comes from recognizing that we would have the same pde if the expected rate of change in s was $\hat{\alpha} \equiv \alpha - \lambda$ but there were no λ term, i.e., the expected rate of return on holding A was equal to the riskless interest rate r. This means securities sell for the same price as they would in a risk neutral world in which s followed the 'risk adjusted' stochastic process of equation (1) with $\hat{\alpha}$ replacing α . This $\hat{\alpha}$ is termed the

risk adjusted growth rate in s. One consequence of this property is that one may, for example, perform a Monte Carlo simulation of the *risk adjusted* process for a state variable, then estimate A by simply calculating average present values of the cash flows arising from the security.

Equation (5), as it stands, does not uniquely determine a function $A(s,t)$. Many different functions can satisfy the relation, corresponding to the fact that there are many types of securities whose value could be a deterministic function of s and t. To obtain a unique solution one must impose additional restrictions on A. These are lumped together under the name *boundary conditions*, and are what distinguishes one s-dependent security from another. The term arises from the fact that A is presumed to satisfy a differential equation only in an open (though possibly unbounded) region in the (s,t) plane. Characteristics at the boundary of the region 'pin down' what function it is.

One type of boundary condition occurs if the security is one that matures or expires, and its maturity value is a known function of s. This is termed an *initial value* condition since it is convenient in such contexts to let $t = 0$ represent that time and suppose that time runs backwards from a positive value down to 0. For example, a default free bonds satisfies $A(s,0) = 1.0$ for all s if its maturity value is one dollar. A claim to one ounce of gold satisfies $A(s,0) = s$ if s is the spot price of gold. Addition types of boundary conditions will be introduced as we proceed.

The Black-Scholes case

An important simplification occurs when the state variable is itself the *price of a traded asset.* This is the case when, for example, s is the price of the shares in some company and A is price of an option to purchase or sell a share at a particular exercise price, investigated by Fischer Black and Myron Scholes (1973). Let the asset B be the underlying stock, i.e., $B(s,t) = s$. Then $B_s = 1$, $B_{ss} = 0$, and $B_t = 0$. Making these substitutions into the valuation pde reduces it to

$$\alpha - \lambda = rs \tag{7}$$

Black and Scholes also assume that s follows a constant proportional volatility process, i.e., $ds(t) = \alpha(s,t)\,dt + \sigma s\,dz(t)$ where σ is a constant. Substituting rs for $\alpha - \lambda$ into the pde for the derivative asset A reduces (5) to

$$\tfrac{1}{2}\sigma^2 s^2 A_{ss} + rsA_s - A_t - rA = 0 \tag{8}$$

We adopt the convention here of letting t denote the time remaining to expiry of the option. Since t is now declining as time moves forward this

reverses the sign on the A_t term in the pde.

Since neither α nor λ enter the pde, the value of the option in terms of the stock price s is independent of both risk attitudes and the expected rate of change in the stock price! Put another way, the current value of s embodies all that is needed about these things to determine the value of A.

For some types of derivative securities (equivalently some types of boundary conditions), an explicit solution to (8) can be obtained. If the security is a European call option with exercise price X, maturing in T years, on the asset whose price is s, then the boundary condition is

$$A(s,0) = \max\{0, s(0) - X\} \tag{9}$$

The function A satisfying (8) subject to (9) is

$$A(s,T) = sN(d) - Xe^{-rT}N(d - \sigma T^{1/2}) \tag{10}$$

in which $N()$ denotes the cumulative normal distribution function and

$$d \equiv \frac{\ln(\frac{s}{X}) + (r + \frac{\sigma^2}{2})T}{\sigma T^{1/2}} \tag{11}$$

Exercises

1. The cumulative normal distribution function is defined as $N(y) = (2\pi)^{-1/2} \int_{-\infty}^{y} e^{-y^2/2} dy$. Use the chain rule to get the partial derivatives of A given in (10) and show that A satisfies the pde (8).

2. Extend the arbitrage argument of this section to the case where the underlying asset whose price is s pays a dividend continuously at a rate $c(s,t)$, and the derivative security whose price is A pays dividends at a continuous rate $q(s,t)$. Show that the valuation pde (8) in this more general case is

$$\tfrac{1}{2}\sigma^2 s^2 A_{ss} + (rs - c)A_s + q - A_t - rA = 0$$

Multi-factor arbitrage valuation

Let there be n state variables denoted by the vector $s = (s_1(t), \ldots s_n(t))$. Suppressing the time argument, let each follow a diffusion process

$$ds_i = \alpha_i \, dt + \sigma_i dz_i$$

where the instantaneous drift and volatility may be (well-behaved) functions of (s,t), and dz_i is the increment in a standard Weiner process. These increments can be correlated: ρ_{ij} denotes the instantaneous correlation between dz_i and dz_j.

Suppose there are n locally linearly independent assets (to be clarified below) whose prices are deterministic functions $A^k(s,t)$ of t and $s(t)$. Applying Ito's lemma tells us asset k follows the process

$$dA^k = (\tfrac{1}{2}\textstyle\sum_i \sum_j \rho_{ij}\sigma_i\sigma_j A^k_{ij} + \sum_i \alpha_i A^k_i + A_t)\,dt + \sum_i \sigma_i A^k_i\,dz_i$$

Subscripts on A indicate appropriate partial derivatives.

Construct n portfolios X^i combining riskless bonds yielding r and these assets such that: (1) each portfolio has 0 current value, and (2) the derivative of the value of portfolio i with respect to dz_k equals σ_i for $k=i$, 0 for $k\neq i$. Letting β be the matrix of amounts β_{ij} of asset j held in portfolio i, and A_s be the Jacobian $[\partial A^i/\partial s_j]$, this means

$$\beta A_s = S \equiv \mathrm{diag}(\sigma_i)$$

S is a matrix with diagonal elements σ_i, 0 elsewhere. This construction is possible if A_s is not singular, with $\beta = SA_s^{-1}$. Portfolio X^i thus has unit risk exposure to s_i-risk. The expected return, or instantaneous drift, of these portfolios can be found from the drifts of A^k — very messy. Let us denote the expected return per unit time on portfolio X^i by λ_i. Hence $dX^i = \lambda_i\,dt + \sigma_i\,dz_i$.

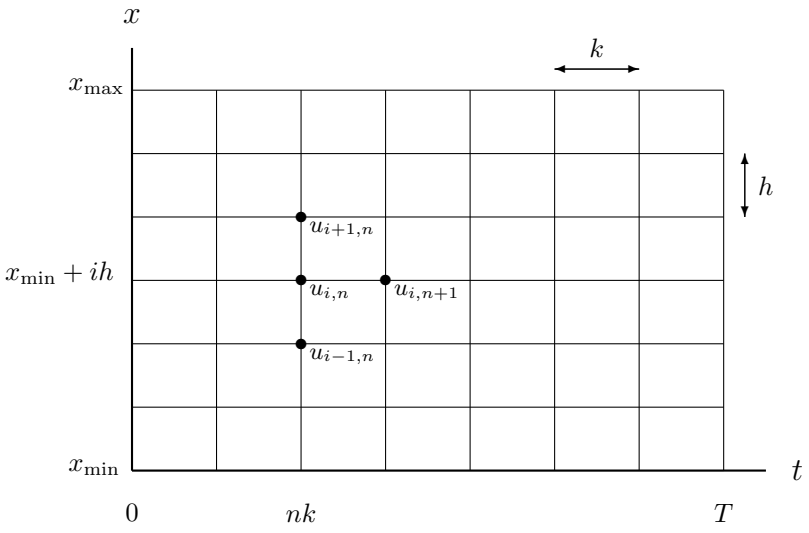
Now consider any other asset with price $P(s,t)$. Construct a portfolio V consisting of one unit of this asset, $-P$ dollars of riskless bonds to pay for it, and $-\partial P/\partial s_i$ units of each portfolio X^i constructed above. The latter have zero current value so require no further financing. But they precisely offset the effect of the dz_i's on P in the portfolio, rendering it riskless. For there not to be an arbitrage opportunity, the return on V over an interval dt must thus be zero:

$$dV = \left(\tfrac{1}{2}\textstyle\sum_i \sum_j \rho_{ij}\sigma_i\sigma_j P^k_{ij} + \sum_i \alpha_i P^k_i + P_t - rP - \sum_i \lambda_i P_i\right)\,dt = 0$$

Dividing by dt and rearranging, equilibrium P must thus satisfy

$$\tfrac{1}{2}\textstyle\sum_i \sum_j \rho_{ij}\sigma_i\sigma_j P^k_{ij} + \sum_i(\alpha_i - \lambda_i)P^k_i + P_t - rP = 0$$

Figure 1: pde solution grid



3. Numerically Solving PDE's: Crank-Nicholson Algorithm

This note provides a brief introduction to finite difference methods for solving partial differential equations. We focus on the case of a pde in one state variable plus time. Suppose one wishes to find the function $u(x,t)$ satisfying the pde

$$au_{xx} + bu_x + cu - u_t = 0 \tag{1}$$

subject to the initial condition $u(x,0) = f(x)$ and other possible boundary conditions. This initial condition will correspond to a maturity or expiry date value condition in our applications and t will denote time left to maturity. Thus time will run backwards down to 0, explaining the negative u_t term in (1). The coefficients a, b, c may themselves be functions of x, t. Any parameters are presumed given, so that one can calculate a numerical value for these functions for any given x, t.

Obviously one cannot 'calculate' the entire function u. What we shall consider a solution is the numerical values that u takes on a grid of x, t values placed over some domain of interest. Once we have these u values, since u is assumed to be smooth almost everywhere, we can interpolate within this grid to get values for arbitrary x, t.

To this end, suppose the domain we will work on is rectangular with x ranging from $x_{\min}$ to $x_{\max}$ and t ranging from 0 to T. Divide $[0,T]$ into N equally spaced intervals at t values indexed by $n = 0, 1, \ldots, N$, and $[x_{\min}, x_{\max}]$ into I intervals at x values indexed by $i = 0, 1, \ldots, I$. The length of these intervals is k in the time direction and h in the x direction. We seek an approximation to the true values of u at the $(N+1) \times (I+1)$ gridpoints. Let $u_{i,n}$ denote our approximation at the gridpoint where $x = x_{\min} + ih$, $t = nk$.

The next step — and this is what makes this procedure a *finite difference* method — is to approximate the partial derivatives of u at each gridpoint by difference expressions in the as yet unknown $u_{i,n}$'s. We can calculate $u_{i,0}$ for each i directly from the initial value condition f. Thus it is natural to start from this boundary are work outward, calculating the $u_{i,n+1}$'s from $u_{i,n}$. Focussing on an arbitrary internal gridpoint in, one could approximate the partial derivatives at that point by the following:

$$\begin{aligned} u &= u_{i,n} \\ \partial u/\partial t &= \frac{u_{i,n+1} - u_{i,n}}{k} \\ \partial u/\partial x &= \frac{u_{i+1,n} - u_{i-1,n}}{2h} \\ \partial^2 u/\partial x^2 &= \frac{u_{i+1,n} - 2u_{i,n} + u_{i-1,n}}{h^2} \end{aligned} \tag{2}$$

The differences in the x, or state, direction have been centered around the point in to give 'second order' accuracy to the approximation. These expressions could then be substituted into the pde (1). Solving the resulting equation for $u_{i,n+1}$ gives the explicit solution

$$u_{i,n+1} = (\frac{k}{h^2}a + \frac{k}{2h}b)u_{i+1,n} + (1 + kc - \frac{2k}{h^2}a)u_{i,n} + (\frac{k}{h^2}a - \frac{k}{2h}b)u_{i-1,n} \tag{3}$$

One could then proceed to calculate all the $u_{i,n+1}$'s from the $u_{i,n}$'s and recursively obtain u for the entire grid. Since equation (3) applies only to the interior gridpoints, we at each time step would have to make use of some other boundary conditions (e.g., at $x_{\min}$ and $x_{\max}$) to identify all the u's. More will be said about this later.

The result of this is called an *explicit* finite difference solution for u. It is second order accurate in the x direction, though only first order accurate in the t direction, and easy to implement. Unfortunately the numerical solution is *unstable* unless the ratio k/h^2 is sufficiently small. By unstable is meant that small errors due either to arithmetic inaccuracies or to the approximate

nature of the derivative expressions will tend to accumulate and grow as one proceeds rather than dampen out. Loosely speaking, this will occur when the difference equation system described by (3) has eigenvalues greater than one in absolute value. See a numerical analysis book such as Vemuri and Karplus (1981) or Lapidus and Pinder (1982) for discussion of stability issues.

The instability problem can be handled by instead using and *implicit* finite difference scheme. The recommended method for most problems in the Crank-Nicholson algorithm, which has the virtues of being unconditionally stable (i.e., for all k/h^2) and also is second order accurate in both the x and t directions (i.e., one can get a given level of accuracy with a coarser grid in the time direction, and hence less computation cost). This is the algorithm implemented by the routines `CNSET` and `CNSTEP` handed out in class.

The algorithm entails proceeding as before, but using difference expressions for the partial derivatives which are centered around $t + k/2$ rather than around t. Thus the expressions for u, u_x and u_{xx} are averages of what we had in (3) for times n and $n + 1$:

$$\begin{aligned} u &= \frac{u_{i,n} + u_{i,n+1}}{2} \\ \partial u/\partial t &= \frac{u_{i,n+1} - u_{i,n}}{k} \\ \partial u/\partial x &= \frac{u_{i+1,n} - u_{i-1,n} + u_{i+1,n+1} - u_{i-1,n+1}}{4h} \\ \partial^2 u/\partial x^2 &= \frac{u_{i+1,n} - 2u_{i,n} + u_{i-1,n} + u_{i+1,n+1} - 2u_{i,n+1} + u_{i-1,n+1}}{2h^2} \end{aligned} \tag{4}$$

Substituting the above into the pde, multiplying through by $4h^2k$ to eliminate the denominators, and collecting all the terms involving the unknown $u_{\cdot,n+1}$'s on the left hand side results in

$$\begin{aligned} &\overbrace{-(2ka + khb)}^{A_i} u_{i+1,n+1} + \overbrace{(4h^2 + 4ka - 2h^2kc)}^{B_i} u_{i,n+1} \overbrace{-(2ka - khb)}^{C_i} u_{i-1,n+1} \\ &= \\ &\qquad \underbrace{(2ka + khb)u_{i+1,n} + (4h^2 - 4ka + 2h^2kc)u_{i,n} + (2ka - khb)u_{i-1,n}}_{D_i} \end{aligned} \tag{5}$$

for each $i = 1, \ldots, I - 1$. It is apparent that that the $u_{\cdot,n+1}$'s cannot individually be written as simple linear combinations of the $u_{\cdot,n+1}$'s, but are simultaneously determined as the solution to this system of linear equations. However this system has a very convenient structure. Written in

matrix form, (5) provides the interior equations of

$$\begin{bmatrix} B_0 & C_0 & 0 & 0 & \cdots & 0 \\ A_1 & B_1 & C_1 & 0 & \cdots & 0 \\ 0 & A_2 & B_2 & C_2 & 0 & \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & & 0 & A_I & B_I \end{bmatrix} \begin{pmatrix} u_{0,n+1} \\ u_{1,n+1} \\ u_{2,n+1} \\ \vdots \\ u_{I,n+1} \end{pmatrix} = \begin{pmatrix} D_0 \\ D_1 \\ D_2 \\ \vdots \\ D_I \end{pmatrix} \tag{6}$$

This sort of system is most efficiently solved by Gaussian elimination. It requires just $8I$ floating point operations to determine the unknown $u_{\cdot,n+1}$'s (e.g., see Press et al., 1986, p.40 or the listing of routine TRIDAG at the end of CNSTEP).

Boundary conditions

We almost have a procedure for recursively determining the entire grid of $u_{i,n}$'s starting from the given initial values. Substitution of the the difference expressions into the differential equation only gave us a linear equation for each interior point in the grid. That gives $I-1$ equations at each time step, which is not sufficient to determine the $I+1$ unknowns. The missing two equations must be provided by boundary conditions applied at each time step. It would be desirable for these to be representable in a form that preserves the tridiagonal form of the system and thus the efficiency of solution. Three possibilities are considered.

The most convenient to apply would be knowledge of the value of u on the boundaries $x_{\max}$ and $x_{\min}$. In (6), this could be embodied by setting the coefficients $B_0 = 1$, $C_0 = 0$, $A_I = 0$ and $B_I = 1$, then setting D_0 and D_I on the right hand side of (6) equal to the known values $u(x_{\min}, t)$ and $u(x_{\max}, t)$ respectively. This could be appropriate, for example, if u was the value of a call option on a stock whose current price was x, and the assumed stochastic process for x had $0 = x_{\min}$ as an absorbing state (if x ever hit 0, it never again moves away from it). The call option would then be worth nothing — i.e., $u(x_{\min}, t) = 0$.

Alternatively, from the nature of the security being valued we may have knowledge of the derivative u_x at one or the other boundary. This could be expressed in the form of (6) by setting $B_0 = -1$, $C_0 = 1$ and $D_0 = hu_x(x_{\min}, t)$. The upper boundary slope is handled analogously. This could be appropriate, for example, if a boundary was in the region where an option would rationally be exercised immediately, and if the derivative of the exercise value of the option with respect to x was known. The difference

approximation we are using the u_x is actually centered around a point $h/2$ inside the boundary. If u was known to be linear in the neighbourhood of the boundary this should be adequate. If not, then a more elaborate approximation to the slope at the boundary (e.g., that of a quadratic through the boundary gridpoint and the two interior ones) should be used. This would initially upset the tridiagonal form of (6) by introducing a non-zero term just to the right of C_0. However a manual elimination of this term by subtracting a appropriate multiple of the second equation from the first would restore the system to the desired form.

In each of the above cases, the boundary at which you know the value or slope of u might be at $x = \infty$. In that case you might consider transforming the domain for u to some finite interval. For example, if x ranges from 0 to ∞, then $y = x/(1 + x)$ ranges from 0 to 1. Transforming the pde into one in $v(y, t) \equiv u(x(y), t)$ and solving for v would let you use the known boundary information at $y = 1$. See Brennan and Schwartz (1979) for an example. One disadvantage of this approach is that a fairly large grid may be needed to get a desired level of fineness in the levls of x of interest. In many situations you will do better with the next approach.

The third possibility is that we have no simple information about u at the boundaries we have chosen, yet we feel that the behaviour at these boundaries should have little impact on the behaviour of u in that part of the grid in which we are interested. This could be the case if the probability of reaching the boundary from relevant starting levels of the state in time $t_{\max} - t_{\min}$ is small. This could occur if, for example, the state variable followed a mean-reverting process. A procedure that I have found often works in this case is to determine the boundary values of u by simple extrapolation of the immediately interior values. Focussing on the lower boundary, quadratic extrapolation of u_3, u_2, u_1 to u_0 implies

$$(x_3 - x_2) - (x_2 - x_1) = (x_2 - x_1) - (x_1 - x_0)$$

or, equivalently,

$$x_0 - 3x_1 + 3x_2 - x_3 = 0$$

Using this relation as the first equation in system (6) gives us

$$\begin{bmatrix} 1 & -3 & 3 & -1 & \cdots & 0 \\ A_1 & B_1 & C_1 & 0 & \cdots & 0 \\ 0 & A_2 & B_2 & C_2 & 0 & \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & & 0 & A_I & B_I \end{bmatrix} \begin{pmatrix} u_{0,n+1} \\ u_{1,n+1} \\ u_{2,n+1} \\ \vdots \\ u_{I,n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ D_1 \\ D_2 \\ \vdots \\ D_I \end{pmatrix} \tag{7}$$

which is no longer tridiagonal. However the offending elements of the first row of our matrix can be eliminated as follows: Define

$$G \equiv \frac{C_1}{B_2 + 3C_2}$$

Multiply the first row by GC_2. Add G times the third row to the first, then subtract the second row from the first. A little algebra verifies that the third and fourth elements of the first row are now 0. The first two elements and corresponding right hand side element are now

$$\begin{aligned} B_0 &= GC_2 - A_1 \\ C_0 &= G(A_2 - 3C_2) - B_1 \\ D_0 &= GD_2 - D_1 \end{aligned} \tag{8}$$

Extrapolation at the $x_{\max}$ boundary could be similarly handled. Treating the boundaries in this fashion may not be stricly correct for a given problem. We are in effect imposing a boundary condition $u_{xxx} = 0$. This is consistent with u being truly linear or constant in the neighbourhood of the boundary. If used, one should check to see how much damage it may be causing by trying out a few differing levels of $x_{\min}$ and $x_{\max}$ to make sure the solution in the domain of interest is not sensitive to it.

In the routine CNSET this last procedure is the default, implemented if the flags ISMN or ISMX are set to 0. If either is set to 1 then it is presumed that a known boundary value is being provided via functions FMIN(T) or FMAX(T) respectively.

Interpretation

Viewed as a numerical procedure, what we have done is convert a partial differential equation that holds everywhere in some domain into a system of simultaneous linear equations to get an approximate solution. However what is going on can be given economic interpretation. This is most readily apparent in the explicit approximation.

Separating out the term involving c in equation (3) lets us write $u_{i,n+1}$ as

$$\begin{aligned} u_{i,n+1} = & \\ & ((\frac{k}{h^2}a + \frac{k}{2h}b)u_{i+1,n} + (1 - \frac{2k}{h^2}a)u_{i,n} + (\frac{k}{h^2}a - \frac{k}{2h}b)u_{i-1,n}) + kcu_{i,n} \end{aligned} \tag{9}$$

The coefficient c will usually equal $-r$ in valuing securities. The final term is thus $-rku_{i,n}$, the interest lost by holding u for an amount of time k. The

remaining coefficients on the right hand side add up to one. They can be viewed as the probabilites of the state x moving up by h, staying unchanged, and moving down by h respectively. Indeed, if these were the only moves possible over the interval $n+1$ to n, one may verify that the expected change in x is kb and its variance is $2ka$. The distribution of x one period ahead thus agrees with that of the risk adjusted process for x in terms of mean and variance.

Equation (9) calculates $u_{i,n+1}$ as the probability weighted average of the possible values u may take on next period, with allowance for the opportunity cost of holding the security. This is just the expected discounted value of the security at the next time point. The values at the time n had been calculated in turn as the discounted expectation of their possible values at time $n-1$, and so on down to the asset's maturity. This permits a more intuitive interpretation of the valuation pde that we started out to solve: It states that the equilibrium of an asset equals the expectation of the discounted present value of its cash flows to maturity, where the expectation is with respect to the risk adjusted process for the state variables. Cox, Ingersoll and Ross (1985) provide a formal demonstration of this interpretation.

One can also see the limits of the explicit difference. If k is too large relative to h, some of the alleged probabilities will be negative. This reflects the maximum volatility of the state variable that can be represented by simply moving up or down by h over an interval k. Negative probabilities in this interpretation correspond to instability of the difference scheme.

We can also interpret the implicit Crank-Nicholson scheme as a jump process approximation to the risk adjusted x process. The important difference is that that approximation permits jumps to *any* point in the x grid over the interval k, and the implicit probabilities are all positive. Thus an arbitrarily large volatility can always be represented and large k may be used. We leave it as an exercise to work out the details. Or you may consult Brennan and Schwartz (1978).