ma414l10.tex Lecture 10. 15.3.2012

Theorem (Brownian Martingale Representation Theorem). Let $M = (M(t))_{t\geq 0}$ be a RCLL local martingale with respect to the Brownian filtration (\mathcal{F}_t) . Then

$$M(t) = M(0) + \int_0^t H(s) dW(s), \ t \ge 0$$

with $H = (H(t))_{t\geq 0}$ a progressively measurable process such that $\int_0^t H(s)^2 ds < \infty$, $t \geq 0$ with probability one. That is, all Brownian local martingales may be represented as stochastic integrals with respect to Brownian motion (and as such are continuous).

As mentioned above, the economic relevance of the representation theorem is that it shows that the Black-Scholes model is complete – that is, that every contingent claim (modelled as an appropriate random variable) can be replicated by a dynamic trading strategy. Mathematically, the result is purely a consequence of properties of the Brownian filtration. The desirable mathematical properties of BM are thus seen to have hidden within them desirable economic and financial consequences of real practical value.

The next result, which is an example for the rich interplay between probability theory and analysis, links stochastic differential equations (SDEs) with partial differential equations (PDEs). Such links between probability and stochastic processes on the one hand and analysis and partial differential equations on the other are very important, and have been extensively studied. Suppose we consider a stochastic differential equation,

$$dX(t) = \mu(t, X(t))ds + \sigma(t, X(t))dW(t) \qquad (t_0 \le t \le T), \qquad X(t_0) = x.$$

For suitably well-behaved functions μ, σ , this stochastic differential equation will have a unique solution $X = (X(t) : t_0 \leq t \leq T)$. Taking existence of a unique solution for granted for the moment, consider a smooth function F(t, X(t)) of it. By Itô's lemma,

$$dF = F_t dt + F_x dX + \frac{1}{2} F_{xx} d\langle X \rangle,$$

and as $d\langle X \rangle = \langle \mu dt + \sigma dW \rangle = \sigma^2 d\langle W \rangle = \sigma^2 dt$, this is $dF = F_t dt + F_x(\mu dt + \sigma dW) + \frac{1}{2}\sigma^2 F_{xx} dt = (F_t + \mu F_x + \frac{1}{2}\sigma^2 F_{xx})dt + \sigma F_x dW.$ Now suppose that F satisfies the partial differential equation

$$F_t + \mu F_x + \frac{1}{2}\sigma^2 F_{xx} = 0$$

with boundary condition,

$$F(T, x) = h(x).$$

Then the above expression for dF gives

$$dF = \sigma F_x dW,$$

which can be written in stochastic-integral rather than stochastic-differential form as

$$F(s, X(s)) = F(t_0, X(t_0)) + \int_{t_0}^s \sigma(u, X(u)) F_x(u, X(u)) dW(u).$$

Under suitable conditions, the stochastic integral on the right is a martingale, so has constant expectation, which must be 0 as it starts at 0. Then

$$F(t_0, x) = E(F(s, X(s)) | X(t_0) = x).$$

For simplicity, we restrict to the time-homogeneous case: $\mu(t, x) = \mu(x)$ and $\sigma(t, x) = \sigma(x)$, and assume μ and σ Lipschitz, and $h \in C_0^2$ (*h* twice continuously differentiable, with compact support). Then the stochastic integral is a martingale, and replacing t_0 , s by t, T we get the stochastic representation F(t, x) = E(F(X(T))|X(t) = x) for the solution F. Conversely, any solution F which is in $C^{1,2}$ (has continuous derivatives of order one in t and two in x) and is bounded on compact t-sets arises in this way. This gives:

Theorem (Feynman-Kac Formula. For $\mu(x), \sigma(x)$ Lipschitz, the solution F = F(t, x) to the partial differential equation

$$F_t + \mu F_x + \frac{1}{2}\sigma^2 F_{xx} = 0$$

with final condition F(T, x) = h(x) has the stochastic representation

$$F(t,x) = E[h(X(T))|X(t) = x],$$

where X satisfies the stochastic differential equation

$$dX(s) = \mu(X(s))ds + \sigma(X(s))dW(s) \qquad (t \le s \le T)$$

with initial condition X(t) = x.

The Feynman-Kac formula gives a stochastic representation to solutions of partial differential equations (e.g., the Black-Scholes PDE). *Application.*

One classical application of the Feynman-Kac formula is to Kac's proof of Lévy's arc-sine law for Brownian motion. Let τ_t be the amount of time in [0,t] for which BM takes positive values. Then the proportion τ_t/t has the *arc-sine* law - the law on [0,1] with density $1/(\pi\sqrt{x(1-x)})$ ($x \in [0,1]$).

5. Stochastic Differential Equations

Perhaps the most basic general existence theorem for SDEs is Picard's theorem, for an ordinary differential equation (non-linear, in general)

$$dx(t) = b(t, x(t))dt, \qquad x(0) = x_0,$$

or to use its alternative and equivalent expression as an integral equation,

$$x(t) = x_0 + \int_0^t b(s, x(s)) ds.$$

If one assumes the Lipschitz condition

$$|b(t,x) - b(t,y)| \le K|x - y|$$

for some constant K and all $t \in [0,T]$ for some T > 0, and boundedness of b on compact sets, one can construct a unique solution x by the *Picard iteration*

$$x^{(0)}(t) := x_0, \qquad x^{(n+1)}(t) := x_0 + \int_0^t b(s, x^{(n)}(s)) ds.$$

See any textbook on analysis or differential equations. (The result may also be obtained as an application of Banach's contraction-mapping principle in functional analysis.)

Naturally, stochastic calculus and stochastic differential equations contain all the complications of their non-stochastic counterparts, and more besides. Thus by analogy with PDEs alone, we must expect study of SDEs to be complicated by the presence of more than one concept of a solution. The first solution concept that comes to mind is that obtained by sticking to the non-stochastic theory, and working pathwise: take each sample path of a stochastic process as a function, and work with that. This gives the concept of a *strong* solution of a stochastic differential equation. Here we are given the probabilistic set-up – the filtered probability space in which our SDE arises – and work within it. The most basic results, like their non-stochastic counterparts, assume regularity of coefficients (e.g., Lipschitz conditions), and construct a unique solution by a stochastic version of Picard iteration. Consider the stochastic differential equation

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \qquad X(0) = \xi,$$

where b(t, x) is a *d*-vector of drifts, $\sigma(t, x)$ is a $d \times r$ dispersion matrix, W(t) is an *r*-dimensional Brownian motion, ξ is a square-integrable random *d*-vector independent of W, and we work on a filtered probability space satisfying the usual conditions on which W and ξ are both defined. Suppose that the coefficients b, σ satisfy the following global Lipschitz and growth conditions:

$$\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| \le K \|x - y\|,$$
$$\|b(t,x)\|^2 + \|\sigma(t,x)\|^2 \le K^2 (1 + \|x\|^2),$$

for all $t \ge 0$, $x, y \in \mathbf{R}^d$, for some constant K > 0. **Theorem**. Under the above Lipschitz and growth conditions, (i) the Picard iteration $X^{(0)}(t) := \xi$,

$$X^{(n+1)}(t) := \xi + \int_0^t b(s, X^{(n)}(s)) ds + \int_0^t \sigma(s, X^{(n)}(s)) dW(s)$$

converges, to X(t) say;

(ii) X(t) is the unique strong solution to the stochastic differential equation

$$X(0) = \xi, \qquad X(t) = \xi + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s);$$

(iii) X(t) is square-integrable, and for each T > 0 there exists a constant C, depending only on K and T, such that X(t) satisfies the growth condition

$$E(||X(t)||^2) \le C(1 + E(||\xi||^2))e^{Ct} \quad (0 \le t \le T).$$

Unfortunately, it turns out that not all SDEs have strong solutions. However, in many cases one can nevertheless solve them, by setting up a filtered probability space for oneself, setting up an SDE of the required form on it, and solving the SDE there. The resulting solution concept is that of a *weak* solution. Naturally, weak solutions are distributional, rather than pathwise, in nature. However, it turns out that it is the weak solution concept that is often more appropriate for our purposes. This is particularly so in that we will often be concerned with convergence of a sequence of (discrete) financial models to a (continuous) limit. The relevant convergence concept here is that of *weak convergence*. In the continuous setting, the price dynamics are described by a stochastic differential equation, in a discrete setting by a stochastic difference equation. One seeks results in which weak solutions of the one converge weakly to weak solutions of the other.

The Ornstein-Uhlenbeck Process.

The most important example of a stochastic differential equation for us is that for geometric Brownian motion. We close here with another example.

Consider a model of the velocity V(t) of a particle at time t ($V(0) = v_0$), moving through a fluid or gas, which exerts a force on the particle:

(i) a frictional drag, assumed proportional to the velocity,

(ii) a noise term resulting from the random bombardment of the particle by the molecules of the surrounding fluid or gas.

The basic model for processes of this type is given by the (linear) stochastic differential equation

$$dV = -\beta V dt + \sigma dW,$$

whose solution is called the Ornstein-Uhlenbeck (velocity) process with relaxation time $1/\beta$ and diffusion coefficient $D := \frac{1}{2}\sigma^2/\beta^2$. It is a stationary Gaussian Markov process (not stationary-increments Gaussian Markov like Brownian motion), whose limiting (ergodic) distribution is $N(0, \beta D)$ (this is the classical Maxwell-Boltzmann distribution of statistical mechanics) and whose limiting correlation function is $e^{-\beta|.|}$.

If we integrate the Ornstein-Uhlenbeck velocity process to get the Ornstein-Uhlenbeck displacement process, we lose the Markov property (though the process is still Gaussian). Being non-Markov, the resulting process is much more difficult to analyze.

The Ornstein-Uhlenbeck process is important in many areas, including: (i) statistical mechanics, where it originated,

(ii) mathematical finance, where it appears in the *Vasicek model* for the termstructure of interest-rates.

 $e^{-\beta t}$ solves the corresponding homogeneous DE $dV = -\beta V dt$. So by

variation of parameters, take a trial solution $V = Ce^{-\beta t}$. Then

$$dV = -\beta C e^{-\beta t} dt + e^{-\beta t} dC = -\beta V dt + e^{-\beta t} dC,$$

so V is a solution of (OU) if $e^{-\beta t}dC = \sigma dW$, $dC = \sigma e^{\beta t}dW$, $C = c + \int_0^t e^{\beta u}dW$. So with initial velocity v_0 ,

$$V = v_0 e^{-\beta t} + \sigma e^{-\beta t} \int_0^t e^{\beta u} dW_u.$$

This approach to solving linear SDEs can be generalized.

6. Semi-martingales.

The martingale concept, though crucial, is a little too restrictive, and one needs to generalize it. We will be brief here. First, a local martingale M = (M(t)) is a process such that, for some sequence of stopping times $S_n \to \infty$, each stopped process $M^{(n)} = (M(t \land S_n))$ is a martingale. This localization idea can be applied elsewhere: a process (A(t)) (adapted to our filtration, understood) is locally of finite variation if each $(A(t \land S_n))$ is of finite variation for some sequence of stopping times $S_n \to \infty$. A semi-martingale (Meyer, 1976) is a process (X(t)) expressible as

$$X(t) = M(t) + A(t)$$

with (M(t)) a local martingale and (A(t)) locally of finite variation (the concept is due to Meyer).

Lévy Processes as Semi-martingales.

The Gaussian component $X^{(1)}$ is a martingale; so too is the compensated sum of (small) jumps process $X^{(3)}$, while the sum of large jumps process $X^{(2)}$ is (locally) of finite variation, being compound Poisson. Thus a Lévy process $X = X^{(1)} + X^{(2)} + X^{(3)}$ is a semi-martingale. Indeed, Lévy processes are the prototypes, and motivating examples, of semi-martingales. The natural domain of stochastic integration is predictable integrands and semi-martingale integrators. Thus, stochastic integration works with a general Lévy process as integrator. Here, however, the theory simplifies considerably. *Previsible (= Predictable) Processes.*

The crucial difference between left-continuous (e.g., càglàd) functions and right-continuous (e.g., càdlàg) ones is that with *left*- continuity, one can 'predict' the value at t – 'see it coming' – knowing the values before t.

We write \mathcal{P} , called the *predictable* (or previsible) σ -algebra, for the σ -

algebra on $\mathbf{R}_+ \times \Omega$ (\mathbf{R}_+ for time $t \ge 0$, Ω for randomness $\omega \in \Omega$ – we need both for a stochastic process $X = (X(t, \omega))$) for the σ -field generated by (= smallest σ -field containing) the adapted càglàd processes. (We shall almost always be dealing with adapted processes, so the operative thing here is the *left*-continuity.) We also write $X \in \mathcal{P}$ as shorthand for 'the process X is \mathcal{P} -measurable', and $X \in b\mathcal{P}$ if also X is bounded.

Predictability and Semi-Martingales.

Let us confess here why we need to introduce the last two concepts. One can develop a theory of stochastic integrals, $\int_0^t H(s,\omega)dM(s,\omega)$ or $\int_0^t H_s dM_s$, where H, M are stochastic processes and the integrator M is a semimartingale, the integrand H is previsible (and bounded, or L_2 , or whatever). This can be done; see e.g. [P] for details. More: this theory is the most general theory of stochastic integration possible, if one demands even reasonably good properties (appropriate behaviour under passage to the limit, for example). For emphasis:

Integrands: previsible; Integrators: semimartingales.

Prototype: H is left-continuous (and bounded, or L_2 , etc.); M is BM. Economic Interpretation. Think of the integrator M as, e.g., a stock-price process. The increments over [t, t+u] (u > 0, small) represent 'new information'. Think of the integrand H as the amount of stock held. The investor has no advance warning of the price change $M_{t+dt} - M_t$ over the immediate future [t, t + dt], but has to commit himself on the basis of what he knows already. So H needs to be predictable at H before t (e.g., left- continuity will do), hence predictability of integrands. By contrast, $M_{t+dt} - M_t$ represents new price-sensitive information, or 'driving noise'. The value process of the portfolio is the limit of sums of terms such as $H_{t-}(M_{t+dt} - M_t)$, the stochastic integral $\int_0^t H_s dM_s$. This is the continuous-time analogue of the martingale transform in discrete time.

Poisson Stochastic Calculus.

Recall that the prototypes of Lévy processes are Brownian motion and the Poisson process, also that the essence of Itô calculus for BM is $(dW_t)^2 = dt$. Now the Poisson process N is a point process with jumps of size 1, so $(dN_t)^2 = dN_t$ (both sides are 1 at a jump and 0 elsewhere). This suggests that a Poisson-based stochastic calculus can be developed, and indeed it can. Lévy stochastic calculus.

With both Brownian and Poissonian calculus to hand, this suggests that stochastic calculus for Lévy processes can be developed – and indeed it can. For, Lévy processes are semimartingales, and we saw above that stochastic calculus has as its natural domain that of predictable integrands and semimartingale integrators. The resulting Lévy calculus is very flexible and useful, but we cannot develop it here. It extends Black-Scholes theory to allow prices to have *jumps*, which they do in reality if looked at closely enough. *Lévy finance.*

We close with some comments on the use of Lévy processes for modelling in mathematical finance. There are three main objections to the use of Brownian-based models, as in Black-Scholes theory.

(i) Gaussian distributions are symmetric, and have extremely thin tails. Real financial data show skew, and have much fatter tails than Gaussian. For example, with return distributions on stock, the tail behaviour depends on the length of the return interval. For monthly returns, say, returns are approximately Gaussian. This is because of *aggregational Gaussianity*: the Central Limit Theorem applies. The rule of thumb is that 16 trading days suffice here. High-frequency ('tick') data typically gives heavy tails – tails decreasing like a power; daily returns are intermediate (e.g., hyperbolic distributions).

(ii) Brownian models are *complete* (see the Brownian Martingale Representation Theorem, above). Real markets are *incomplete*. One can see this in, e.g., the *bid-ask spread* – real prices are not unique, but fill an interval.

(iii) Brownian motion is continuous, but real prices *jump*. This is partly because prices are quoted in terms of money, which is quantised. Also, the very act of trading shifts prices, as it affects the current balance of supply and demand. In Black-Scholes theory, one assumes that financial agents are price takers and not price makers – true to a good approximation for small traders (or small trades), but not for large ones. Where there is no trading, there is no price. Where there is trading, there are prices rather than a price. Take, for instance, the price evolution of a heavily traded (and so highly liquid) stock under normal market conditions. There will be very many individually small trades, resulting in what is called *jitter*. Lévy processes of *infinite activity* – infinitely many jumps in finite time – are well suited to modelling such things. What was once pure Probability Theory for its own sake has now become an everyday modelling tool for the financial practitioner.

The mathematics of markets under crisis conditions is of course very interesting and topical, but we cannot develop it here.

NHB, March 2012