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,

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

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., $\mathrm{E}[\Delta z^2] = \Delta t$.

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

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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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:

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

Dividing by dt and rearranging, equilibrium P must thus satisfy

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