Important concepts from today: solving linear ODEs, solving linear SDEs, converting linear SDE into ODEs satisfied by the mean and variance of the solution.

1. Review of last time

Girsanov’s theorem: if $B_t$ is a Brownian motion on $(\Omega, \mathbb{P})$ and $Q$ is a measure defined by

$$
\frac{dQ}{d\mathbb{P}}(\omega) = \exp \left( - \int_0^T u(s) dB_s(\omega) - \frac{1}{2} \int_0^T u(s)^2 ds \right)
$$

then under $Q$, the process $\tilde{B}_t = B_t + \int_0^t u(s) ds$ is a standard Brownian motion on the interval $[0, T]$. As a special case, if $u(s) = q$ is constant, then

$$
\frac{dQ}{d\mathbb{P}}(\omega) = \exp \left( -qB_T(\omega) - \frac{1}{2} q^2 T \right).
$$

In the process of proving this result, we saw that in order to look at events on $F_t$ (the $\sigma$-field of $B_t$), we needed to know not the full change of measure $dQ/d\mathbb{P}$, but the likelihood ratio

$$
L_t = \mathbb{E} \left( \frac{dQ}{d\mathbb{P}} | F_t \right).
$$

We checked that $L_t$ was a martingale under $\mathbb{P}$, and used that to show that $\tilde{B}_t$ was a martingale under $Q$, or equivalently that $\tilde{B}_tL_t$ was a martingale under $\mathbb{P}$:

$$
\mathbb{E}_\mathbb{P} \left( \tilde{B}_t \frac{dQ}{d\mathbb{P}} | F_s \right) = \mathbb{E}_\mathbb{P} \left( \tilde{B}_t \mathbb{E} \left( \frac{dQ}{d\mathbb{P}} | F_t \right) | F_s \right)
$$

$$
= \mathbb{E}_\mathbb{P} \left( \tilde{B}_s \mathbb{E} \left( \frac{dQ}{d\mathbb{P}} | F_t \right) | F_s \right)
$$

$$
= \mathbb{E}_\mathbb{P} \left( \tilde{B}_s \frac{dQ}{d\mathbb{P}} | F_s \right).
$$

That is, in order to compute the conditional expectations on the $\sigma$-field $F_s$, we needed the likelihood ratio averaged out over events in that $\sigma$-field.

We also saw the converse to Girsanov’s theorem: any likelihood ratio that’s adapted to the filtration of Brownian motion has to take the form of an exponential like the one appearing in Girsanov’s theorem. In particular, if the change of measure is coming from the underlying Brownian motion, it can’t change the volatility of a Brownian motion, only its drift.

Date: 12/3/2015.
2. Stochastic differential equations

Our goal today is to say a few words about solutions to stochastic differential equations, i.e. equations that look like this:

\[ dX_t = A_1(t, X_t)dt + A_2(t, X_t)dB_t, \quad X_t = X_0 + \int_0^t A_1(s, X_s)ds + \int_0^t A_2(s, X_s)dB_s. \]

We are interested in solutions where \( X_t \) is adapted to the filtration of Brownian motion; these are called strong solutions. This means that \( X_t \) is a function of the path of the Brownian motion up to time \( t \). (You could also look for “\( X_t \) and the integral on the right have the same distribution”, that’s called a weak solution.) To “solve” an SDE means to write down an expression for \( X_t \) that isn’t circular – we could have integrals in the answer that we don’t know how to evaluate, but we won’t have “\( X_t \) is the integral of \( X_t \)”.

**Theorem 2.1** (Existence and uniqueness of strong solutions). There exists a unique strong solution to an SDE (with initial conditions) if:

1. The initial condition \( X_0 \) is independent of the Brownian motion;
2. The coefficient functions \( A_1(t, X_t) \) and \( A_2(t, X_t) \) are continuous (in each variable);
3. The coefficient functions are Lipschitz in \( X_t \).

Lipschitz formally means:

\[ |a(t, y) - a(t, x)| \leq C|y - x|, \quad \forall x, y. \]

A good way of thinking about it is that it’s continuous, differentiable almost everywhere, and the derivative is bounded.

**Example 1.**

All linear functions are Lipschitz.

The functions \( |x|, (x - K)_+ \) are Lipschitz.

The function \( f(x) = x^2 \) is Lipschitz if \( x \) is bounded, but not on \( \mathbb{R} \) (the derivative tends to \( \infty \) as \( x \to \infty \)). It would work as the coefficient function if \( |X_t| \) is guaranteed to be bounded for some reason.

The function \( f(x) = \sqrt{x} \) is Lipschitz on \([\epsilon, \infty)\) for any \( \epsilon \), but is not Lipschitz on \([0, 1]\). It would work as the coefficient function if \( X_t \) is guaranteed to be not too small.

The most general SDE that we will learn to solve is the general linear SDE:

\[ dX_t = (c_1(t)X_t + c_2(t))dt + (\sigma_1(t)X_t + \sigma_2(t))dB_t. \]

This covers many interesting financial processes – the simplest place you’d see this is if you’re modeling time-dependent rate of return and volatility.

2.1. Interlude on solving linear ODEs. Before dealing with stochastic SDEs, let’s look at the linear ODE,

\[ \frac{dx}{dt} = c_1(t)x + c_2(t). \]

The general plan of solution is as follows:

1. Let \( p(t) = -\int_0^t c_1(s)ds \). Make sure to get the sign right!
2. The integrating factor is \( y(t) = e^{p(t)} \).
3. The point is that

\[ \frac{d}{dt} \left( e^{p(t)}x \right) = e^{p(t)} \left( c_1(t)x + c_2(t) \right) + \left( -c_1(t) \right) e^{p(t)}x = e^{p(t)}c_2(t). \]
Therefore,
\[ e^{p(t)}x(t) - e^{p(0)}x(0) = \int_0^t e^{p(s)}c_2(s)ds, \quad x(t) = e^{-p(t)} \left( x(0) + \int_0^t e^{p(s)}c_2(s)ds \right). \]

2.2. Solving the linear SDE.

Example 2 (Vasicek interest rate model). In the Vasicek interest rate model,
\[ dr_t = C(\mu - r_t)dt + \sigma dB_t. \]
When \( r_t \) gets away from \( \mu \), it tends to come back to it at rate proportional to \( \mu - r_t \), but there are fluctuations of order \( \sigma \) (the volatility).

Let’s solve: \( p(t) = \int_0^t Cds = Ct \), so
\[ e^{Ct}r_t - r_0 = \int_0^t e^{Cs}C\mu ds + \int_0^t e^{Cs}\sigma dB_s = \mu(e^{Ct} - 1) + \sigma \int_0^t e^{Cs}dB_s, \]
or
\[ r_t = e^{-Ct}r_0 + \mu(1 - e^{-Ct}) + \sigma e^{-Ct} \int_0^t e^{Cs}dB_s. \]

Let’s compute the mean and variance of the solution, using the fact that the mean of an integral \( dB \) is 0, and the variance of the deterministic part is also zero.
\[ E[r_t] = e^{-Ct}E[r_0] + \mu(1 - e^{-Ct}), \]
and
\[ \text{Var}(r_t) = \text{Var}\left( \sigma e^{-Ct} \int_0^t e^{Cs}dB_s \right) = \sigma^2 e^{-2Ct} \int_0^t e^{2Cs}ds = \sigma^2 \frac{1}{2C}(1 - e^{-2Ct}). \]
Notice that the mean and variance tend to a limit as \( t \to \infty \); this is because \( r_t \) converges to a (normal) random variable in the limit of \( t \to \infty \).

Let’s now do an example where \( X_t \) appears in the \( dB \) and \( dt \) terms, but without any extra additive noise:

Example 3 (Multiplicative noise). Suppose
\[ dX_t = c(t)X_tdt + \sigma(t)X_tdB_t. \]
We expect the solution to be exponential, so we try looking at \( Y_t = \ln X_t \). Then
\[ dY_t = \frac{1}{X_t}dX_t + \frac{1}{2} \frac{-1}{X_t^2}(dX_t)^2 \]
\[ = c(t)dt + \sigma(t)dB_t - \frac{1}{2}(\sigma(t))^2dt \]
so
\[ \ln X_t = \ln X_0 + \int_0^t c(s) - \frac{1}{2}(\sigma(s))^2ds + \int_0^t \sigma(s)dB_s, \]
or
\[ X_0 = X_0 \exp \left( \int_0^t c(s) - \frac{1}{2}(\sigma(s))^2ds + \int_0^t \sigma(s)dB_s \right), \]
a variant on geometric Brownian motion.

By combining the two, you can solve the general linear SDE.
Example 4 (General linear SDE). Suppose
\[ dX_t = (c_1(t)X_t + c_2(t))dt + (\sigma_1(t)X_t + \sigma_2(t))dB_t. \]
Multiplying by the right Itô exponential will kill \(c_1(t)X_tdt + \sigma_1(t)X_tdB_t\). Specifically: let
\[ dY_t = c_1(t)Y_tdt + \sigma_1(t)Y_tdB_t \quad \text{helper equation} \]
and consider
\[
\begin{align*}
\frac{X_t}{Y_t} &= \frac{1}{Y_t} \frac{dX_t}{\partial x} - \frac{X_t}{Y_t^2} \frac{dY_t}{\partial y} \\
&= \frac{1}{Y_t} \left( 0dX_t^2 + \frac{1}{2} \frac{-2X_t}{Y_t^2} \frac{dY_t^2}{\partial y} + \frac{-1}{Y_t^2} \frac{dX_t dY_t}{\partial y} (\sigma_1(t)Y_t + \sigma_2(t)\sigma_1(t)Y_t)dt \right) \\
&= \frac{1}{Y_t} \left( c_2(t) - \sigma_1(t)\sigma_2(t) \right) dt + \frac{1}{Y_t} \sigma_2(t)dB_t.
\end{align*}
\]
I am skipping some algebra where I expand terms and cancel them; you should check for yourself that it actually works out!

This is now an equation that we know how to solve:
\[
\frac{X_t}{Y_t} = \frac{X_0}{Y_0} + \int_0^t \frac{1}{Y_s} \left( c_2(s) - \sigma_1(s)\sigma_2(s) \right) ds + \int_0^t \frac{1}{Y_s} \sigma_2(t)dB_t,
\]
where
\[
Y_t = \begin{cases} 
Y_0 & \text{set to 1 for simplicity} \\
\exp \left( \int_0^t \left( c_1(s) - \frac{1}{2} \sigma_1(s)^2 \right) ds + \int_0^t \sigma_1(s)dB_s \right). 
\end{cases}
\]

3. Mean and Variance of Solution of Linear SDE

While we “solved” SDEs above, we may not get an easy to work with representation, so let’s find some statistics of it.

Suppose
\[ dX_t = (c_1(t)X_t + c_2(t))dt + (\sigma_1(t)X_t + \sigma_2(t))dB_t, \]
meaning
\[
X_t = X_0 + \int_0^t (c_1(s)X_s + c_2(s))ds + \int_0^t (\sigma_1(s)X_s + \sigma_2(s))dB_s.
\]
This tells us that
\[
\mathbb{E}X_t = \mathbb{E}X_0 + \mathbb{E} \int_0^t (c_1(s)X_s + c_2(s))ds = \mathbb{E}X_0 + \int_0^t (c_1(s)\mathbb{E}X_s + c_2(s))ds.
\]
This means that the mean of the SDE satisfies an ordinary differential equation,
\[
m(t) = m(0) + \int_0^t (c_1(s)m(s) + c_2(s))ds, \quad m'(t) = c_1(t)m(t) + c_2(t).
\]
This is a general linear ODE, which we know how to solve (find the integrating factor...)

Note that if we have an equation for \(dX_t\), we can write down an equation for \(d(X_t^2)\) (this is not the square of the differential, this is the differential of \(X_t^2\)). If we figure out \(\mathbb{E}[X_t^2]\), we’ll be able to get the variance of the solution.

We know
\[
d(X_t^2) = 2X_t dX_t + (dX_t)^2 = (2c_1(t) + \sigma_1(t)^2) X_t^2 dt + (2c_2(t)X_t + \sigma_1(t)\sigma_2(t)X_t) dt + 2(\sigma_1(t)X_t^2 + \sigma_2(t)X_t)dB_t.
\]
so if we let \( q(t) = \mathbb{E}[X_t^2] \), we get
\[
q'(t) = (2c_1(t) + \sigma_1(t)^2)q(t) + 2(c_2(t) + \sigma_1(t)\sigma_2(t))m(t) + \sigma_2^2(t).
\]
Notice that the function \( m(t) \) comes into the ODE, so you have to first solve for \( m(t) \).

In the same way we could in principle find differential equations satisfied by \( \mathbb{E}[X_t^3] \), \( \mathbb{E}[X_t^4] \), and so on, although it becomes increasingly unlikely that we’ll be able to write down a nice solution of them.

**Example 5** (Vasicek interest rate model). Recall
\[
dr = C(\mu - r)dt + \sigma dB
\]
had the solution
\[
r = e^{-Ct}r_0 + \mu(1 - e^{-Ct}) + \sigma e^{-Ct} \int_0^t e^{Cs} dB.
\]
The ODE for the mean is
\[
m'(t) = C(\mu - m(t))dt.
\]
This is a very nice equation whose solution is negative exponential: formally, let \( u(t) = \mu - m(t) \), then \( u'(t) = -Cudt \), so \( u(t) = u(0)e^{-Ct} \) and
\[
m(t) = \mu - u(0)e^{-Ct} = \mu - (\mu - m(0))e^{-Ct} = e^{-Ct}m(0) + \mu(1 - e^{-Ct}).
\]

Now, to find the variance, we write
\[
d(r^2_t) = 2r_1dr_1 + (dr_1)^2
\]
\[
= 2r_1(C(\mu - r_1)dt + \sigma dB) + \sigma^2 dt
\]
\[
= (-2Cr_1^2 + 2C\mu r_1 + \sigma^2) dt + 2r_1\sigma dB.
\]
Setting \( q(t) = \mathbb{E}r_t^2 \), we’re allowed to kill the \( dB \) term, and take expectations of the \( dt \) term:
\[
q'(t) = -2Cq(t) + (2C\mu m(t) + \sigma^2)
\]
This means that \( w(t) = e^{2Ct}q(t) \) has derivative
\[
w'(t) = e^{2Ct} (2C\mu m(t) + \sigma^2),
\]
which is a nice integrable thing involving a bunch of exponentials. You can finish the computation yourself; remember that this is \( \mathbb{E}[r_t^2] \), not \( \text{Var}(r_t) \)!

## 4. Higher-dimensional models

In real life, markets have more than one stock in them. A more realistic model is this: we have one riskless bond, whose price is given by \( d\beta_t = r(t)\beta_t dt \), and \( n \) stocks whose prices satisfy
\[
dS_i(t) = S_i(t)c_i(t)dt + S_i(t)\sum_{j=1}^n \sigma_{ij}(t)dB_j^1.
\]
Of course, we also have initial conditions for all of these. Here, \( B^1, B^2, \ldots, B^n \) are independent Brownian motions. \( r(t) \) is the (time-dependent instantaneous) volatility, \( c_i(t) \) is the (instantaneous) mean rate of return, and \( \sigma_{ij}(t) \) are dispersion coefficients. This is saying that the rate of growth of a stock is proportional to its current price, with noise that’s also proportional to price but may be correlated for different stocks; the best way to model that is to start with independent Brownian motions and put a covariance structure on top. In particular, each \( S_i(t) \) is still a geometric Brownian motion here (with mean
rate of return $c_i(t)$, and volatility $\sum_j \sigma_{ij}^2(t)$, but they are correlated among themselves. (See today’s homework!)

To deal with multidimensional models, we add one more rule to our arsenal of Itô rules:

$$d t^2 = 0, \quad dB_i dt = 0, \quad dB^2_i = dt, \quad dB_i dB_j = 0 \text{ when } i \neq j.$$  

(To derive this: $\mathbb{E}[dB_i(t)dB_j(t)] = \mathbb{E}[dB_i(t)]\mathbb{E}[dB_j(t)] = 0$, and $\text{Var}(dB_i(t)dB_j(t)) = dt^2$; when you integrate, because increments of Brownian motion are independent across time, $\text{Var}(\int_0^t dB_i(t)dB_j(t)) = \int_0^t \text{Var}(dB_i(t)dB_j(t)) = O(dt)$. Here I’m treating the integral as a sum, and saying that the variance of the sum is the sum of variances.) To do Itô calculus with these objects, you take partial derivatives in all variables out to second order (remembering that $1/2$ goes in front of any second partial with respect to a single variable, and doesn’t go in front of mixed partials), and expand the products of differentials according to those rules. For example, if $dX_t = dt + dB_1(t) + dB_2(t)$, then

$$d(B_1^2(t) \sin(X_t)) = 2B_1(t) \sin(X_t) dB_1(t)$$

$$+ \left(\frac{dB_1(t)^2}{dt}\right) \sin(X_t) + B_1^2(t) \cos(X_t) dX_t - \frac{1}{2} B_1^2(t) \sin(X_t) \frac{d(\sin(X_t))^2}{dt} = (dB_1(t))^2 + (dB_2(t))^2 = 2dt$$

$$+ 2B_1(t) \cos(X_t) \left(\frac{dB_1(t)}{dt} dB_1(t) + dB_1(t) dX_t\right).$$

A portfolio in the multidimensional market consists of $n + 1$ numbers, $N_i(t)$, corresponding to the number of shares of each asset that I hold. I will take $N_0(t)$ to be the number of bonds. Thus,

$$V_t = N_0(t) \beta(t) + \sum_{i=1}^n N_i(t) S_i(t).$$

It’s self-financing if

$$dV_t = N_0(t) d\beta(t) + \sum_{i=1}^n N_i(t) dS_i(t).$$

The analog of Girsanov’s theorem in higher dimensions asserts that we can change to an equivalent risk-neutral measure in which all the mean rates of return are $r(t)$. The risk-neutral measure is given by

$$\frac{dQ}{dP}(\omega) = \exp \left( -\sum_{i=1}^n \int_0^T \theta_i(s) dB_i^k - \frac{1}{2} \sum_{i=1}^n \int_0^t \theta_i(s)^2 \right),$$

where

$$\theta_i(t) = \sum_{j=1}^n (\sigma(t)^{-1})_{ij} (b_j(t) - r(t)).$$

Here, $\sigma(t)$ is the matrix with entries $\sigma_{ij}(t)$.

If we change to this measure, then we can price contingent claims as follows:

$$e^{-\int_0^T r(s) ds} V_t = \mathbb{E}_Q \left( e^{-\int_0^T r(s) ds} V_T | \mathcal{F}_t \right).$$

Here, we take $V_T$ as given, as a function of (all the) terminal prices $S_T$. We’ve discounted by the risk-free interest rate, and then taken conditional expectation, because the discounted value of the portfolio should be a martingale under $Q$.

It is also possible to set up a PDE for $V_t = u(T - t, S^1_t, \ldots, S^n_t)$ using Itô rules of differentiation, and then hope to solve the PDE. In the PDE version, we would use the
fact that coefficients on $dt, dS_t^1, dS_t^2, \ldots, dS_t^n$ are all uniquely determined. This means that when you write down

$$dV(T-t, S_1(t), S_2(t), \ldots, S_n(t)) = N_0(t)d\beta(t) + \sum_{i=1}^n N_i(t)dS_i(t),$$

you have $n+1$ equations here, which you can solve for $N_i(t)$ in terms of partial derivatives of $V$. The one remaining equation

$$V_t = N_0(t)\beta(t) + \sum_{i=1}^n N_i(t)S_i(t)$$

then gives you a PDE for $V$.

We casually wrote $\sigma(t)^{-1}$ in the Girsanov change-of-measure; this assumes that the matrix $\sigma(t)$ is invertible. If it isn’t invertible, then one of the stock prices is a (predictable) function of the other stocks, because one of the underlying Brownian motions is a linear combination of the others. This means that very likely there’s arbitrage in the market, unless the one stock was priced exactly right. In practice, if you see covariance matrices that have very small eigenvalues (close to 0), you start looking for the linear combinations that have particularly low variance, and there might be arbitrage opportunities there.

5. INCOMPLETE MARKETS

We talked about what happens when you have fewer independent sources of information than tradeable assets; let’s now say something about the opposite case, a market where there are some extra non-tradeable sources of information. For example, you may wish to price a contract based on weather information (travel insurance) – the weather itself isn’t directly tradeable, it only provides information. In other words, you’ve increased your filtration, but not your hedging flexibility. For a simple example, suppose that in the binomial market model we had

$$S_{t+1} = \begin{cases} U \cdot S_t \\ M \cdot S_t \\ D \cdot S_t \end{cases}$$

($M$ for middle). Then there would typically be many distributions on $U, M, D$ that make the stock prices into a martingale, leading to many possible option prices; and there would typically be no way to replicate contracts using just stocks and bonds. For example, the contract that pays out $S_t^2$ at time 1 would be impossible to replicate: we cannot solve three equations

$$a + bU = U^2, \quad a + bM = M^2, \quad a + bD = D^2$$

for two unknowns $a$ and $b$.

Any of the risk-neutral measures lead to arbitrage-free prices, so in an incomplete market there are many possible prices for a contract. The market will decide on one of them somehow, using something other than the requirement of no arbitrage.

As soon as I introduce some freely-traded options into this market, I start to be able to price options relative to each other. The mathematics in the Black-Scholes model gets a bit more involved, because option prices have different distributions than stock prices; but in a binomial market model it’s an entirely reasonable computation. For example, if I have a stock whose price I don’t tell you, but I tell you the price of an option on the stock, then you should be able to invert the Black-Scholes formula and work out the price of the stock from the price of the option.