pfsl25(15).tex
Lecture 25. 8.12.2015

Interpretation. Martingales model fair games. Submartingales model favourable games. Supermartingales model unfavourable games.

Martingales represent situations in which there is no drift, or tendency, though there may be lots of randomness. In the typical statistical situation where we have *data = signal + noise,* martingales are used to model the noise component. It is no surprise that such decompositions occur constantly in more advanced work (with 'semi-martingales').

Closed martingales.

As before, some martingales are of the form

$$X_t = E[X|\mathcal{F}_t] \qquad (t \geq 0)$$

for some integrable random variable X. Then X is said to *close* (X_t), which is called a *closed* (or *closable*) martingale, or a *regular* martingale. As before, closed martingales have specially good convergence properties:

$$X_t \to X_\infty \qquad (t \to \infty) \qquad \text{a.s. and in } \ L_1,$$

and then also

$$X_t = E[X_\infty|\mathcal{F}_t], \qquad a.s.$$

Again, this property is equivalent also to *uniform integrability* (UI):

$$\sup_t \int_{\{|X_t|>x\}} |X_t| dP \to 0 \qquad (x \to \infty).$$

These are the mgs that are crucial in mathematical finance. Here, the closing random variable is the *payoff* of the option. The option price is what one would expect – the (conditional) expectation of the payoff, given what one knows. This intuition is exactly right (and part of the crucial Fundamental Theorem of Asset Pricing), *provided* that one can bring martingale theory to bear. For this, one needs to change from the real-world measure to the *equivalent martingale measure* (EMM) – the measure making discounted prices martingales (recall: EMM exists iff no arbitrage; EMM unique iff market complete).

Doob-Meyer Decomposition.

One version in continuous time of the Doob decomposition in discrete time – called the Doob-Meyer (or the Meyer) decomposition – follows next but needs one more definition. A process X is called of class (D) if

$$\{X_\tau : \tau \text{ a finite stopping time}\}$$

is uniformly integrable. Then a (càdlàg, adapted) process Z is a submartingale of class (D) if and only if it has a decomposition

$$Z = Z_0 + M + A$$

with M a uniformly integrable martingale and A a predictable increasing process, both null at 0. This composition is unique.

Among the contrasts with the discrete case, we mention that the Doob-Meyer decomposition, easy in discrete time, is a deep result in continuous time.

Square-integrable Martingales.

For $M = (M_t)$ a martingale, write $M \in \mathcal{M}^2$ if M is L_2-bounded:

$$\sup_t E(M_t^2) < \infty,$$

and $M \in \mathcal{M}_0^2$ if further $M_0 = 0$. Write $c\mathcal{M}^2$, $c\mathcal{M}_0^2$ for the subclasses of continuous M.

As before, L_p-bounded mgs are convergent for $p > 1$. So for $M \in \mathcal{M}^2, M$ is convergent:

$$M_t \to M_\infty \qquad \text{a.s. and in mean square}$$

for some random variable $M_\infty \in L_2$. One can recover M from M_∞ by

$$M_t = E[M_\infty|\mathcal{F}_t].$$

The bijection

$$M = (M_t) \leftrightarrow M_\infty$$

is in fact an isometry, and as $M_\infty \in L_2$, a Hilbert space, so too is $\mathcal{M}^2$.

Quadratic Variation.

A non-negative right-continuous submartingale is of class (D). So it has a Doob-Meyer decomposition. We specialize this to X^2, with $X \in c\mathcal{M}^2$:

$$X^2 = X_0^2 + M + A,$$

with M a continuous martingale and A a continuous (so predictable) and increasing process. We write

$$\langle X \rangle := A$$

here, and call $\langle X \rangle$ the *quadratic variation* of X. We shall see later that this is a crucial tool for the stochastic integral. There is a variant on $\langle X \rangle$ (the 'angle-bracket process'), called $[X]$ (the 'square-bracket process'), needed to handle jumps.

Quadratic Covariation.

We write $\langle M, M \rangle$ for $\langle M \rangle$, and extend $\langle . \rangle$ to a bilinear form $\langle ., . \rangle$ with two different arguments by the *polarization identity*:

$$\langle M, N \rangle := \frac{1}{4}(\langle M + N, M + N \rangle - \langle M - N, M - N \rangle.$$

(The polarization identity reflects the Hilbert-space structure of the inner product $\langle ., . \rangle$.) If N is of finite variation, $M \pm N$ has the same quadratic variation as M, so $\langle M, N \rangle = 0$.

Where there is a Hilbert-space structure, one can use the language of projections, of Pythagoras' theorem etc., and draw diagrams as in Euclidean space. The right way to treat the Linear Model of statistics is in such terms (analysis of variance = ANOVA, sums of squares etc.)

L_1, L_2 *and* L_p.

We quote from Functional Analysis: for $p \in (1, \infty)$, define the *conjugate index* $q \in (1, \infty)$ by

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then L_p and L_q are *dual*: each continuous linear functional on L_p can be identified with a function $g \in L_q$, acting on functions $f \in L_p$ by

$$f \mapsto (f, g) := \int fg$$

($fg \in L_1$, by Hölder's inequality). So for $p = 2$, $q = 2$ also: L_2 is *self-dual.* L_2 is *Hilbert space*, H, which has an *inner product*, $(f, g) := \int fg$ (or $(f, g) := \int f\overline{g}$ in the complex case). This is one reason why L_2 is the nicest of the L_p-spaces, and why L_p for $p \in (1, \infty)$ is nicer than L_1.

For $p > 1$, L_p-mgs are UI, and so 'nice'. For $p = 1$, this no longer holds: what is needed instead is the "$L \log L$" condition,

$$E[|X| \log^+ |X|] < \infty.$$

Also important in Functional Analysis are the *Hardy spaces*, H_p. H_p can be identified with a subspace of L_p. For $p \in (1, \infty)$, the dual of H_p is H_q, as with L_p-spaces. But H_1 has dual BMO, the space of functions of *bounded mean oscillation*, which has many connections with martingale theory.

4. Other classes of process

Gaussian Processes

A vector $\mathbf{X} \in \mathbb{R}^n$ has the *multivariate normal distribution* in n dimensions if all linear combinations $\mathbf{a}'\mathbf{X} = \sum_{i=1}^n a_i X_i$ of its components are normally distributed (in one dimension). Such a distribution is determined by a vector μ of means and a non-negative definite $n \times n$ matrix $\mathbf{\Sigma}$ of covariances, and is written $N(\mu, \mathbf{\Sigma})$. Then $\mathbf{X}$ has distribution $N(\mu, \mathbf{\Sigma})$ if and only if it has characteristic function

$$\phi_{\mathbf{X}}(\mathbf{t}) := E[\exp\{i\mathbf{t}' \cdot \mathbf{X}\}] = \exp\{i\mathbf{t}' \cdot \mu - \frac{1}{2}\mathbf{t}'\mathbf{\Sigma}\mathbf{t}\} \qquad (\mathbf{t} \in \mathbb{R}^n).$$

Further, if $\mathbf{\Sigma}$ is positive definite (so non-singular), $\mathbf{X}$ has density

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\mathbf{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right\},$$

by Edgeworth's formula. A process $X = (X(t))_{t \geq 0}$ is *Gaussian* if all its finite-dimensional distributions are Gaussian. Such a process can be specified by:
(i) a measurable function $\mu = \mu(t)$ with $E(X(t)) = \mu(t)$, the *mean function*;
(ii) a non-negative definite function $\sigma(s,t)$ with $\sigma(s,t) = cov(X(s), X(t))$, the *covariance function*.

Gaussian processes have many interesting properties. Among these, we quote *Belayev's dichotomy*: with probability one, the paths of a Gaussian process are either continuous, or extremely pathological: for example, unbounded above and below on any time interval, however short. Naturally, we shall confine ourselves here to continuous Gaussian processes.

Examples.

1. *Brownian motion* (VI.1): mean 0, covariance $\sigma(s,t) := \min(s,t)$.
2. *Brownian bridge* ('Brownian motion started at 0 and conditioned to be at 0 at time 1'): mean 0, covariance $\min(s,t) - st$.
3. *Ornstein-Uhlenbeck process* (the prototypical stationary Gaussian Markov process): mean 0, covariance $e^{-\beta|s-t|}$.

Markov Processes

X is *Markov* if for each t, each $A \in \sigma(X(s) : s > t)$ (the 'future') and $B \in \sigma(X(s) : s < t)$ (the 'past'),

$$P(A|X(t), B) = P(A|X(t)).$$