Math 288 - Probability Theory and Stochastic Process

Taught by Horng-Tzer Yau
Notes by Dongryul Kim

Spring 2017

This course was taught by Horng-Tzer Yau. The lectures were given at MWF 12-1 in Science Center 310. The textbook was *Brownian Motion, Martingales, and Stochastic Calculus* by Jean-François Le Gall. There were 11 undergraduates and 12 graduate students enrolled. There were 5 problem sets and the course assistant was Robert Martinez.

Contents

1 January 23, 2017

The textbook is J. F. Le Gall, *Brownian Motion, Martingales, and Stochastic Calculus*.

A Brownian motion is a random function $X(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$. This is a model for the movement of one particle among many in "billiard" dymanics. For instance, if there are many particles in box and they bump into each other. Robert Brown gave a description in the 19th century. $X(t)$ is supposed to describe the trajectory of one particle, and it is random.

To simplify this, we assume $X(t)$ is a random walk. For instance, let

$$S_N = \sum_{i=1}^{N} X_i, \qquad X_i = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2. \end{cases}$$

For non-integer values, you interpolate it linearly and let

$$S_N(t) = \sum_{i=1}^{\lfloor Nt \rfloor} X_i + (Nt - \lfloor Nt \rfloor) X_{\lfloor Nt \rfloor + 1}.$$

Then $\mathbb{E}(S_N^2) = N$. As we let $N \to \infty$, we have three properties:

- $\lim_{N\to\infty} \mathbb{E}\Big(\dfrac{S_N(t)}{\sqrt{N}}\Big)^2 = t$
- $\lim_{N\to\infty} \dfrac{\mathbb{E}(S_N(t)S_N(s))}{\sqrt{N}^2} = t \wedge s = \min(t, s)$
- $S_N(t)/\sqrt{N} \to N(0, t)$

This is the model we want.

1.1 Gaussians

Definition 1.1. X is a standard **Gaussian random variable** if

$$P(X \in A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-x^2/2} dx$$

for a Borel set $A \subseteq \mathbb{R}$.

For a complex variable $z \in \mathbb{C}$,

$$\mathbb{E}(e^{zX}) = \frac{1}{\sqrt{2\pi}} \int e^{zx} e^{-x^2/2} dx = e^{z^2/2}.$$

This is the **moment generating function**. When $z = i\xi$, we have

$$\sum_{k=0}^{\infty} \frac{(-\xi^2)^k}{k!} = e^{-\xi^2/2} = 1 + i\xi\mathbb{E}X + \frac{(i\xi)^2}{2}\mathbb{E}X^2 + \cdots.$$

From this we can read off the moments of the Gaussian $\mathbb{E}X^n$. We have

$$\mathbb{E}X^{2n} = \frac{(2n)!}{2^n n!}, \quad \mathbb{E}X^{2n+1} = 0.$$

Given $X \sim \mathcal{N}(0,1)$, we have $Y = \sigma X + m \sim \mathcal{N}(m, \sigma^2)$. Its moment generating function is given by

$$\mathbb{E}e^{zY} = e^{im\xi} - \sigma^2\xi^2/2.$$

Also, if $Y_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ and $Y_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ and they are independent, then

$$\mathbb{E}e^{z(Y_1+Y_2)} = \mathbb{E}e^{zY_1}\,\mathbb{E}e^{zY_2} = e^{i(m_1+m_2)\xi - (\sigma_1^2+\sigma_2^2)\xi^2/2}.$$

So $Y_1 + Y_2 \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Proposition 1.2. *If $X_n \sim \mathcal{N}(m_n, \sigma_n^2)$ and $X_n \to X$ in L^2, then*

(i) X is Gaussian with mean $m = \lim_{n\to\infty} m_n$ and variance $\sigma^2 = \lim_{n\to\infty} \sigma_n^2$.

(ii) $X_n \to X$ in probability and $X_n \to X$ in L^p for $p \geq 1$.

Proof. (i) By definition, $\mathbb{E}(X_n - X)^2 \to 0$ and so $|\mathbb{E}X_n - \mathbb{E}X| \leq (\mathbb{E}|X_n - X|^2)^{1/2} \to 0$. So $\mathbb{E}X = m$. Similarly because L^2 is complete, we have $\operatorname{Var} X = \sigma^2$. Now the characteristic function of X is

$$\mathbb{E}e^{i\xi X} = \lim_{n\to\infty} e^{i\xi X_n} = \lim_{n\to\infty} e^{im_n\xi - \sigma_n^2\xi^2/2} = e^{im\xi - \sigma^2\xi^2/2}.$$

So $X \sim \mathcal{N}(m, \sigma^2)$.

(ii) Because $\mathbb{E}X^{2p}$ can be expressed in terms of the mean and variance, we have $\sup_n \mathbb{E}|X_n|^{2p} < \infty$ and so $\sup_n \mathbb{E}|X_n - X|^{2p} < \infty$. Define $Y_n = |X_n - X|^p$. Then Y_n is bounded in L^2 and hence is uniformly integrable. Also it converges to 0 in probability. It follows that $Y_n \to 0$ in L^1, which means that $X_n \to X$ in L^p. It then follows that $X_n \to X$ in probability. $\square$

Definition 1.3. We say that X_n is **uniformly integrable** if for every $\epsilon > 0$ there exists a K such that $\sup_n \mathbb{E}[|X_n| \cdot 1_{|X_n|>K}] < \epsilon$.

Proposition 1.4. *A sequence of random variables X_n converge to X in L^1 if and only if X_n are uniformly integrable and $X_n \to X$ in probability.*

1.2 Gaussian vector

Definition 1.5. A vector $X \in \mathbb{R}^n$ is a **Gaussian vector** if for every $n \in \mathbb{R}^n$, $\langle u, X\rangle$ is a Gaussian random variable.

Example 1.6. Define $X_1 \sim \mathcal{N}(0,1)$ and $\epsilon = \pm 1$ with probability $1/2$ each. Let $X_2 = \epsilon X_1$. Then $X = (X_1, X_2)$ is not a Gaussian vector.

Note that $u \mapsto \mathbb{E}[u \cdot X]$ is a linear map and so $\mathbb{E}[u \cdot X] = u \cdot m_X$ for some $m_X \in \mathbb{R}^n$. We call m_X the **mean** of X. If $X = \sum_{i=1}^n X_i e_i$ for an orthonormal normal basis e_i of $\mathbb{R}^n$, then the mean of X is

$$m_X = \sum_{i=1}^N (\mathbb{E}X_i) e_i.$$

Likewise, the map $u \mapsto \operatorname{Var} u \cdot X$ is a quadratic form and so $\operatorname{Var} uX = \langle u, q_X u \rangle$ for some matrix q_X. We all q_X the **covariant matrix** of X. We write $q_X(n) = \langle u, q_X u \rangle$. Then the characteristic function of X is

$$\mathbb{E}e^{zi(u \cdot X)} = \exp(izm_X - z^2 q_X(u)/2).$$

Proposition 1.7. *If $\operatorname{Cov}(X_j, X_k)_{1 \le j,k \le n}$ is diagonal, and $X = (X_1, \dots, X_n)$ is a Gaussian vector, then X_is are independent.*

Proof. First check that

$$q_X(u) = \sum_{i,j=1}^n u_i u_j \operatorname{Cov}(X_i, X_j).$$

Then

$$\begin{aligned} \mathbb{E}\exp(iu \cdot X) = \exp\Big(-\frac{1}{2} q_X(u)\Big) &= \exp\Big(-\frac{1}{2} \sum_{i=1}^n u_i^2 \operatorname{Var}(X_i)\Big) \\ &= \prod_{i=1}^n \exp\Big(-\frac{1}{2} u_i^2 \operatorname{Var}(X_i)\Big). \end{aligned}$$

This gives a full description of X. □

Theorem 1.8. *Let $\gamma : \mathbb{R}^n \to \mathbb{R}^n$ be a positive definite symmetric $n \times n$ matrix. Then there exists a Gaussian vector X with $\operatorname{Cov} X = \gamma$.*

Proof. Diagonalize γ and let $\lambda_j \ge 0$ be the eigenvalues. Choose $v_1, \dots, v_n$ an orthonormal basis of eigenvectors of $\mathbb{R}^n$ and let $w_j = \lambda_j v_j = \gamma v_j$. Choose Y_j to be Gaussians $\mathcal{N}(0, 1)$. Then the vector

$$X = \sum_{j=1}^n \lambda_j v_j Y_j$$

is a Gaussian vector with the right covariance. □

2 January 25, 2017

We start with a probability space $(\Omega, \mathcal{F}, P)$ and joint Gaussians (or a Gaussian vector) $X_1, \dots, X_d$. We have

$$\mathrm{Cov}(X_j, X_k) = \mathbb{E}[X_j X_k] - \mathbb{E}[X_j]\mathbb{E}[X_k], \quad \gamma_X = (\mathrm{Cov}(X_j, X_k))_{jk} = C.$$

Then the probability distribution is given by

$$P_X(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det C}} e^{-\langle x, C^{-1}x\rangle/2}$$

where $C > 0$ and the characteristic function is

$$\mathbb{E}[e^{itx}] = e^{-\langle t, Ct\rangle/2}.$$

2.1 Gaussian process

A (centered) **Gaussian space** is a closed subspace of $L^2(\Omega, \mathcal{F}, P)$ containing only Gaussian random variables.

Definition 2.1. A **Gaussian process** indexed by $T(\in \mathbb{R}_+)$ is one such that $\{X_{t_1}, \dots, X_{t_k}\}$ is always a Gaussian vector.

Definition 2.2. For a collection S of random variables, we denote by $\sigma(S)$ the smallest σ-algebra on Ω such that all variables in S are measurable.

Proposition 2.3. *Suppose $H_1, H_2 \subseteq H$ where H is the centered Gaussian space. Then $H_1 \perp H_2$ if and only if H_1 and H_2 are independent, i.e., the generated σ-algebras $\sigma(H_1)$ and $\sigma(H_2)$ are independent.*

Example 2.4. If we take $X_1 = X$ to be a Gaussian and $X_2 = \epsilon X$ where $\epsilon = \pm 1$ with probability $1/2$ each, then $\mathbb{E}[X_1 X_2] = \mathbb{E}[\epsilon X^2] = 0$. That is, X_1 and X_2 are uncorrelated. But X_1 and X_2 are dependent. This is because $\{X_1, X_2\}$ are not jointly Gaussian.

2.2 Conditional expectation for Gaussian vectors

Lt X be a random variable in a Gaussian space H. If H is finite dimensional with (jointly Gaussian) basis $X_1, \dots, X_d$, then we may write $X = \sum_j a_j X_j$. Let $K \subseteq H$ be a closed subspace. We want to look at the conditional expectation

$$\mathbb{E}[X|\sigma(K)] = \mathbb{E}[X|K],$$

where $\sigma(K)$ is the σ-algebra generated by elements in K.

Definition 2.5. The **conditional expectation** is defined as $\mathbb{E}[X|K] = \varphi(K)$, the measurable function with respect to K so that for any function $g(K)$ we have

$$\mathbb{E}[Xg(K)] = \mathbb{E}[\varphi(K)g(K)], \text{ or alternatively, } \mathbb{E}[(X - \varphi(K))g(K)] = 0,$$

i.e., $\varphi(K)$ is the L_2-orthogonal projection of X to the space $\sigma(K)$.

We also write $\mathbb{E}[X|\sigma(K)] = p_K(X)$. Let $Y = X - p_K(X)$. Note that Y is a Gaussian that is orthogonal to K. Then

$$\mathbb{E}[f(X)|\sigma(K)] = \mathbb{E}[f(Y + p_K(X))|\sigma(K)] = \int f(y + p_K(X)) \frac{e^{-y^2/2\sigma^2}}{\sqrt{2\pi}\sigma}$$
$$= \int f(y) \frac{e^{-(y-p_K(X))^2/2\sigma^2}}{\sqrt{2\pi}\sigma} dy,$$

where σ^2 is the variance of $Y = X - p_K(X)$.

We remark that if X is a Gaussian with $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] = \sigma^2$, then

$$\mathbb{E}[Xf(X)] = \sigma^2 \mathbb{E}[f'(X)]$$

by integration by parts. It follows that $\mathbb{E}[X^{2n}]$ corresponds to the number of ways to pair elements in a set of size $2n$. There is a more general interpretation. In general, we have

$$\mathbb{E}[Yf(X_1, \ldots, X_d)] = \sum_j \mathrm{Cov}(Y, X_j) \mathbb{E}[\partial_j f].$$

3 January 27, 2017

For a Gaussian process X_s indexed by $s \in \mathbb{R}_+$, let us define

$$\Gamma(s,t) = \mathrm{Cov}(X_s, X_t) \geq 0,$$

where being positive semidefinite means

$$\mathrm{Var}\biggl(\int \alpha(s) X_s ds\biggr) = \iint \alpha(s)\Gamma(s,t)\alpha(t) ds dt \geq 0$$

for any α.

Theorem 3.1. *Suppose $\Gamma(s,t)$ is a positive semidefinite function. Then there exists a Gaussian process with Γ as the covariant matrix.*

Theorem 3.2 (Kolmogorov extension theorem)**.** *Given a family of consistent finite dimensional distributions, there exists a measure with this family of finite dimensional distributions as the finite dimensional marginals.*

Consistency means that if we are given the distribution of $(X_{s_1}, X_{s_2}, X_{s_3})$, then the distribution of (X_{s_1}, X_{s_3}) must be the integration. Kolmogorov's theorem is giving you an abstract construction of the measure.

So we are not going to be worry too much about the existence of these infinite dimensional measure spaces.

Example 3.3. Suppose μ is a probability measure on $\mathbb{R}$. Then

$$\Gamma(s,t) = \int e^{i\xi(s-t)} \mu(d\xi)$$

is positive semidefinite. This is because

$$\iiint \alpha(s) e^{i\xi(s-t)} \alpha(t) ds dt \mu(d\xi) = \int \biggl(\int \alpha(s) e^{i\xi s}\biggr)\biggl(\int \alpha(t) e^{-i\xi t}\biggr)\mu(d\xi) \geq 0.$$

3.1 Gaussian white noise

Suppose μ is a σ-finite measure.

Definition 3.4. We call a Gaussian space G a **Gaussian white noise** with intensity μ if

$$\mathbb{E}(G(f)G(g)) = \int fg d\mu$$

for every $f, g \in L^2(\mu)$.

Informally, this is the same as saying $\mathbb{E}[G(x)G(y)] = \delta(x-y)$ with the definition $G(x) = G(\delta_x)$. Then

$$\begin{aligned}\mathbb{E}[G(f)G(g)] &= \mathbb{E}\biggl[\int G(x)f(x)dx \int G(y)g(y)dy\biggr] \\ &= \iint f(x)\mathbb{E}[G(x)G(y)]g(y)dx\,dy = \int f(x)g(x)dx.\end{aligned}$$

Theorem 3.5. *For every μ a σ-finite measure there exists a Gaussian white noise with intensity μ.*

Proof. Assume $L^2(\mu)$ is separable. There then exists an orthonormal basis in $L^2(\mu)$. For any f there exists a presentation $f = \sum_{i=1}^{\infty} \alpha_i f_i$. Now define

$$G(f) = \sum_{i=1}^{\infty} \alpha_i X_i,$$

where $X_1, X_2, \ldots$are independent in $\mathcal{N}(0,1)$. Then we obvious have the desired condition. □

3.2 Pre-Brownian motion

Definition 3.6. Let G be a Gaussian white noise on $(\mathbb{R}^+, dx)$. Define $B_t = G(1_{[0,t]})$. Then B_t is a **pre-Brownian motion**.

Proposition 3.7. *The covariance matrix is $\mathbb{E}[B_t B_s] = s \wedge t$. (Here, $s \wedge t$ denotes* $\min(s,t)$.*)*

Proof. We have

$$\mathbb{E}[B_t B_s] = \mathbb{E}[G([0,t])G([0,s])] = s \wedge t.$$ □

Because $B_t = G([0,t])$, we may formally write $dB_t = G(dt)$. Then we can also write

$$G(f) = \int f(s) dB_s.$$

We can also think

$$G([0,t]) = \int_0^t dB_s = B_t - B_0 = B_t.$$

Roughly speaking, dB_t is the white noise. This idea can be rigorously formulated as that if $0 = t_0 < t_1 < \cdots < t_n$ then $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent Gaussian.

Proposition 3.8. *The following statements are equivalent:*

(i) (X_t) *is a Brownian motion.*

(ii) (X_t) *is a centered Gaussian with covariance $s \wedge t$.*

(iii) $(X_t - X_s)$ *is independent of $\sigma(X_r, r \leq s)$ and $X_t - X_s \sim \mathcal{N}(0, t-s)$.*

(iv) $\{X_{t_i} - X_{t_{i-1}}\}$ *are independent Gaussian with variance $t_i - t_{i-1}$.*

Proof. (ii) $\Rightarrow$ (iii) We have

$$\mathbb{E}[(B_t - B_s)B_r] = t \wedge r - s \wedge r = 0$$

for $r \leq s$. The other implications are obvious. □

Note that the probability density of $B_{t_1}, \ldots, B_{t_n}$ is given by

$$p(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2 - t_1)\cdots}} \prod_i e^{-(x_i - x_{i-1})^2/2(t_i - t_{i-1})}.$$

It is important to understand this as a Wiener measure.

4 January 30, 2017

Let me remind you of some notions. A Gaussian space is a closed linear subspace of $L^2(\Omega, P)$. A random process is a sequence $(X_t)_{t\geq 0}$ of random variables. A Gaussian process is a process such that $(X_{t_1}, \dots, X_{t_k})$ is Gaussian for any choice $t_1, \dots, t_k$. A Gaussian white noise with intensity μ is a map $G : L^2(\mu) \to L^2(P)$ such that $G(f)$ is Gaussian and $\mathbb{E}[G(f)G(g)] = \int fgd\mu$. A pre-Brownian motion is a process with $B_t = G(1_{[0,t]})$.

Sometimes people write $B_t(\omega)$. In this case, this ω is an element of the probability space Ω. So then $B_t(\omega)$ is actually something concrete map.

4.1 Kolmogorov's theorem

Theorem 4.1 (Kolmogorov). *Let $I = [0, 1]$ and $(X_t)_{t\in I}$ be a process. Suppose for all $t, s \in I$, $\mathbb{E}[|X_t - X_s|^q] \leq C|t - s|^{1+\epsilon}$. Then for all $0 < \alpha < \epsilon/q$ there exists a "modification" $\tilde{X}$ of X such that*

$$|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)| \leq C_\alpha(\omega)|s - t|^\alpha$$

almost surely.

Here we say that $\tilde{X}$ is a **modification** of X if for every $t \in T$,

$$P(\tilde{X}_t = X_t) = 1.$$

Definition 4.2. $\tilde{X}$ and X are **indistinguishable** if there exists a set N of probability zero such that for all t and $\omega \in N$, $\tilde{X}_t(\omega) = X_t(\omega)$.

There is a slight difference, in the case of indistinguishability, N is some set that does not depend on t. That is, indistinguishable implies being a modification. If X_t and $\tilde{X}_t$ have continuous sample paths with probability 1, the converse holds. But nobody worries about these.

For a pre-Brownian motion, we have $X_t - X_s \sim \mathcal{N}(0, t - s)$ and so

$$\mathbb{E}[|X_t - X_s|^q] = C_\alpha|t - s|^{q/2}.$$

Then letting $\epsilon = q/2 - 1$ and $\alpha < \epsilon/q = 1/2 - 1/q$, we get $\alpha \nearrow 1/2$ as $q \to \infty$. That is, $t \to X_t(\omega)$ is Hölder continuous of order ϵ.

Corollary 4.3. *A Brownian motion sample path is Hölder of order $\alpha < 1/2$.*

Proof of Theorem 4.1. Let $D_n = \{0/2^n, 1/2^n, \dots, (2^n - 1)/2^n\}$. We have

$$\begin{aligned} P\biggl(\bigcup_{i=1}^{2^n}(|X_{i/2^n} - X_{(i-1)/2^n}| \geq 2^{-n\alpha})\biggr) &\leq 2^n\bigl(\mathbb{E}|X_{i/2^n} - X_{(i-1)/2^n}|^q\bigr)2^{n\alpha q} \\ &\leq C2^{nq\alpha-(1+\epsilon)n}2^n = C2^{nq\alpha-\epsilon n}. \end{aligned}$$

So $\epsilon - q\alpha > 0$ implies

$$\sum_n P\left(\bigcup_i (|X_{i/2^n} - X_{(i-1)/2^n}| \geq 2^{-n\alpha})\right) < \infty.$$

Borel–Cantelli then implies that there exists a set N with $P(N) = 0$ such that for every $\omega \in N^c$ there exists an $n_0(\omega)$ such that

$$\left|X_{i/2^n}(\omega) - X_{(i-1)/2^n}(\omega)\right| < 2^{-n\alpha}$$

for all $n > n_0(\omega)$. Now

$$K_\alpha(\omega) = \sup_{n\geq 1}\sup_i \left|\frac{X_{i/2^n}(\omega) - X_{(i-1)/2^n}(\omega)}{2^{-n\alpha}}\right| < \infty$$

almost surely. □

We will finish the proof next time.

5 February 1, 2017

We start with Kolmogorov's theorem.

5.1 Construction of a Brownian motion

Proof of Theorem 4.1. We have proved that

$$K_\alpha(\omega) = \sup_{n\geq 1}\sup_i \frac{|X_{i/2^n} - X_{(i-1)/2^n}|}{2^{-n\alpha}} < \infty$$

almost surely using Markov's inequality and the Borel–Cantelli lemma.

We look at the set

$$D = \left\{\frac{i}{2^n} : 0 \leq i \leq 2^n, n = 1, \dots\right\}.$$

Suppose we have a function$f : D \to \mathbb{R}$ such that $|f(i/2^n) - f((i-1)/2^n)| \leq K2^{-n\alpha}$. This condition implies

$$|f(s) - f(t)| \leq \frac{2K}{1-2^{-\alpha}}|t-s|^\alpha$$

for $t, s \in D$. You can prove this using successive approximations.

Now define $\tilde{X}_t(\omega) = \lim_{s\to t, s\in D} X_s(\omega)$ if $K_\alpha(\omega) < \infty$ and otherwise just assign an arbitrary value. You can check that

$$|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)| \leq C_\omega |t-s|^\alpha$$

for some constant C_ω depending only on ω, i.e., $\tilde{X}_t$ is Hölder continuous of order α.

Finally we see that

$$P(|X_t - \tilde{X}_t| > \epsilon) = \lim_{s\to t, s\in D} P(|X_t - X_s| > \epsilon) = 0$$

and so $P(X_t = \tilde{X}_t) = 1$. □

This gives a construction of a **Brownian motion**, which is the modification of the pre-Brownian motion. We first constructed a Gaussian white noise in an abstract space, then constructed a pre-Brownian motion by $X_t = G([0,t])$. Then by Kolmogorov's theorem defined a Brownian motion. Then a sample path is Hölder continuous up to order 1/2.

There is another construction of a Brownian motion, due to Lévy. This construction is Exercise 1.18.

5.2 Wiener measure

The space of continuous paths $C(\mathbb{R}_+, \mathbb{R})$ can be given a structure of a σ-algebra by declaring that $w \mapsto w(t)$ is measurable for all t. Now there is a measurable map

$$\Omega \to C(\mathbb{R}_+, \mathbb{R}), \quad w \mapsto (t \mapsto B_t(\omega)).$$

The **Wiener measure** is the push-forward of this measure. Don't think too hard about what this looks like.

Theorem 5.1 (Blumenthal's zero-one law)**.** *Let $\mathcal{F} = \sigma(B_s, s \leq t)$ and let $\mathcal{F}_{0^+} = \bigcap_{t>0} \mathcal{F}_t$. Then $\mathcal{F}_{0^+}$ is trivial, i.e., for any $A \in \mathcal{F}_{0^+}$, either $P(A) = 0$ or $P(A) = 1$.*

Example 5.2. Let $A = \bigcap_n \{\sup_{0 \leq s \leq n^{-1}} B_s > 0\}$. Then $P(A) = 0$ or $P(A) = 1$. The right answer is $P(A) = 1$. This is because

$$P(A) = \lim_{n\to\infty} P\Big(\sup_{0\leq s\leq n^{-1}} B_s > 0\Big) \geq \lim_{n\to\infty} P(B_{n^{-1}} > 0) = \frac{1}{2}.$$

Sketch of proof. Take a $A \in \mathcal{F}_{0^+}$ and let g be a bounded measurable function. If I can show that

$$\mathbb{E}[1_A g(B_{t_1}, \dots, B_{t_k})] = \mathbb{E}[1_A]\mathbb{E}[g(B_{t_1}, \dots, B_{t_k})]$$

then $\mathcal{F}_{0^+} \perp \sigma(B_t, t > 0)$. Taking the limit, we see that $\mathcal{F}_{0^+}$ is independent of itself. Then we get the result. $\square$

6 February 3, 2017

Using Kolmogorov's theorem we defined a Brownian motion by modifying the pre-Brownian motion. A sample path of a Brownian motion is then Hölder continuous with order α for all $\alpha < 1/2$. Using it we defined the Wiener measure on $C(\mathbb{R}_+, \mathbb{R})$.

6.1 Blumenthal's zero-one law

Let us define

$$\mathcal{F}_0^+ = \bigcap_{t>0} \mathcal{F}_t = \bigcap_{t>0} \sigma(B_s, s \leq t).$$

Theorem 6.1 (Blumenthal). *The σ-algebra $\mathcal{F}_0^+$ is independent of itself. In other words, $P(A) = 0$ or 1 for every $A \in \mathcal{F}_0^+$.*

Proof. Take an arbitrary bounded continuous function g. For an $A \in \mathcal{F}_0^+$, we claim that

$$\mathbb{E}[1_A g(B_{t_1}, \dots, B_{t_k})] = \mathbb{E}[1_A]\mathbb{E}[g(B_{t_1}, \dots, B_{t_k})],$$

for $0 < t_1 < \cdots < t_k$. Let us only work with the case $k = 1$ for simplicity. Note that $B_{t_1} - B_\epsilon$ is independent of $\mathcal{F}_\epsilon$. So we have

$$\mathbb{E}[1_A g(B_{t_1})] = \lim_{\epsilon\to 0} \mathbb{E}[1_A g(B_{t_1} - B_\epsilon)] = \lim_{\epsilon\to 0} P(A)\mathbb{E}[g(B_{t_1})] = \mathbb{E}[1_A]\mathbb{E}[g(B_{t_1})].$$

This implies that $\mathcal{F}_0^+$ is orthogonal to $\sigma(B_t, t > 0)$. So $\mathcal{F}_0^+ \perp \mathcal{F}_t$, but $\mathcal{F}_0^+ \subseteq \mathcal{F}_t$. So $\mathcal{F}_0^+ \perp \mathcal{F}_0^+$. □

Proposition 6.2. *Define $T_a = \inf\{t \geq 0, B_t = a\}$. Then almost surely for all $a \in \mathbb{R}$, $T_a < \infty$. That is, $\limsup_{t\to\infty} B_t = \infty$ and $\liminf_{t\in\infty} B_t = \infty$.*

Proof. We have

$$1 = P\Big(\sup_{0\leq s\leq 1} B_s > 0\Big) = \lim_{\delta\searrow 0} P\Big(\sup_{0\leq s\leq 1} B_s > \delta\Big).$$

But by the scaling property of the Brownian motion, $B_t \sim \lambda^{-1} B_{\lambda^2 t}$, we can write

$$P\Big(\sup_{0\leq s\leq 1} B_s > \delta\Big) = P\Big(\sup_{0\leq s\leq \delta^{-2}} B_s > 1\Big).$$

It follows that $P(\sup_{0\leq s} B_s > 1) = 1$. □

6.2 Strong Markov property

Definition 6.3. A random variable $T \in [0,\infty]$ is a **stopping time** if the set $\{T \le t\}$ is in $\mathcal{F}_t$. Intuitively, a stopping time cannot depend on the future.

Example 6.4. The random variable $T_a = \inf\{t : B_t = a\}$ is a stopping time.

Definition 6.5. We define the σ-field

$$\mathcal{F}_T = \{A : A \cap \{T \le t\} \in \mathcal{F}_t \text{ for all } t \ge 0\}$$

generated by T. Roughly speaking, this is all the information collected up to the stopping time.

Theorem 6.6. *Define $B_t^{(T)} = 1_{\{T<\infty\}} \cdot (B_{T+t} - B_T)$. Then under the measure $P(\cdot : T < \infty)$, the process $B_t^{(T)}$ is a Brownian motion independent of $\mathcal{F}_T$.*

We will prove this next time.

7 February 6, 2017

Theorem 7.1. *Suppose $P(T < \infty) > 0$. Define $B_t^{(T)} = 1_{T<\infty}(B_{T+t} - B_T)$. Then respect to $P(\cdot : T < \infty)$, the process $B_t^{(T)}$ is a Brownian motion independent of $\mathcal{F}_T$.*

Here, $A \in \mathcal{F}_T$ is defined as $A \cap \{T \le t\} \in \mathcal{F}_t$.

Proof. Let $A \in \mathcal{F}_T$. We want to show that for any continuous bounded function F,

$$\mathbb{E}[1_A F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)})] = \mathbb{E}[1_A]\mathbb{E}[F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)}].$$

Define $[t]_n = \min\{k/n^2 : k/n^2 > t\}$. We have $F(B_t^{(T)}) = \lim_{n\to\infty} F(B_t^{([T]_n)})$. Then

$$\begin{aligned} \mathbb{E}[1_A F(B_t^{(T)})] &= \lim_{n\to\infty} \mathbb{E}[1_A F(B_t^{([T_n])})] \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbb{E}[1_A 1_{\{(k-1)2^{-n} < T \le k2^{-n}\}} F(B^{(k/2^n)} t)] \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbb{E}[1_A 1_{\{(k-1)2^{-n} < T \le k2^{-n}\}}] \cdot \mathbb{E}[F(B_t^{(k/2^n)})] \\ &= \cdots = \mathbb{E}[1_A] \cdot \mathbb{E}[F(B_t^{(T)})]. \end{aligned}$$

Here we are using the fact that $1_A 1_{\{(k-1)2^{-n} < T \le k2^{-n}\}}$ is in $\mathcal{F}_{k2^{-n}}$ and so is independent to $B_t^{(k2^{-n})} = B_{k2^{-n}+t} - B_{k2^{-n}}$. □

7.1 Reflection principle

Let $S_t = \sup_{s\le t} B_s$.

Theorem 7.2 (Reflection principle)**.** *We have $P(S_t \ge a, B_t \le b) = P(B_t \ge 2a - b)$.*

The idea is that you can reflect with respect to the first tie you meet a.

Proof. Let T_a = first time $S_t \ge a$. Then $B_t^{(T)}$ is independent of $\mathcal{F}_{T_a}$ by the strong Markov property. This is why reflection works. □

Proposition 7.3. *The probability density of $(S_t = a, B_t = b)$ is*

$$\frac{2(2a-b)}{\sqrt{2\pi t^3}} e^{-(2a-b)^2/2t} 1_{(a>0, b<a)}.$$

The probability density of T_a is

$$f(t) = \frac{a}{\sqrt{2\pi t^3}} e^{-a^2/2t}.$$

Proof. Exercise. We have

$$\begin{aligned}P(T_a \leq t) &= P(S_t \geq a) = P(S_t \geq a, B_t \leq a) + P(S_t \geq a, B_t \geq a) \\ &= 2P(B_t \geq a) = P(tB_1^2 \geq a^2).\end{aligned}$$

Then compute the density. □

Proposition 7.4. *Let* $0 < t_0^n < \cdots < t_{p_n}^n = t$. *Then*

$$\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 \to t$$

in L^2 *as* $n \to \infty$ *and* $\max_i |t_i^n - t_{i-1}^n| \to 0$.

This is saying that $\int_0^t (dB_t)^2 \to t$.

Proof. We have

$$\begin{aligned}\mathbb{E}\Big[\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 - t\Big]^2 &= \mathbb{E}\Big[\sum_{i=1}^{p^n} \big((B_{t_i^n} - B_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big)\Big]^2 \\ &= \sum_{i=1}^{p^n} \mathbb{E}\big[(B_{t_i^n} - B_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)\big]^2 \leq C \sum_{i=1}^{p_n} (t_i^n - t_{i-1}^n)^2 \to 0,\end{aligned}$$

since the $(B_{t_i^n} - B_{t_{i-1}^n})^2 - (t_i^n - t_{i-1}^n)$ are independent with expectation 0. □

7.2 Filtrations

Now we are going to talk about martingales.

Definition 7.5. A **filtration** is an increasing sequence $\mathcal{F}_t$. Define

$$\mathcal{F}_{t^+} = \bigcap_{s>t} \mathcal{F}_s \supseteq \mathcal{F}_t.$$

If $\mathcal{F}_{t^+} = \mathcal{F}_t$, then $\mathcal{F}_t$ is called **right continuous**. By definition, $\mathcal{F}_{t^+}$ is right continuous.

If you work with $\mathcal{F}_{t^+}$, then you are going to be fine.

8 February 8, 2017

The first homework is 1.18, 2.29, 2.31, 2.32.

Last time we talked about the strong Markov property for Brownian motion. This implies the reflection principle, and so we were able to effectively compute the distributions of $\max_{0\leq s\leq t} B_s$.

8.1 Stopping time of filtrations

A filtration consists of an increasing sequence $(\mathcal{F}_t)_{t=0}^{\infty}$, and we define $\mathcal{F}_t^+ = \bigcap_{s>t}\mathcal{F}_s = \mathcal{G}_t$.

Definition 8.1. X_t is **adapted** with respect to $(\mathcal{F}_t)$ if X_t is $\mathcal{F}_t$ measurable. X_t is **progressively measurable** if $(\omega, s)\mapsto X_s(\omega)$ is measurable with respect to $\mathcal{F}_t \times \mathscr{B}([0,t])$.

Proposition 8.2. *If the sample path is right continuous almost surely, then X_t is adapted if and only if X_t is progressively measurable.*

Proof. It is obvious that progressively measurable implies adapted. Assume that it is adapted. Fix a time t. Define

$$X_s^n = X_{kt/n} \quad \text{for } s \in \Big[\frac{(k-1)t}{n}, \frac{kt}{n}\Big).$$

Then X_s^n is now discrete, and $\lim_{n\to\infty} X_s^n = X_s$ almost surely. Clearly $(\omega, s)\to X_s^n(\omega)$ is measurable with respect to $\mathcal{F}_t\times\mathscr{B}([0,t])$. $\square$

A **stopping time** is a random variable T such that $\{T\leq t\}\in\mathcal{F}_t$. We similarly define $\mathcal{F}_T = \{A : A\cap\{T\leq t\}\in\mathcal{F}_t\}$.

Proposition 8.3. *T is a stopping time with respect $\mathcal{G}_t = \mathcal{F}_t^+$ if and only if $\{T<t\}\in\mathcal{F}_t$ for all $t>0$.*

Proof. We have $\{T\leq t\} = \bigcap_{s<t}\{T<s\}$. Likewise $\{T<t\} = \bigcup_{s<t}\{T\leq s\}$. $\square$

We define $\mathcal{G}_t = \{A : A\cap\{T\leq t\}\in\mathcal{F}_T\} = \mathcal{F}_{T^+}$.

Here are some facts:

- If S and T are stopping times, then $S\vee T$ and $S\wedge T$ are also stopping times.
- If S_n is a sequence of stopping times and increasing, then $S=\lim_{n\to\infty} S_n$ is also a stopping time.
- If S_n is a decreasing sequence of stopping times converging to S, then S is a stopping time with respect to $\mathcal{G}$.

8.2 Martingales

Definition 8.4. A process X_t is a **martingale** if $\mathbb{E}[|X_t|] < \infty$ and $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ for every $t \geq s$. If $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$, then it is called a **submartingale**.

Example 8.5. B_t is a martingale. But B_t^2 is not a martingale because $\mathbb{E}[B_t^2|\mathcal{F}_s] = (t-s) + B_s^2$. But $B_t^2 - t$ is a martingale.

Example 8.6. The process $e^{\theta B_t^2 - \theta^2/2t}$ is a martingale for all θ. This is because

$$\mathbb{E}[e^{\theta B_t^2 - \theta^2 t/2}|\mathcal{F}_s] = \mathbb{E}[e^{\theta(B_t - B_s)^2 + 2\theta(B_t - B_s)B_s}|\mathcal{F}_s]e^{\theta^2 B_s^2 - \theta^2 t/2}.$$

Example 8.7. For any $f(s) \in L^2[0,\infty)$, the process

$$\int_0^t f(s)dB_s = \lim_{n\to\infty} \sum_i f(s_{i-1})(B_{s_i} - B_{s_{i-1}})$$

is a martingale. Note that this is well-defined for simple functions f and you can extend for general $f \in L^2$ because

$$\mathbb{E}\biggl(\int_0^t f(s)dB_s\biggr)^2 = \int_0^t f(s)^2 ds$$

if f is simple. Furthermore,

$$\biggl(\int_0^t f(s)dB_s\biggr)^2 - \int_0^t f(s)^2 ds$$

is also a martingale.

9 February 10, 2017

Chapter VII of Shirayayev, *Probability* is a good reference for discrete martingales.

9.1 Discrete martingales - Doob's inequality

We are now going to look at martingales with discrete time.

Theorem 9.1 (Optional stopping theorem). *Suppose X_n is a submartingale and $\sigma \leq \tau < \infty$ (almost surely) are two stopping times. If $\mathbb{E}[|X_n| 1_{\tau > n}] \to 0$ and $\mathbb{E}[|X_\tau| + |X_\sigma|] < \infty$, then $\mathbb{E}[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma$. If X_n is a martingale, then $\mathbb{E}[X_\tau | \mathcal{F}_\sigma] = X_\sigma$.*

Proof. To prove $\mathbb{E}[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma$, it suffices to prove $\mathbb{E}[X_\tau A] \geq \mathbb{E}[X_\sigma A]$ for all $A \in \mathcal{F}_\sigma$. In this case, A will look like

$$A = \bigcup_{k=1}^{\infty} A_k, \quad \text{where } A_k = A \cap 1_{\sigma = k} \in \mathcal{F}_k.$$

So it suffices to check for A_m where m is fixed.

Define $\tau_n = \tau \wedge n$ and $\sigma_n = \sigma \wedge n$. We automatically have

$$\mathbb{E}[X_\sigma A_m] = \lim_{n \to \infty} \mathbb{E}[X_{\sigma_n} A_m],$$

because the sequence on the right side is eventually constant. Also

$$\begin{aligned} \mathbb{E}[X_\tau A_m] &= \mathbb{E}[X_\tau 1_{\tau > n} A_m] + \mathbb{E}[X_\tau 1_{\tau \leq n} A_m] = \mathbb{E}[X_\tau 1_{\tau > n} A_m] + \mathbb{E}[X_{\tau_n} 1_{\tau \leq n} A_m] \\ &= \mathbb{E}[X_\tau 1_{\tau > n} A_m] - \mathbb{E}[X_n 1_{\tau > n} A_m] + \mathbb{E}[X_{\tau_n} A_m]. \end{aligned}$$

Here as $n \to \infty$, the first two terms goes to 0 because of the condition. It is also clear that $\mathbb{E}[X_{\sigma_n} A_m] \leq \mathbb{E}[X_{\tau_n} A_m]$, because we can only care about finitely many outcomes of σ and τ. So

$$\mathbb{E}[X_\sigma A_m] = \lim_{n \to \infty} \mathbb{E}[X_{\sigma_n} A_m] \leq \lim_{n \to \infty} \mathbb{E}[X_{\tau_n} A_m] = \mathbb{E}[X_\tau A_m]. \qquad \square$$

Theorem 9.2 (Doob inequality). *Let X_j be a nonnegative submartingale, and let $M_n = \max_{0 \leq j \leq n} X_j$.*
(1) For $a > 0$, $aP(M_n > a) \leq \mathbb{E}[1_{M_n \geq a} X_n]$.
(2) for $p > 1$,

$$\|M_n\|_p^p \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_n|^p].$$

Proof. (1) Let $\tau = \min\{j \leq n : X_j \geq a\}$. This is a stopping time, and $\tau = n$ if $M_n < a$. Then

$$\begin{aligned} \mathbb{E}[X_n] \geq \mathbb{E}[X_\tau] &= \mathbb{E}[1_{M_n \geq a} X_\tau] + \mathbb{E}[1_{M_n < a} X_\tau] \\ &\geq aP(M_n \geq a) + \mathbb{E}[1_{M_n < a} X_n]. \end{aligned}$$

(2) It follows from the following lemma. $\square$

Lemma 9.3. *If $\lambda P(Y \geq \lambda) \leq \mathbb{E}[X 1_{Y \geq \lambda}]$ for $X, Y \geq 0$, then*

$$\mathbb{E}[Y^p] \leq \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[X^p].$$

Proof. We have

$$\begin{aligned}\mathbb{E}[Y^p] &= \int_0^\infty p\lambda^{p-1} P(Y \geq \lambda) d\lambda \leq \mathbb{E} \int_0^\infty X p \lambda^{p-2} 1_{Y \geq \lambda} d\lambda \\ &= \mathbb{E}\Big[X \frac{p}{p-1} Y^{p-1}\Big] \leq \frac{p}{p-1} \|X\|_p \|Y\|_p^{p-1}.\end{aligned}$$

□

Theorem 9.4 (Doop upcrossing inequality)**.** *Suppose X_n is a submartingale. Let $\beta(a,b)$ be the number of upcrossings, i.e., the number times the martingale goes from below a to above b. Then*

$$\mathbb{E}[\beta(a,b)] \leq \frac{\mathbb{E}[\max(X_n - a, 0)]}{b - a}.$$

Note that if X is a submartingale, then $\max(X - a, 0)$ is a submartingale. This is because

$$\mathbb{E}[f(X_n) | \mathcal{F}_{n-1}] \geq f(\mathbb{E}[X_n | \mathcal{F}_{n-1}]) \geq f(X_{n-1}).$$

10 February 13, 2017

10.1 More discrete martingales - upcrossing inequality

In the optional stopping theorem, we needed the stopping time τ to satisfy the inequality $\mathbb{E}[X_n 1_{\tau>n}] \to 0$ as $n \to \infty$.

Proposition 10.1. *If X_n is uniformly integrable, i.e.,*

$$\lim_{M\to\infty} \sup_n \int_{|X_n|\geq M} |X_n| = 0,$$

then $\mathbb{E}[X_n 1_{\tau>n}] \to 0$.

Proof. We have

$$\begin{aligned}\lim_{n\to\infty} \int |X_n| 1_{\tau>n} &\leq \lim_{n\to\infty} M 1_{\tau>n} + \lim_{n\to\infty} \int |X_n| 1_{|X_n|\geq M} 1_{\tau>n} \\ &\leq \lim_{n\to\infty} |X_n| 1_{|X_n|\geq M} = 0.\end{aligned}$$

□

Corollary 10.2. *Let X_n be a submartingale. Suppose (1) $\mathbb{E}[|X_0|] < \infty$, (2) $\mathbb{E}[\tau] < \infty$, and (3) $\sup_n \mathbb{E}[|X_{n+1} - X_n||\mathcal{F}_n] \leq C$. Then $\mathbb{E}[X_\tau] \geq \mathbb{E}[X_0]$ and $\mathbb{E}[X_\tau|\mathcal{F}_0] \geq X_0$.*

Proof. We want to check the three conditions in the optional stopping inequality. We have

$$\begin{aligned}\mathbb{E}[|X_\tau - X_m|] \leq \sum_{n=m}^{\infty} \mathbb{E}[|X_{n+1} - X_n| 1_{\tau\geq n+1}] &= \sum_{n=m}^{\infty} \mathbb{E}[1_{\tau\geq n+1} \mathbb{E}[|X_{n+1} - X_n|\mathcal{F}_n|] \\ &\leq C \sum_{n=m}^{\infty} \mathbb{E}[1_{\tau\geq n+1}] \leq C\mathbb{E}[\tau] < \infty.\end{aligned}$$

This shows $\mathbb{E}[|X_\tau|] < \infty$. We also have

$$\lim_{n\to\infty} \mathbb{E}[|X_n| 1_{\tau\geq n+1}] \leq \lim_{n\to\infty} \mathbb{E}[|X_\tau - X_n|] + \lim_{n\to\infty} \mathbb{E}[|X_\tau| 1_{\tau\geq n+1}] = 0.$$

□

Theorem 10.3 (Upcrossing inequality)**.** *For $a < b$, let $\beta(a, b)$ be the number of upcrossings in $(X_j)_{j=1}^n$. Then*

$$\mathbb{E}[\beta(a, b)] \leq \frac{\mathbb{E}[\max(X_n - a, 0)]}{b - a}.$$

Proof. Replace X_j by $\max(X_j - a, 0)$. This is still a submartingale, and so we can assume $X_i \geq 0$ and $a = 0$.

Define a random variable ϕ_i such that $\phi_i = 1$ if there exists a $k \leq i$ such that $X_k = 0$ and $X_j < b$ for all $k \leq j \leq i$. Then we get

$$\begin{aligned}\mathbb{E}[\beta(a, b)] \leq b^{-1}\mathbb{E} \sum_{i=1}^{n-1} \phi_i (X_{i+1} - X_i) &= b^{-1} \sum_i \mathbb{E}[\phi_i \mathbb{E}[X_{i+1} - X_i|\mathcal{F}_i]] \\ &\leq b^{-1} \sum_i \mathbb{E}[X_{i+1} - X_i] = b^{-1}\mathbb{E}[X_n],\end{aligned}$$

because $\mathbb{E}[X_{i+1} - X_i | \mathcal{F}_i] \geq 0$ and so we can replace ϕ_i by 1. $\square$

Theorem 10.4 (Doob's convergence theorem)**.** *If X_n is a submartingale with $\sup_i \mathbb{E}[|X_i|] < \infty$, then X_n converges almost surely to some X.*

Proof. We have

$$P(\limsup X_n > \liminf X_n) = P \bigcup_{\substack{a<b \\ a,b\in\mathbb{Q}}} (\limsup X_n > b > a > \liminf X_n) = 0$$

by the upcrossing inequality. $\square$

Note that $X_n \to X$ in L^1 may fail. For instance, take Y_i be independent Bernoulli with $\{0, 2\}$ and $X_n = Y_1 \cdots Y_n \to 0$. Then $\mathbb{E}X_n = 1$.

Proposition 10.5. *(1) Suppose $X_n \to X$ almost surely and $\{X_n\}$ is uniformly integrable. Then $X_n \to X$ in L^1.*
(2) Suppose $X_n \geq 0$, $X_n \to X$ almost surely, and $\mathbb{E}X_n = \mathbb{E}X$. Then $X_n \to X$ in L^1.

Theorem 10.6 (Levi)**.** *Suppose $X \in L^1$ be a random variable and let $\mathcal{F}_n$ be a filtration with $\mathcal{F}_\infty = \bigcup \mathcal{F}_n$. Then $\mathbb{E}[X|\mathcal{F}_n] \to \mathbb{E}[X|\mathcal{F}_\infty]$ almost surely and L^1. Conversely, if $\{X_n\}$ is uniformly integrable and $\mathcal{F}_\infty$ is the full σ-algebra $\mathcal{F}$, then there exists a $X \in L^1$ such that $X_n = \mathbb{E}[X|\mathcal{F}_n]$.*

11 February 15, 2017

So far we have looked that Doob's inequality, the optional stopping time, and Doob's upcrossing inequality. There is a supermartingale version of the upcrossing inequality:

Theorem 11.1. *For a discrete supermartingale X, the number of upcrossings M^n_{ab} has expectation value*

$$\mathbb{E}[M^n_{ab}(x)] \leq \mathbb{E}\Big[\frac{\max\{0, a - x_n)\}}{b-a}\Big].$$

Proof. Let

$$\tau_1 = n \wedge \min\{j > 0 : x_j \leq a\}, \quad \tau_2 = n \wedge \min\{j > \tau_1 : x_j \geq b\},$$

and so forth. Then all τ_k are stopping times. Let

$$D = (X_{\tau_2} - X_{\tau-1}) + (X_{\tau_4} - X_{\tau_3}) + \cdots .$$

Then $(b-a)M^n_{ab} + R \leq D$, where $R = 0$ if the tail is a down-crossing and $R = X_n - a$ otherwise. Because X is a supermartinagle, we have $\mathbb{E}[X_{\tau_2} - X_{\tau_1}] \leq 0$ and so

$$(b-a)\mathbb{E}M^n_{ab} \leq -\mathbb{E}R \leq \mathbb{E}[\max(0, a - X_n)].$$

□

11.1 Levi's theorem

Theorem 11.2 (Levi)**.** *Suppose $X \in L^1$ and $\mathcal{F}_n$ be a filtration, with $\mathcal{F}_\infty = \bigcup \mathcal{F}_n$. Then $\mathbb{E}[X|\mathcal{F}_n] \to \mathbb{E}[X|\mathcal{F}_\infty]$ in L^1 and almost surely. Conversely, if a martingale X_n is uniformly integrable, then there is an X such that $\mathbb{E}[X|\mathcal{F}_n] = X_n$.*

Proof. Let $X_n = \mathbb{E}[X|\mathcal{F}_n]$. We claim that X_n is uniformly integrable, and you can check this. Also X_n is a martingale. So the convergence theorem tells us that $X_n \to X_\infty$ almost surely and in L^1. Now the question is whether $X_\infty = \mathbb{E}[X|\mathcal{F}_\infty]$. Because X_n is a martingale, for any $A \in \mathcal{F}_n$ and $m \geq n$,

$$\int_A X_n d\mu = \int_A X_m d\mu.$$

Taking $m \to \infty$, we get

$$\int_A X d\mu = \int_A X_n d\mu = \int_A X_\infty d\mu$$

because $X_m \to X_\infty$ in L^1. So this is true for all $A \in \mathcal{F}_\infty$. So $X_\infty = \mathbb{E}[X|\mathcal{F}_\infty]$.

Now let us do the second half. Suppose X_n is uniformly integrable. Then $X_n \to X$ for some X almost surely and in L^1. So basically by the same argument,

$$\int_A X_n d\mu = \lim_{m\to\infty} \int_A X_m d\mu = \int_A X d\mu.$$

This shows that $X_n = \mathbb{E}[X|\mathcal{F}_n]$. □

So essentially all martingales come from taking a filtration and looking at the conditional expectations.

11.2 Optional stopping for continuous martingales

Theorem 11.3. *Suppose X_t is a submartingale with respect to a right continuous filtration. If $\sigma \leq \tau \leq C$, then $\mathbb{E}[X_\tau|\mathcal{F}_\sigma] \geq X_\sigma$.*

Proof. Let $\tau_n = (\lfloor n\tau \rfloor + 1)/n$. Then τ_n is a stopping time and $\mathcal{F}_{\tau_n} \downarrow \mathcal{F}_\tau$ as $n \to \infty$, because $\mathcal{F}$ is right continuous. By the discrete version of optional stopping time, we see that $X_{\tau_n} \to X$ implies that $\mathbb{E}[X_C|\mathcal{F}_{\tau_n}] \geq X_{\tau_n}$. Then X_{τ_n} is uniformly integrable and $X_{\tau_n} \to X_\tau$ almost surely and in L^1. Similarly $X_{\sigma_n} \to X_\sigma$. So

$$\begin{aligned}
\mathbb{E}[X_{\tau+\epsilon}|\mathcal{F}_\sigma] &= \lim_{m\to\infty} \mathbb{E}[X_{\tau+\epsilon}|\mathcal{F}_{\sigma_m}] \\
&= \lim_{m\to\infty} \lim_{n\to\infty} \mathbb{E}[X_{\tau_n+\epsilon}|\mathcal{F}_{\sigma_m}] \geq \lim_{m\to\infty} X_{\sigma_m} = X_\sigma. \qquad \square
\end{aligned}$$

12 February 17, 2017

We proved the upcrossing inequality and Doob's inequality. We also proved optional stopping theorem. There is also the martingale convergence theorem that says that if (X_j) is a submartingale or a supermartingale and $\sup_j \mathbb{E}[|X_j|] \leq C$, then $\lim_{J\to\infty} X_j = X_\infty$ for some X_∞ almost surely. If (X_j) is uniformly integrable, then $X_j \to X_\infty$ in L^1.

Theorem 12.1 (3.18)**.** *Suppose $\mathcal{F}_t$ is right continuous and complete. Let X_t be a super-martingale and assume that $t \mapsto \mathbb{E}[t]$ is right continuous. Then X has a modification such that the sample paths are right continuous with left limits, which is also a supermartingale.*

Lemma 12.2 (3.16)**.** *Let D be a dense set and f be defined on D such that*

(i) f is bounded on $D \cap [0,T]$,

(ii) $M^f_{ab}(D \cap [0,T]) < \infty$.

Then $f_+(t) = \lim_{s\downarrow t} f(s)$ and $f_-(t) = \lim_{s\uparrow t} f(s)$ exist for all $t \in D \cap [0,T]$. Define $g(t) = f_+(t)$. Then $g(t)$ is right continuous with left limits.

Theorem 12.3 (3.21)**.** *Suppose a martingale (X_t) has right continuous sample paths. Then the following are equivalent:*

(i) X_t is closed, i.e., there exists a $Z \in L^1$ such that $\mathbb{E}[Z|\mathcal{F}_t] = X_t$.

(ii) X_t is uniformly integrable.

(iii) $X_t \to W$ almost surely and L^1 as $t \to \infty$ for some W.

There is an application. Let $T_a = \inf\{t : B_t = a\}$ for $a > 0$ and $\lambda > 0$.

Proposition 12.4. $\mathbb{E}[e^{-(\lambda^2/2)T_a}] = e^{\lambda a}$.

13 February 22, 2017

Definition 13.1. Let (X_t) by a submartingale or a supermartingale with right continuous sample paths. Assume $X_t \to X_\infty$ almost surely. Suppose T is a stopping time which can take ∞ with positive probability. We define

$$X_T(\omega) = 1_{(T(\omega)<\infty)} X_T(\omega) + 1_{T(\omega)=\infty)} X_\infty(\omega).$$

Theorem 13.2 (Optional stopping theorem)**.** *Let (X_t) be a submartingale. Let $S \le T$ almost surely and $P(T < \infty) = 1$.*
(1) If the submartingale (X_t) is uniformly integrable, then $\mathbb{E}[|X_n| 1_{(T>n)}] \to 0$ and $\mathbb{E}[|X_T| + |X_S|] < \infty$.
(2) If $\mathbb{E}[|X_n| 1_{(T>n)}] \to 0$ and $\mathbb{E}[|X_T| + |X_S|] < \infty$ then $\mathbb{E}[X_T|\mathcal{F}_S] \ge X_S$.
(3) If $S \le T \le C$ almost surely, then $\mathbb{E}[X_T|\mathcal{F}_S] \ge X_S$.
(4) Suppose a martingale (X_n) is uniformly integrable, and let $S \le T$ be stopping times taking values in $\mathbb{R} \cup \{\infty\}$. Then $\mathbb{E}[|X_n| 1_{(T>n)}] \to 0$ and $\mathbb{E}[|X_T| + |X_S|] < \infty$.

Theorem 13.3 (Levi)**.** *Let (X_n) be a uniformly integrable martingale. Then there exists a $X \in L^1$ such that $\mathbb{E}[X|\mathcal{F}_n] = X_n$.*

If X_n is a submartingale and T is a stopping time with $T \in \mathbb{R} \cup \{\infty\}$, then $X_{T\wedge n}$ is a submartingale. That is, for $m < n$, $\mathbb{E}[X_{T\wedge n}|\mathcal{F}_m] \ge X_{T\wedge m}$.

Example 13.4. Let $X_1 = 1$ and X_n be a symmetric random walk. Let T be the first hitting time of 0 and let $Y_n = X_{T\wedge n}$. Then $\mathbb{E}[Y_n] = \mathbb{E}[Y_1] = 1$ but $Y_n \to 0$ almost surely. This means that Y_n does not converge to 0 in L_1, and hence Y_n is not uniformly integrable.

13.1 Submartingale decomposition

Theorem 13.5 (Doob submartingale decomposition)**.** *Suppose (X_n) is a submartingale. Then there exists a martingale M_n and a predictable increasing sequence A_n such that $X_n = M_n + A_n$. Here we say that A_n is* ***predictable*** *if and only if A_n is a measurable with respect to $\mathcal{F}_{n-1}$. If X_n is a supermartingale, then A_n is decreasing. Furthermore, such a decomposition is unique.*

Proof. Because we want $X_2 = M_2 + A_2$, we get $\mathbb{E}[X|\mathcal{F}_1] = M_1 + A_2 = X_1 + A_2$. So $A_2 = \mathbb{E}[X_2|\mathcal{F}_1] - X_1$. Similarly we can define

$$A_k = \mathbb{E}[X_k|\mathcal{F}_{k-1}] - M_{k-1}, \quad M_k = X_k - A_k.$$

You can check that A_k increases. For instance,

$$A_2 = \mathbb{E}[X_2|\mathcal{F}_1] - M_1 = \mathbb{E}[X_2|\mathcal{F}_1] - [X_1 - A_1] = \mathbb{E}[X_2|\mathcal{F}_1] - X_1 + A_1. \quad \square$$

14 February 24, 2017

There is another version of optional stopping time

Theorem 14.1 (Le Gall 3.25). *Suppose X is a nonnegative supermartingale with right continuous sample paths. Let $S \leq T$ be two stopping times. Then $\mathbb{E}[X_T|\mathcal{F}_S] \geq X_S$.*

Proof. Consider only the discrete case, since we can approximate the continuous case with the discrete case. A nonnegative supermartingale is in L^1 in the sense that $\mathbb{E}[X_j] \leq C < \infty$. The upcrossing inequality tells us that $X_j \to X_\infty = X$ almost surely (but maybe not in L^1).

For any $M > 0$, we know that $S \wedge M \leq T \wedge M$ are stopping times. You can check that $X_{S\wedge M} \to X_S$ almost surely. Then by Fatou's lemma,

$$\lim_{M\to\infty} \mathbb{E}[X_{S\wedge M}] \geq \mathbb{E}[X_S].$$

So $X_S \in L^1$.

For any $A \in \mathcal{F}_X$, we want to prove $\mathbb{E}[1_A X_S] \geq \mathbb{E}[1_A X_T]$. This would immediately imply $\mathbb{E}[X_T|\mathcal{F}_S] \geq X_S$. Define the stopping time

$$S^A(\omega) = \begin{cases} S(\omega) & \text{if } \omega \in A \\ \infty & \text{if } \omega \notin A. \end{cases}$$

and similarly $T^A(\omega)$. Then

$$\mathbb{E}[X_{S^A\wedge M}] \geq \mathbb{E}[X_{T^A\wedge M}]$$

because both are finite. Now

$$\begin{aligned} &\mathbb{E}[1_{A\cap\{S\leq M\}} X_{S\wedge M}] + \mathbb{E}[1_{A\cap\{S>M\}} X_M] + \mathbb{E}[1_{A^c} X_M] = \mathbb{E}[X_{S^A\wedge M}] \\ &\quad \geq \mathbb{E}[X_{T^A\wedge M}] = \mathbb{E}[1_{A\cap\{S\leq M\}} X_{T\wedge M}] + \mathbb{E}[1_{A\cap\{S>M\}} X_M] + \mathbb{E}[1_{A^c} X_M]. \end{aligned}$$

Hence we get $\mathbb{E}[1_{A\cap\{S\leq M\}} X_S] \geq \mathbb{E}[1_{A\cap\{S\leq M\}} X_{T\wedge M}]$.

Taking the limit $M \to \infty$, we get

$$\begin{aligned} \mathbb{E}[1_{A\cap\{S<\infty\}} X_S] &= \lim_{M\to\infty} \mathbb{E}[1_{A\cap\{S\leq M\}} X_S] \geq \lim_{M\to\infty} \mathbb{E}[1_{A\cap\{S\leq M\}} X_{T\wedge M}] \\ &\geq \mathbb{E}[1_{A\cap\{S<\infty\}} X_T] \end{aligned}$$

by dominated convergence and Fatou's lemma. On the other hand, we have $\mathbb{E}[1_{A\cap\{S=\infty\}} X_S] = \mathbb{E}[1_{A\cap\{S=\infty\}} X_T]$ clearly. This implies that $\mathbb{E}[1_A X_S] \geq \mathbb{E}[1_A X_T]$. ☐

14.1 Backward martingales

In this case, we have a filtration

$$\cdots \subseteq \mathcal{F}_{-3} \subseteq \mathcal{F}_{-2} \subseteq \mathcal{F}_{-1} \subseteq \mathcal{F}_0.$$

Definition 14.2. A process (X_n) is a **backward supermartingale** if

$$\mathbb{E}[X_n|\mathcal{F}_m] \leq X_m$$

for any $m \leq n \leq 0$.

Theorem 14.3. *if* $\sup_{n<0} \mathbb{E}[|X_n|] \leq C$ *then there exists an* $X_{-\infty} \in L^1$ *such that* $X_n \to X_{-\infty}$ *almost surely and in* L^1*.*

Note that we also have L^1 convergence, which we did not have in the forward case.

Proof. The almost surely convergence is almost identical to the forward case, as a consequence of the upcrossing inequality.

To show convergence in L^1, define

$$A_n = \mathbb{E}[X_{-1} - X_{-2}|\mathcal{F}_2] + \cdots + \mathbb{E}[X_{-(n-1)} - X_{-n}|\mathcal{F}_{-(n-1)}].$$

Note that this is negative. Then $M_n = X_n + A_n$ is a martingale. Then $A_n \to A_{-\infty}$ almost surely and in L^1, which comes from the monotone convergence theorem.

Now M_n being a martingale simply means

$$M_{-n} = \mathbb{E}[M_{-1}|\mathcal{F}_{-n}].$$

This is just the setting of Levi's theorem and so immediately we have L^1 convergence of M^{-n}. □

14.2 Finite variance processes

We assume that sample paths are continuous.

Definition 14.4. A function $a : [0, T] \to \mathbb{R}$ is of **finite variance** if there exists a signed measure μ with finite $|\mu|$ such that $a(t) = \mu([0, t])$ for all $t \leq T$.

Definition 14.5. A process (A_t) is called a **finite variance process** if all sample paths have finite variance.

Definition 14.6. A sequence X_t with $X_0 = 0$ is called a **local martingale** if there exist increasing stopping times $T_n \uparrow \infty$ such that $X_{t\wedge T_n}$ is a uniform integrable martingale for all n. If $X_0 \neq 0$, then X_t is a **local martingale** if $X_t - X_0$ is a local martingale.

15 February 27, 2017

For a finite variance process $A_s(\omega)$, we can write $A_s(\omega) = \mu([0, s])$ for some finite measure μ. So we can define

$$\int H_s(\omega) dA_s(\omega) = \int H_s(\omega) \mu(ds)(\omega).$$

Here we require

$$\int |H_s(\omega)||dA_s(\omega)| < \infty.$$

This is Section 4.1 and you can read about it.

15.1 Local martingales

We are always going to assume that everything is adapted to $\mathcal{F}_t$ and sample paths are continuous.

Definition 15.1. A sequence (M_t) is a **local martingale** with $M_0 = 0$ if there exists an increasing sequence of stopping times $T_n \to \infty$ such that $M_{t \wedge T_n}$ is an uniformly integrable martingale for each n. If $M_0 \neq 0$, then M_t is called a **local martingale** if $M_t - M_0 = N_t$ is a local martingale.

Proposition 15.2. *(i) If A is a nonnegative local martingale with $M_0 \in L^1$, then it is a supermartingale.*
(ii) If there exists a $Z \in L^1$ such that $|M_t| \leq Z$ for all t, then the local martingale M_t is a uniformly integrable martingale.
(iii) If $M_0 = 0$, then $T_n = \inf\{t : M_t \geq n\}$ reduces the martingale, i.e., satisfies the condition in the definition.

Proof. (i) Let us write $M_s = M_0 + N_s$. By definition, there exists a $T_n \uparrow \infty$ such that

$$\lim_{n\to\infty} \mathbb{E}[M_0 + N_{t\wedge T_n} | \mathcal{F}_s] = \lim_{n\to\infty} M_0 + N_{s \wedge T_n} = M_s.$$

Then by Fatou's lemma, we get $\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s$.

(ii) This is clear because M_t is uniformly integrable.

(iii) This basically truncating the local martingale. □

Proposition 15.3. *A finite variance local martingale with $M_0 = 0$ is identical to 0. (Again we are assuming continuous sample paths.)*

Proof. Define a stopping time

$$\tau_n = \inf\{t : \textstyle\int_0^t |dM_s| \geq n\},$$

and let $N_t = M_{\tau_n \wedge t}$. Then we have

$$|N_t| \leq \left| \int_0^t dM_s \right| \leq n.$$

For any t and $0 = t_0 < \cdots < t_p = t$,

$$\mathbb{E}[N_t^2] = \mathbb{E}\left(\sum_i \delta N_i\right)^2 = \mathbb{E}\sum_i (\delta N_i)^2 = \mathbb{E}\left[\sup_i |\delta N_i| \sum_j |\delta N_j|\right] \leq n\mathbb{E}\left[\sup_i |\delta N_i|\right].$$

Taking the mesh to go to 0, we get that the right hand side goes to 0. □

Theorem 15.4. *Suppose that M_t is a local martingale. Then there exists a unique increasing process $\langle M, M\rangle_t$ such that $M_t^2 - \langle M, M\rangle_t$ is a local martingale. This $\langle M, M\rangle_t$ is called the **quadratic variation** of M.*

Example 15.5. If $M_t = B_t$ then $\langle M, M\rangle_t = t$. Then $B_t^2 - t$ is a martingale.

Proof. Let us first prove uniqueness. Suppose $M_t^2 - A_t$ and $M_t^2 - A_t'$ are local martingales. Then $A_t - A_t'$ is also a local martingale. By definition, A_t and A_t' are increasing and less than ∞ almost surely. So $A_t - A_t'$ is of finite variance. This shows that $A_t - A_t' = 0$ for all t.

Proving existence is rather difficult. Suppose $M_0 = 0$ and M_t is bounded, i.e., $\sup_t |M_t| < C$ almost surely. Fix and interval $[0, k]$ and subdivide it to $0 = t_0 < \cdots < t_n = K$. Define

$$X_t^n = \sum_i M_{t_{i-1}}(M_{t_i \wedge t} - M_{t_{i-1} \wedge t}).$$

You can check that this is a local martingale, and you can also check

$$M_{t_j}^2 - 2X_{t_j}^n = \sum_{i=1}^j (M_{t_i} - M_{t_{i-1}})^2.$$

Now define

$$\langle M, M\rangle_t = \lim_{n\to\infty} \sum_i (M_{t_i} - M_{t_{i-1}})^2.$$

We are going to continue next time. □

16 March 1, 2017

16.1 Quadratic variation

As always, we assume continuous sample paths. This is the theorem we are trying to prove:

Theorem 16.1. *Let M be a local martingale. Then there exists an increasing process $\langle M, M\rangle_t$ such that $M_t^2 - \langle M, M\rangle_t$ is a local martingale.*

We proved uniqueness using the fact that if there are two such A, A' then $A - A'$ is a finite variation process that is also a local martingale.

For existence, fix K and subdivide $[0, K]$ into $0 = t_0 < \cdots < t_n = K$. We define

$$X_t^n = M_0(M_{t_1\wedge t} - M_{0\wedge t}) + \cdots + M_{t_{n-1}}(M_{t_n\wedge t} - M_{t_{n-1}\wedge t}).$$

Then we immediately have

$$M_{t_j}^2 - 2X_{t_j}^n = \sum_{i=1}^{j} (M_{t_i} - M_{t_{i-1}})^2.$$

Lemma 16.2. *Assuming that $|M_t| \leq C$ (so that M is actually a martingale),*

$$\lim_{m,n\to\infty,\text{mesh}\to 0} \mathbb{E}[(X_k^n - X_k^m)^2] = 0.$$

Proof. Let us consider the simple case where $n = 5$ and $m = 3$ and the subsequence is given by t_1, t_3, t_5. Then we compute

$$\begin{aligned}
\mathbb{E}[(X_k^n - X_k^m)^2] &= \mathbb{E}\biggl[\biggl\{\sum_{j=1}^{5} M_{j-1}(M_j - m_{j-1}) - M_3(M_5 - M_3) - M_1(M_3 - M_1)\biggr\}^2\biggr] \\
&= \mathbb{E}[((M_4 - M_3)(M_5 - M_4) + (M_2 - M_1)(M_3 - M_2))^2] \\
&= \mathbb{E}[(M_4 - M_3)^2(M_5 - M_4)^2 + (M_2 - M_1)^2(M_3 - M_2)^2] \\
&\leq \mathbb{E}\Bigl[\max_j |M_j - M_{j-1}|^2 \sum_j (M_j - M_{j-1})^2\Bigr] \\
&\leq \mathbb{E}[\max(M_j - M_{j-1})^4]^{1/2} \mathbb{E}\Bigl[\Bigl(\sum_j (M_j - M_{j-1})^2\Bigr)^2\Bigr]^{1/2}.
\end{aligned}$$

Here, we see that

$$
\begin{aligned}
\mathbb{E}\bigg[\bigg(\sum_j (M_j - M_{j-1})^2\bigg)^2\bigg] &\leq \sum_j \mathbb{E}[(M_j - M_{j-1})^4] \\
&\quad + 2\mathbb{E}(M_1 - M_0)^2 \sum_{j=2}^{5} (M_j - M_{j-1})^2 + \cdots \\
&\leq 4C^2 \sum_j \mathbb{E}[(M_j - M_{j-1})^2] \\
&\quad + 2\sum_j \mathbb{E}\bigg[(M_j - M_{j-1})^2 \mathbb{E}\bigg[\sum_{k=j+1}^{n} (M_k - M_{k-1})^2 \bigg| \mathcal{F}_j\bigg]\bigg] \\
&= 4C^2 \sum_j \mathbb{E}[(M_j - M_{j-1})^2] + 2\mathbb{E}[(M_j - M_{j-1})^2 \mathbb{E}[(M_K - M_j)^2 | \mathcal{F}_j]] \\
&\leq 12C^2 \mathbb{E}\bigg[\sum_j (M_j - M_{j-1})^2\bigg] = 12C^2 \mathbb{E}[(M_K - M_0)^2] \leq 48C^4.
\end{aligned}
$$

So taking $n, m \to \infty$ we get $\mathbb{E}[\max(M_j - M_{j-1})^4] \to 0$. □

Now let us go back to the original theorem. We check that $M_t^2 - 2X_t^n$ is an increasing process. Take $n \to \infty$. By the lemma, $X_{t\wedge K}^n \to Y_{t\wedge K}$ in L^2. You can also check that $Y_{t\wedge K}$ has continuous sample paths.

Define

$$\langle M, M \rangle_t = M_{t\wedge K}^2 - 2Y_{t\wedge K}.$$

If $\{t_i\}$ is a subdivision of $[0, t]$, we have

$$\sum_{i=1}^{n} (M_{t_i^n} - M_{t_{i-1}^n})^2 \to \langle M, M \rangle_t$$

in L^2. Also, $X_{t\wedge K}$ is a local martingale.

This construction works for all fixed K. For every K, we get a A_t^K such that $M_{t\wedge K}^2 - A_{t\wedge K}^K$ is a martingale. Then by uniqueness, $A_t^{K+1} = A_t^K$ if $t \leq K$ almost surely. This means that we can just take the union of all the A_t^K.

17 March 3, 2017

Last class we almost finished proved the following theorem.

Theorem 17.1. *Let M be a local martingale. Then there exists a unique increasing process $\langle M, M\rangle_t$ such that $M_t^2 - \langle M, M\rangle_t$ is a local martingale.*

In the proof, we haven't check that the L^2 limit $X_t^n \to X_t$ has continuous sample paths. To show this, it suffices to check that

$$\mathbb{E}\Big[\sup_{s\le t}|X_s^m - X_s^n|^2\Big] \to 0$$

as $n, m \to \infty$. By Doob's inequality,

$$\mathbb{E}\Big[\sup_{s\le t}|X_s^m - X_s^n|^2\Big] \le 4\mathbb{E}[|X_t^m - X_t^n|^2].$$

Then this is clear from the L^2 convergence.

For the general case, define $T_n = \inf\{t : |M_t| \ge \}$. Then M^{T_n} is a bounded martingale and so there exists an increasing process A_t^n such that $(M^{T_n})^2 - A_t^n$ is a martingale. By uniqueness, $A_{t\wedge T_n}^{n+1} = A_t^n$. Now we can take the limit $A_t = \lim_{n\to\infty} A_t^n$. Then we have

$$(M_{T_n\wedge t})^2 - A_{T_n\wedge t} = (M^{T_n})_t^2 - A_{t\wedge T_n}^n$$

is a martingale. This means that $M_t^2 - A_t$ is a local martingale.

17.1 Consequences of the quadratic variation

Corollary 17.2. *Let M be a local martingale with $M_0 = 0$. Then $M = 0$ if and only if $\langle M, M\rangle_t = 0$.*

Proof. Suppose $\langle M, M\rangle_t = 0$. Then M^2 is a nonnegative local martingale. Then M_t^2 is a supermartingale. Since $M_0 = 0$, we immediately get $M = 0$. □

Corollary 17.3. *Assume M is a local martingale and $M_0 \in L^2$. Then $\mathbb{E}\langle M, M\rangle_\infty < \infty$ if and only if M is a true martingale and $\mathbb{E}M_t^2 < C$ for all t.*

Proof. Suppose M is a martingale with $M_0 = 0$ and $\mathbb{E}M_t^2 < C$. Then

$$\mathbb{E}\sup_{t\le T} M_t^2 \le 4\mathbb{E}M_T^2 \le 4C.$$

Let $S^n = \inf\{t : \langle M, M\rangle_t \ge n\}$. Then

$$M_{S^n\wedge t} - \langle M, M\rangle_{S^n\wedge t} \le \sup_{t\ge 0} M_t^2.$$

But $\sup_t M_t^2$ is in L^1. Then

$$\lim_{n\to\infty} \mathbb{E}[M_{S^n\wedge t}^2 - \langle M, M\rangle_{S^n\wedge t}] = \mathbb{E}[M_0^2 - \langle M, M\rangle_0] = 0.$$

By monotone convergence, we have $\mathbb{E}\langle M, M\rangle_t = \mathbb{E}M_t^2 < C$.

Conversely, suppose $\mathbb{E}\langle M, M\rangle_\infty < \infty$. Let $T_n = \inf\{t : |M_t| \geq n\}$. We have

$$M_{T_n \wedge t}^2 - \langle M, M\rangle_{T_n \wedge t} \leq n^2 \in L^1.$$

Then $M_{T_n \wedge t}^2 - \langle M, M\rangle_{T_n \wedge t}$ is a uniformly integrable martingale and so

$$\lim_{n\to\infty} \mathbb{E}[M_{T_n \wedge t}^2 - \langle M, M\rangle_{T_n \wedge t}] = 0$$

and so $\mathbb{E}[M_t^2] \leq \mathbb{E}[\langle M, M\rangle_t] < C$. This also implies that M is a true martingale.

$\square$

18 March 6, 2017

Let us try to finish this chapter.

Theorem 18.1. *If* $\sup_t \mathbb{E}[M_t^2] < \infty$*, then* $\mathbb{E}\langle M, M\rangle_\infty < \infty$ *and* $M_t^2 - \langle M, M\rangle_t$ *is a uniformly integrable martingale. If* M *is a local martingale and* $\mathbb{E}\langle M, M\rangle_\infty < \infty$ *then* $\sup_t \mathbb{E}[M_t^2] < \infty$*.*

18.1 Bracket of local martingales

Definition 18.2. If M and N are two local martingales, we define their **bracket** as

$$\langle M, N\rangle_t = \frac{1}{2}[\langle M+N, M+N\rangle - \langle M, M\rangle - \langle N, N\rangle].$$

Proposition 18.3 (Le Gall 4.15)**.** *In probability,*

$$\lim \sum_i (M_{t_i} - M_{t_{i-1}})(N_{t_i} - N_{t_{i-1}}) = \langle M, N\rangle_t.$$

Furthermore, if M *and* N *are two* L^2 *martingales, then* $\mathbb{E}\langle M, N\rangle_\infty = \mathbb{E}[M_\infty N_\infty]$*.*

Definition 18.4. We say that M is **orthogonal** to N if $\langle M, N\rangle = 0$.

Theorem 18.5 (Kunita–Watanabe)**.** *If* M *and* N *are two local martingales, then*

$$\int_0^\infty |H_s(\omega)K_s(\omega)||d\langle M, N\rangle_s| \le \left(\int_0^\infty |H_s| d\langle M, M\rangle_s\right)^{1/2} \left(\int_0^\infty |K_s|^2 d\langle N, N\rangle_s\right)^{1/2}.$$

Proof. You can just subdivide the intervals finely and use the formula for $\langle M, N\rangle_t$. □

Definition 18.6. A process X_t is called a **semimartingale** if $X_t = M_t + A_t$ where M_t is a local martingale and A_t is a finite variance process.

Note that such a decomposition is unique since a finite variance local martingale is zero.

Proposition 18.7. *For two semimartingales* $X = M + A$ *and* $Y = N + B$*, we define* $\langle X, Y\rangle = \langle M, N\rangle$*. Then*

$$\sum_i (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}) \to \langle X, Y\rangle = \langle M, N\rangle.$$

Proof. We only have to check that

$$\sum_i (N_{t_i} - N_{t_{i-1}})(A_{t_i} - A_{t_{i-1}}) \to 0$$

We use the Schwartz inequality. We have

$$\sum_i (A_{t_i} - A_{t_{i-1}})^2 \le \sup_i |A_{t_i} - A_{t_{i-1}}| \sum_i |A_{t_i} - A_{t_{i-1}}|.$$

The first term goes to zero and the second term is bounded. □

18.2 Stochastic integration

We define the space

$$\mathbb{H}^2 = \begin{Bmatrix} M \text{ a continuous martingale with} \\ M_0 = 0 \text{ and } \sup_t \mathbb{E}M_t^2 < \infty. \end{Bmatrix}$$

We then have $\langle M, M\rangle_\infty = \mathbb{E}M_\infty^2$ and $[M_\infty|\mathcal{F}_t] = M_t$. We define the inner product as

$$(M, N) = \mathbb{E}[\langle M, N\rangle_\infty] = \mathbb{E}M_\infty N_\infty.$$

Also define the space of **elementary processes** as

$$\mathscr{E} = \left\{ H_s(\omega) = \sum_i H_i(\omega) 1_{(t_i, t_{i+1}]}(s),\ H_i \in \mathcal{F}_{t_i} \right\}.$$

We can then define the integral as

$$\int H_s(\omega) dM_s = \sum_i H_i(\omega)(M_{t_i} - M_{t_{i-1}}).$$

This is just like Riemann integration.

Theorem 18.8. $\mathbb{H}^2$ *is a Hilbert space.*

Proof. Suppose there is a Cauchy sequence $\mathbb{E}[(M_\infty^n - M_\infty^m)^2] \to 0$. Then $M_\infty^n \to Z$ in L^2. Our goal is to find a M_t with continuous sample paths such that $M_t^n \to M_t$ in some sense. □

19 March 8, 2017

We define $\mathbb{H}^2$ to be the space of martingales with $\sup_t \mathbb{E}[M_t^2] < \infty$. The inner product is defined as $(M, N) = \mathbb{E}[M_\infty N_\infty] = \mathbb{E}\langle M, N\rangle_\infty$.

Proposition 19.1. *$\mathbb{H}^2$ is a Hilbert space.*

Proof. Suppose $\lim_{m,n\to\infty} \mathbb{E}[(M_\infty^m, M_\infty^n)] = 0$ so that M^n is a Cauchy sequence. Then there exists a Z such that $M_\infty^n \to Z$ in $L^2(P)$. We also have, for an fixed t, $\mathbb{E}[(M_t^m - M_t^n)^2] \to 0$ and so M_t^∞. What we now need to show is that M^∞ has continuous sample paths.

We have

$$\mathbb{E}[\sup_t (M_t^n - M_t^m)^2] \leq 4\mathbb{E}[(M_\infty^n - M_\infty^m)^2] \to 0.$$

This shows that you can find a subsequence such that

$$\mathbb{E}\left[\sum_{k=1}^\infty \sup_t |M_t^{n_{k+1}} - M_t^{n_k}|\right] \leq 4\sum_{k=1}^\infty \mathbb{E}[(M_\infty^{n_{k+1}} - M_\infty^{n_k})^2]^{1/2} < \infty.$$

In particular,

$$Y_t = \sum_k (M_t^{n_{k+1}} - M_t^{n_k})$$

is a martingale with continuous sample paths. Also $Y_\infty = Z$ and so $M^n \to Y$ in $\mathbb{H}^2$. $\square$

19.1 Elementary process

Definition 19.2. We define the space $\mathscr{E}$ of **elementary processes** as

$$\mathscr{E} = \left\{ \sum_{i=0}^{p-1} H_i(\omega) 1_{(t_i, t_{i+1}]} : H_i \in \mathscr{F}_{t_i}, \text{ with} \right.$$
$$\left. \sum_i H_i^2(\langle M, M\rangle_{t_{i+1}} - \langle M, M\rangle_{t_i}) < \infty \right\}.$$

We also define

$$L^2(M) = \left\{ H_s \text{ such that } \int H_s^2 d\langle M, M\rangle_s < \infty \right\}.$$

Proposition 19.3. *If $M \in \mathbb{H}^2$, then the space $\mathscr{E}$ is dense in $L^2(M)$.*

Theorem 19.4. *For any $M \in \mathbb{H}^2$, we can define*

$$(H \cdot M)_t = \sum_{i=0}^{p-1} H_i(\omega)(M_{t_{i+1}\wedge t} - M_{t_i \wedge t}) = \int_0^t H dM \in \mathbb{H}^2.$$

This map $\mathscr{E} \to \mathbb{H}^2$ given by $H \mapsto H \cdot M$ is an isometry from $\mathscr{E}$ to $\mathbb{H}$:

$$\|H \cdot M\|_{\mathbb{H}^2} = \int H_s^2 d\langle M, M\rangle_s.$$

So there exists a unique extension from $L^2(M) \to \mathbb{H}$, denoted by $\int HdM$. Furthermore,

(1) $\langle H \cdot M, N\rangle = H\langle M, N\rangle$,
(2) $(1_{[0,T]}H) \cdot M = (H \cdot M)^T = H \cdot M^T$.

Roughly speaking, stochastic integration is just Riemann integration on elementary functions, which has a unique extension.

You can check that

$$\langle H \cdot M, H \cdot M\rangle_t = \sum_i H_i^2(\langle M, M\rangle_{t_{i+1}\wedge t} - \langle M, M\rangle_{t_i \wedge t}).$$

Proof. Since $(H \cdot M)^2 - H^2\langle M, M\rangle$ is a martingale, we have

$$\mathbb{E}[(H \cdot M)_\infty^2] = \mathbb{E}\int_0^\infty H_s^2 d\langle M, M\rangle_s.$$

This shows that stochastic integration is and L^2 isometry. Then there exists a unique extension.

Now we want to check that $\langle H \cdot M, N\rangle = H\langle M, N\rangle$. Note that $\langle M, N\rangle$ is a finite variance process $\langle M, M\rangle$ and $\langle N, N\rangle$ are increasing and $\langle M + N, M + N\rangle$ is decreasing, and also $\mathbb{E}|\langle M, N\rangle| < \infty$. We want to check that

$$\left\langle \int_0^t Hd_M, N_t \right\rangle = \int_0^t Hd\langle M, N\rangle.$$

For elementary processes, we have

$$\left\langle \int_0^t H_s dM_s, \int_0^t dN_s \right\rangle = \int_0^t H_s d\langle M, N\rangle_s.$$

Then it extends to all processes. $\square$

20 March 20, 2017

20.1 Review

Today we will review, because all the properties are hard to remember. All processes have continuous sample paths and are adapted.

1. Finite variation processes: A_t is a finite variation process if its sample paths have finite variation almost surely. Classical theory of integration applies here. Then

$$\int_0^t H_s(\omega) dA_s(\omega)$$

   is defined by Lebesgue integration. (Note that dA_s is not positive.)

2. Continuous local martingales: Usually we assume $M_0 = 0$. M_t is a local martingale if there exists an increasing sequence $T_n \uparrow \infty$ almost surely such that $(M^{T_n})_t = M_{T_n \wedge t}$ is an uniformly integrable martingale. In this case, we say that T_n reduces the local martingale.

   Why do we look at these objects ? If X_n is a sub/supermartingale, and $\sup_n \mathbb{E}|X_n|^p < \infty$ for some $p \geq 1$, then $X_n \to X_\infty$ almost surely. Moreover if $p > 1$ then $X_n \to X_\infty$ in L^p (by Doob's inequality). Furthermore, if X_n is uniformly integrable then $X_n \to X_\infty$ also in L^1. We really need uniform integrability to run the theory, and so we have to stick in a stopping time to make things uniformly integrability.

3. Quadratic variation: If M is a local martingale, then there exists a unique increasing process $\langle M, M\rangle_t$ such that $M_t^2 - \langle M, M\rangle_t$ is a local martingale. This is constructed as

$$\langle M, M\rangle_t = \lim \sum_i (M_{t_{i+1}} - M_{t_i})^2.$$

   For $M = B$ a Brownian motion, we have $\langle B, B\rangle_t = t$.

4. Some properties of local martingales:

   - A nonnegative local martingale M with $M_0 \in L_1$ is a submartingale. (This follows from Fatou's lemma.)
   - If there exists a $Z \in L^1$ such that $|M_t| < Z$ for all t, then M is a uniformly integrable martingale.
   - If M with $M_0 = 0$ is both a local martingale and a finite variation process, then $M = 0$.
   - For M a local martingale with $M_0 = 0$, $\langle M, M\rangle = 0$ if and only if $M = 0$.

5. Stochastic integration: We first define the class $\mathbb{H}^2$ of square-integrable martingales with the inner product

$$(M, N)_{\mathbb{H}^2} = \mathbb{E}[M_\infty N_\infty] = \mathbb{E}[\langle M, N\rangle_\infty].$$

Then we defined, for any $M \in \mathbb{H}^2$, the space $L^2(M)$ of processes H such that

$$\mathbb{E}\int_0^\infty H^2 d\langle M, M\rangle < \infty.$$

The space $\mathscr{E}$ is the space of elementary processes, processes of the form $\sum_{i=0}^{p-1} H_i 1_{(t_i, t_{i+1}]}$ for H_i measurable with respect to $\mathcal{F}_{t_i}$ and bounded almost surely. You can check that for every $M \in \mathbb{H}^2$, the space $\mathscr{E}$ is dense in $L^2(M)$. (Note that the inner product on $\mathscr{E}$ depends on M.)

We can now define stochastic integration $\int H dM$ on the space $\mathscr{E}$ by

$$(H \cdot M)_t = \sum_i H_i (M_{t_{i+1}\wedge t} - M_{t_i \wedge t}).$$

This gives an isometry $\mathscr{E} \to \mathbb{H}^2$. So you can extend it to an isometry $L^2(M) \to \mathbb{H}^2$.

$$\begin{tikzcd} L^2(M) \arrow[r, dashed] & \mathbb{H}^2 \\ \mathscr{E} \arrow[u, hook] \arrow[ur] & \end{tikzcd}$$

So far we have defined stochastic integration for $M \in \mathbb{H}^2$. Now we want to define stochastic integration for M a local martingale. Then we would like to define

$$\Big(\int H dM\Big)^{T_n} = \int H dM^{T_n}.$$

We are also going to define, for a semi-martingale $X = M + V$,

$$\int H dX = \int H dM + \int H dV.$$

Next time we are going to quickly go over these definitions and start actually doing some stuff.

21 March 22, 2017

For any fixed $M \in \mathbb{H}^2$, we consider the space

$$L^2(M) = \Big\{\text{adapted processes } (H_t) \text{ with } \mathbb{E}\int_0^\infty H_t^2 d\langle M, M\rangle < \infty\Big\}.$$

There is an isometry

$$(-)\cdot M : \mathscr{E} \to \mathbb{H}^2; \quad \sum_i H_i 1_{(t_i,t_{i+1}]} \mapsto \sum_i H_i(M_{t_{i+1}\wedge t} - M_{t_i\wedge t}).$$

This can be extend to an isometry $L^2(M) \to \mathbb{H}^2$.

21.1 Stochastic integration with respect to a Brownian motion

We now want to look at stochastic integration with respect to Brownian motions. But (B_t) is not in $\mathbb{H}^2$. More generally, we are going to stochastic integration this to local martingales. We are going to do in the setting

(1) M is a local martingale,

(2) $H \in L^2_{\text{loc}}(M)$, i.e., $\int_0^t H_s^2 d\langle M, M\rangle_s < \infty$ almost surely for any fixed t.

Recall that for the Brownian motion, $\langle B, B\rangle_t = t$. This is easy to prove because

$$\mathbb{E}\Big[\sum_i ((B_{t^n_{i+1}} - B_{t^n_i})^2 - (t^n_{i+1} - t^n_i))\Big]^2 = C\sum_{i=0}^{p-1}(t^n_{i+1} - t^n_i)^2 \to 0.$$

Then we can define, for $t \leq T$,

$$\int_0^t H dB = \sum_i H_i(B_{t_{i+1}\wedge t} - B_{t_i\wedge t}).$$

We check that

$$\mathbb{E}\Big[\int_0^T H dB\Big]^2 = \mathbb{E}\Big(\sum_i H_i(B_{t_{i+1}} - B_{t_i})\Big)^2 = \sum_i H_i^2(t_{i+1} - t_i) = \int_0^T H_s ds.$$

That is, $\int_0^t H dB$ is an isometry in L^2. Then you can again extend it to all adapted processes H_s such that $\int_0^T H_s^2 ds < \infty$ almost surely

21.2 Stochastic integration with respect to local martingales

Let M be a local martingale, and we define

$$L^2_{\text{loc}}(M) = \Big\{(H_s) : \int_0^t H_s^2 d\langle M, M\rangle_s < \infty \text{ almost surely for all } t\Big\}.$$

Theorem 21.1. *There exists a unique continuous local martingale $H \cdot M = \int H_s dM_s$ such that*

- $\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle$,
- $H \cdot (K \cdot M) = (HK) \cdot M$,
- $\langle \int H dM, \int K dN \rangle_t = \int_0^t H_s K_s d\langle M, N \rangle_s$,
- $\int H d(\int K dM) = \int (HK) dM$.

Proof. We look at the stopping time

$$T_n = \inf\{t : \int_0^t (1 + H_s^2) d\langle M, M \rangle_s \geq n\}.$$

You can check that they actually are stopping times and $T_n \uparrow \infty$. Consider $(M^{T_n})_t$. You can also check $\langle M^{T_n}, M^{T_n} \rangle = \langle M, M \rangle^{T_n}$. Then

$$\mathbb{E}(M^{T_n})_t^2 \leq \mathbb{E}(\langle M, M \rangle_t^{T_n}) \leq n.$$

That is, $M^{T_n} \in \mathbb{H}^2$, and so $H \cdot M^{T_n}$ is well-defined.

We now want a process $H \cdot M$ such that $H \cdot M^{T_n} = (H \cdot M)^{T_n}$. We will continue next time. □

If X is a semi-martingale, with $X = M + V$ where V is a finite variation process, then we can further define

$$\int H_s dX = \int H_s dM_s + \int H_s dV_s.$$

Theorem 21.2 (Ito's formula)**.** *Let $F \in C^2$ be a function and $X^1, \ldots, X^p$ be semi-martingales. Then*

$$F(X_t^1, \ldots, X_t^p) = F(X_0^1, \ldots, X_0^p) + \sum_{i=1}^p \int_0^t \frac{\partial F}{\partial X^i}(X_s^1, \ldots, X_s^p) dX_s^i + \sum_{i,j=1}^p \int_0^t \frac{\partial^2 F}{\partial X^i \partial X^j}(X_s^1, \ldots, X_s^p) d\langle X^i, X^j \rangle_s.$$

22 March 24, 2017

Theorem 22.1. *Let M be a local martingale, and consider the space*

$$L^2_{\mathrm{loc}}(M) = \Big\{(H_s) : \int_0^t H_s^2(\omega) d\langle M, M\rangle_s < \infty \text{ almost surely for any fixed } t\Big\}.$$

Then there exists a unique integration $\int H_s dM_s$ giving a continuous local martingale such that

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle.$$

Proof. Assume $M_0 = 0$ as always, and consider the stopping time

$$T_n = \inf\Big\{t : \int_0^t (1 + H_s^2) d\langle M, M\rangle_s \geq n\Big\}.$$

These are actually stopping times and $T_n \uparrow \infty$. We see that M^{T_n} is a L^2 martingale because $\int_0^t d\langle M, M\rangle_s \leq n$. Then

$$\mathbb{E}[(M^{T_n})_t^2] = \mathbb{E}[\langle M^{T_n}, M^{T_n}\rangle_t] = \mathbb{E}[\langle M, M\rangle_{t\wedge T_n}] \leq n.$$

So the integration $H \cdot M^{T_n} = \int H dM^{T_n}$ is well-defined.

For $m \geq n$, you can check that $(H \cdot M^{T_m})^{T_n} = H \cdot M^{T_n}$. That is, this is a consistent family and so there exists a local martingale $H \cdot M$ such that $(H \cdot M)^{T_n} = H \cdot M^{T_n}$. □

22.1 Dominated convergence for semi-martingales

Proposition 22.2. *Let X be a continuous semi-martingale, and H be an adapted process with continuous sample paths. For every $t > 0$ fixed, and $0 = t_0^n < \cdots < t_p^n = t$,*

$$\sum_{i=0}^{p-1} H_{t_i^n}(X_{t_{i+1}^n} - X_{t_i^n}) \to \int_0^t H_s dX_s$$

in probability.

Lemma 22.3 (Dominated convergence in stochastic integration)**.** *Let $X = M + V$ be a semi-martingale. Assume that H satisfies $H_s^n \to H_s$ almost surely for $s \in [0, t]$, and there exists a $K_s \geq 0$ such that $|H_s^n| \leq K_s$ for all n and $s \in [0, t]$, with the conditions*

$$\int_0^t K_s^2 d\langle M, M\rangle_s < \infty, \quad \int_0^t K_s |dV_s| < \infty$$

almost surely. Then

$$\int H_s^n dX_s \to \int H_s dX_s$$

in probability.

Proof. The case $M = 0$ is exactly the classical dominated convergence. Let us now assume $V = 0$. Define the stopping time

$$T_p = \Big\{\gamma \le t : \int_0^\gamma K_s^2 d\langle M, M\rangle_s \ge p\Big\} \wedge t.$$

We have

$$\begin{aligned}\mathbb{E}\Big[\Big(\int_0^{T_p} H_s^n dM_s - \int_0^{T_p} H_s dM_s\Big)^2\Big] &\le \mathbb{E}\int_0^{T_p} (H_s^n - H_s)^2 d\langle M, M\rangle_s \\ &\le \mathbb{E}\int_0^{T_p} 2K_s^2 d\langle M, M\rangle_s < C(p).\end{aligned}$$

By dominated convergence, the left hand side goes to 0 as $n \to \infty$. Since $P(T_p = t) \to 1$ as $p \to \infty$, we get the desired result. □

Proof of Proposition 22.2. Define the processes

$$H_s^n = \begin{cases} H_{t_i}^n & t_i < s \le t_{i+1}^n \\ 0 & s > t. \end{cases}$$

Let $K_s = \max_{0 \le r \le s} |H_r|$. Then H has continuous sample paths and so $K_s < \infty$ almost surely. We trivially check that $|H_s^n| \le K_s$ for all n and s. Also

$$\int K_s^2 d\langle M, M\rangle_s < \infty$$

because it is locally bounded. $H_s^n \to H_s$ almost surely as $n \to \infty$, because sample paths are continuous. So we can apply dominated convergence. □

22.2 Itô's formula

Theorem 22.4 (Itô's formula)**.** *Let $F \in C^2$ be a function and $X^1, \dots, X^p$ be semi-martingales. Then*

$$\begin{aligned}F(X^1, \dots, X^p) = F(X_0^1, \dots, X_0^p) &+ \sum_{i=1}^p \int_0^t \frac{\partial F}{\partial X^i}(X_s^1, \dots, X_s^p) dX_s^i \\ &+ \frac{1}{2}\sum_{i,j=1}^p \int_0^t \frac{\partial^2 F}{\partial X^i \partial X^j}(X_s^1, \dots, X_s^p) d\langle X^i, X^j\rangle_s.\end{aligned}$$

Proof. Let us consider only the case $p = 1$. We have

$$\begin{aligned}F(X_t) &= F(X_0) + \sum_{i=0}^{p-1}(F(X_{t_{i+1}^n}^n) - F(X_{t_i}^n)) \\ &= \sum_i F'(X_{t_i})(X_{t_{i+1}^n} - X_{t_i^n}) + \frac{1}{2}\sum_i F''(X_{t_i}^n + \theta(X_{t_{i+1}^n} - X_{t_i^n}))(X_{t_{i+1}^n} - X_{t_i^n})^2\end{aligned}$$

by simple Taylor expansion, with $0 \leq \theta \leq 1$. Now the first term converges to $\int F(X_s)dX_s$ by the Proposition 22.2. To show that the second term goes to $\frac{1}{2}\int F''(X_s)d\langle X,X\rangle_s$, we note that

$$\mathbb{E}\Big[\sum_i F''(B^n_{t_i})[(X_{t^n_{i+1}} - X_{t^n_i})^2 - (\langle X,X\rangle_{t^n_{i+1}} - \langle X,X\rangle_{t^n_i})]\Big]^2$$
$$\leq C\sum(\langle X,X\rangle_{t^n_{i+1}} - \langle X,X\rangle_{t^n_i})^2 \to 0$$

since all the cross terms vanish. So we get the formula. □

23 March 27, 2017

Last time we prove Itô's formula. For $F \in C^2$ and semi-martingales $X^1, \ldots, X^d$, we have

$$F(t, X_t^1, \ldots, X_t^d) = F(0, X_0^1, \ldots, X_0^d) + \int_0^t \sum_{i=1}^d \frac{\partial F}{\partial X^i}(t, X_s^1, \ldots, X_s^d) dX_s^i$$
$$+ \int_0^t \Big(\frac{\partial F}{\partial t} dt + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 F}{\partial X^i \partial X^j}(s, X_s^1, \ldots, X_s^d) d\langle X_i, X_j \rangle_s \Big).$$

In the special case of $d = 1$ and $X = B$, we get

$$F(t, B_0) = F(0, B_0) + \int_0^t \frac{\partial F}{\partial X}(s, B_s) dB_s + \int_0^t \Big(\frac{\partial F}{\partial s} + \frac{\Delta}{2} F \Big)(s, B_s) ds.$$

Example 23.1. Let

$$\mathscr{E}(\lambda M)_t = e^{\lambda M_t - \frac{\lambda^2}{2}\langle M, M \rangle_t},$$

where M is a local martingale. For any $\lambda \in \mathbb{C}$, $\mathscr{E}(\lambda M)_t$ is a local martingale and

$$\mathscr{E}(\lambda M)_t = e^{\lambda M_0} + \lambda \int_0^t \epsilon(\lambda M)_s dM_s.$$

Proof. Let $F(x, r) = e^{\lambda M_t - \frac{\lambda^2}{2} r}$. If we set $G(x, t) = F(x, \langle M, M \rangle_t)$, then Itô's formula tells us that

$$G(M_t, t) = G(M_0, 0) + \int_0^t \frac{\partial G}{\partial M} dM_s + \int_0^t \Big(\frac{\partial G}{\partial s} + \frac{1}{2} \frac{\partial^2 G}{\partial M^2} \Big)(s, x) d\langle M, M \rangle_s.$$

Here, we can treat $\langle M, M \rangle_t$ as an external time dependence, because it is a finite variation process. Computing the second derivative gives

$$\frac{\partial G}{\partial s} ds + \frac{G''}{2} d\langle M, M \rangle_s = \Big[-\frac{\lambda^2}{2} \frac{dr}{ds} ds + \frac{\lambda^2}{2} d\langle M, M \rangle_s \Big] G(s, M_s) = 0.$$

Then we are left with the first derivative. ☐

23.1 Application of Itô calculus

Theorem 23.2 (Levi). *Let $X^1, \ldots, X^d$ be adapted local martingales with continuous ample paths. If $\langle X^i, X^j \rangle_t = \delta_{ij} t$, then $X^1, \ldots, X^d$ are independent Brownian motions.*

Proof. Let us only look at the case $d = 1$. By the previous theorem,

$$\mathscr{E}(t) = e^{i\xi X_t + \frac{\xi^2}{2} t}$$

is a local martingale and further a martingale because it is bounded. This shows that

$$\mathbb{E}[e^{i\xi(X_t-X_s)+\frac{\xi^2}{2}(t-s)}|\mathcal{F}_s] = 1$$

for all ξ. Then $X_t - X_s$ is Gaussian with variance $(t-s)$ when conditioned on $\mathcal{F}_s$. □

Theorem 23.3 (Burkholder–Davis–Gundy)**.** *Let M_t be a local martingale and set $M_t^* = \sup_{0\le s\le t}|M_s|$. For every $p > 0$ there exist constants c_p, C_p depending only on p such that for any time T,*

$$c_p\mathbb{E}[\langle M, M\rangle_T^{p/2}] \le \mathbb{E}[(M_T^*)^p] \le C_p\mathbb{E}[\langle M, M\rangle_T^{p/2}].$$

Proof. Let us first consider the case $p \ge 2$. By Itô's formula,

$$|M_t|^p = \int_0^t p|M_s|^{p-1}\operatorname{sgn} M_s dM_s + \frac{1}{2}\int_0^t p(p-1)|M_s|^{p-2}d\langle M, M\rangle_s.$$

We may assume that M is bounded because we can apply to M^{S_n} with $S_n = \inf\{t : |M_t| \ge n\}$ and take the limit. Then $\int_0^t p|M_s|^{p-1}\operatorname{sgn} M_s dM_s$ is an L^2 martingale.

Hence

$$\begin{aligned}\mathbb{E}[|M_t|^p] &\le p(p-1)\mathbb{E}\Big[\int_0^t |M_s|^{p-2}d\langle M, M\rangle_s\Big] \le p(p-1)\mathbb{E}[|M_t^*|^{p-2}\langle M, M\rangle_t] \\ &\le p(p-1)\mathbb{E}[|M_t^*|^p]^{(p-2)/p}\mathbb{E}[\langle M, M\rangle_t^{p/2}]^{2/p}.\end{aligned}$$

So by Doob's inequality, we get an upper bound on $\mathbb{E}[|M_t^*|^p]$ by $\mathbb{E}[\langle M, M\rangle_t^{p/2}]$. □

24 March 29, 2017

Last time we were proving the Burkholder–Davis–Gundy inequality.

24.1 Burkholder–Davis–Gundy inequality

Theorem 24.1 (Burkholder–Davis–Gundy inequality)**.** *For $0 < p < \infty$, there exists a constant C_p such that*

$$\mathbb{E}[(M_t^*)^p] \leq C_p \mathbb{E}[\langle M, M\rangle_t^{p/2}].$$

We first assumed that M is bounded. This is because we can prove the inequality for M^{T_n} where $T_n = \inf\{t : |M_t| \geq n\}$ and then take the limit $n \to \infty$. Next, we assumed that $p \geq 2$ and applied Itô's formula.

Proof (cont.) Now let us consider the case $0 < p < 2$. Define $T_x = \inf\{t : M_t^2 > x\}$. Then

$$\begin{aligned} P((M_t^*)^2 \geq x) = P(T_x \leq t) = P((M_{T_x \wedge t})^2 \geq x) &\leq \frac{1}{x}\mathbb{E}[(M_{T_x\wedge t})^2] \\ = \frac{1}{x}\mathbb{E}[\langle M, M\rangle_{T_x \wedge t}] &\leq \frac{1}{x}\mathbb{E}[\langle M, M\rangle_t] \end{aligned}$$

because $M^2 - \langle M, M\rangle$ is a martingale.

Let us define $S_x = \{\inf t : \langle M, M\rangle_t \geq x\}$. Our goal is to prove

$$P((M_t^*)^2 \geq x) \leq \frac{1}{x}\mathbb{E}[\langle M, M\rangle_t 1(\langle M, M\rangle_t \leq x)] + 2P(\langle M, M\rangle_t \geq x).$$

Once we prove this, we get for $q = p/2$,

$$\begin{aligned} \mathbb{E}[(M_t^*)^{2q}] &= \int_0^\infty P((M_t^*)^2 \geq x) q x^{q-1} dx \\ &\leq c\int_0^\infty \frac{1}{x}\mathbb{E}[\langle M, M\rangle_t 1_{(\langle M, M\rangle_t \leq x)}] x^{q-1} dx \\ &\quad + c\int_0^\infty P(\langle M, M\rangle_t \geq x) x^{q-1} dx \\ &= c_q \mathbb{E}[\langle M, M\rangle_t^q] + c_q \mathbb{E}[\langle M, M\rangle_t^q]. \end{aligned}$$

Now set us prove the inequality. We have

$$\{(M_t^*)^2 \geq x\} \subseteq \{(M_{S_x\wedge t}^*)^2 \geq x\} \cup \{S_x \leq t\}.$$

Then

$$\begin{aligned} P((M_t^*)^2 \geq x) &\leq P((M_{S_x \wedge t}^*)^2 \geq x) + P(S_x \leq t) \\ &\leq \frac{1}{x}\mathbb{E}[\langle M, M\rangle_{S_x\wedge t}] + P(S_x \leq t) \\ &\leq \frac{1}{x}\mathbb{E}[\langle M, M\rangle_t \wedge x] + P(\langle M, M\rangle_t \geq x) \\ &\leq \frac{1}{x}\mathbb{E}[\langle M, M\rangle_t 1_{(\langle M, M\rangle_t \leq x)}] + 2P(\langle M, M\rangle_t \geq x). \end{aligned}$$

□

Corollary 24.2. *Let M be a local martingale with $M_0 = 0$. If $\mathbb{E}[\langle M, M\rangle_\infty^{1/2}] < \infty$ then M is uniformly integrable.*

Proof. By the theorem, $\mathbb{E}[(M_\infty)^*] < \infty$. Then M_t is bounded by a fixed L^1 variable. $\square$

Previously we have shown that $\mathbb{E}[\langle M, M\rangle_\infty^2] < \infty$ implies that M is L^2 bounded. This is a significantly better criterion for checking that M is uniformly integrable. As a consequence, if $\mathbb{E}[(\int_0^t H_s(\omega)^2 d\langle M, M\rangle_s^{1/2})] < \infty$ then

$$\int_0^s H_\sigma(\omega) dM_\sigma$$

is uniformly integrable for $s \leq t$.

There are a few more things I want to go over. The first one is representation of martingales as stochastic integrals. The second is Girsanov's theorem.

Theorem 24.3. *Let B be a Brownian motion. If $Z \in L^2(\Omega, \mathcal{F}_\infty, P)$, then there exists a unique adapted (progressive) process $h \in L^2(B)$ such that*

$$Z = \mathbb{E}[Z] + \int_0^\infty h_s dB_s.$$

More generally,

Theorem 24.4. *If M is a local martingale with filtration $\mathcal{F}_t$, then there exists a $h_s \in L^2_{\text{loc}}(B)$ such that*

$$M_t = C + \int_0^t h_s dB_s.$$

25 March 31, 2017

25.1 Representation of martingales

Theorem 25.1.

(1) If $Z \in L^2(\Omega, \mathcal{F}_\infty, P)$ where $\mathcal{F}_t$ is a σ-field of the Brownian motion, then there exists a process h with $\mathbb{E}\int_0^t h_s^2 ds < \infty$ such that

$$Z = \mathbb{E}[Z] + \int_0^\infty h_s dB_s.$$

(2) If M_t is a local martingale with respect to $\mathcal{F}_t$, then there exists a process h with $\int_0^t h_s^2 ds < \infty$ such that

$$M_t = C + \int_0^t h_s dB_s.$$

Lemma 25.2. *The vector space generated by $\{\exp(i\sum_{j=1}^p \lambda_j(B_{t_j} - B_{t_{j-1}})\}$ is dense in $L^2_{\mathbb{C}}(\Omega, \mathcal{F}_\infty, P)$.*

Proof. By Fourier transformation, it suffices to show that the set

$$\{F(B_{t_1}, \dots, B_{t_p}) \text{ for } F \text{ smooth with compact support}\}$$

is dense in $L^2(\Omega, \mathcal{F}_\infty, P)$. □

Proof of Theorem 25.1. (1) Define $f(s) = \sum l_j 1_{(t_j, t_{j+1}]}(s)$ so that $\int_0^t f(s)dB_s = \sum_j \lambda_j(B_{t_{j+1}\wedge t} - B_{t_j \wedge t})$. Define

$$\mathcal{E}(t) = \exp\biggl(i\int_0^t f(s)dB_s + \frac{1}{2}\int_0^t f^2(s)ds\biggr).$$

Itô's formula implies

$$d\mathcal{E}(t) = \mathcal{E}(t)\Bigl[if(t)dB_t + \frac{1}{2}\langle if(t)dB_t, if(t)dB_t\rangle + \frac{1}{2}f^2(t)dt\Bigr] = \mathcal{E}(t)if(t)dB_t.$$

This shows that

$$\mathcal{E}(t) = 1 + i\int_0^t \mathcal{E}(s)f(s)dB_s.$$

Now let H be the closure of $\{\int_0^t h_s dB_s\}$. We have shown that $\{\exp(i\sum_j \lambda_j(B_{t_{j+1}} - B_{t_j}))\} \subseteq H$ and so by the lemma, $H = L^2(P)$.

(2) We know that if M_t is an L^2-martingale, then $M_\infty = \int_0^\infty h_s dB_s$. We further hand $\mathbb{E}[M_\infty|\mathcal{F}_t] = \int_0^t h_s dB_s$.

If M_t is a local martingale, then consider M^{T_n} with $T_n = \inf_t\{t : |M_t| \geq n\}$. We then have a h_s^n such that

$$M_t^{T_n} = \int_0^t h_s^n dB_s.$$

Check the consistence of h_s^n and h_s^m and put the together to form h_s. Then $M_t = \int_0^t h_s dB_s$. □

25.2 Girsanov's theorem

Theorem 25.3 (Girsanov)**.** *Let P_t and Q_t be two probability measures with the same $\{\mathcal{F}_t\}$. Assume that $P \ll Q$ (P is absolutely continuous with respect to Q) and $Q \ll P$. Define*

$$D_t = \frac{dQ_t}{dP_t}, \quad D_\infty = \frac{dQ_\infty}{dP_\infty}.$$

Then D_t is a uniformly integrable martingale with $\mathbb{E}[D_\infty|\mathcal{F}_t] = D_t$. Further $D_\infty > 0$ almost surely and there exists a local martingale L such that

$$D_t = e^{L_t - \frac{1}{2}\langle L,L\rangle_t}.$$

If M is a continuous local martingale with respect to P, then

$$\tilde{M} = M - \langle M, L\rangle$$

is a continuous local martingale with respect to Q.

Proof. From Itô's formula,

$$d\log D_t = \frac{1}{D_t} dD_t - \frac{1}{2}\frac{1}{D_t^2} d\langle D, D\rangle_t.$$

So we have

$$\log D_t = \int_0^t \frac{dD_s}{D_s} - \frac{1}{2}\frac{d\langle D, D\rangle_s}{D_s^2}.$$

If we define $L_t = \int_0^t dD_s/D_s$ then

$$\langle L, L\rangle_t = \left\langle \int_0^t \frac{dD_s}{D_s}, \int_0^t \frac{dD_s}{D_s} \right\rangle = \int_0^t \frac{d\langle D, D\rangle_s}{D_s^2}.$$

So $\log D_t = L_t - \frac{1}{2}\langle L, L\rangle_t$.

Now let us prove the main part. If MD is a local martingale with respect to P then M is a local martingale with respect to Q. This is almost a tautology, because for G a measurable function with respect to $\mathcal{F}_s$,

$$\mathbb{E}^P[(MD)_t G] = \mathbb{E}^P[(MD)_s G]$$

and this translates to $\mathbb{E}^Q[M_t G] = \mathbb{E}^Q[M_s G]$.

So it suffices to check that $\tilde{M}D$ is a martingale with respect to P. Then it suffices to check that $d(\tilde{M}D)$ is a stochastic integral of something. This is going to be a half-page computation. □

26 April 3, 2017

Theorem 26.1 (Girsanov's formula)**.** *Let $P \ll Q$ and $Q \ll P$, where P and Q are probability measures with the same $\mathcal{F}_t$. If $D_t = \frac{dQ}{dP}|_{\mathcal{F}_t}$ then D_t is a uniformly integrable martingale. There exists a local martingale L_t such that $D_t = \exp(L_t - \frac{1}{2}\langle L, L\rangle_t)$. Furthermore, if M_t is a local martingale with respect to P then*

$$\tilde{M} = M - \langle M, L\rangle$$

is a local martingale with respect to Q.

Proof. We have checked that if $\tilde{M}D_t$ is a (local) martingale with respect to P, then $\tilde{M}$ is a (local) martingale with respect to Q.

Now by Itô's formula,

$$\begin{aligned} d(\tilde{M}_t D_t) &= d\tilde{M}_t D_t + \tilde{M}_t dD_t + \langle d\tilde{M}, dD\rangle_t \\ &= D_t dM_t - D_t d\langle M, L\rangle_t + \tilde{M}_t dD_t + \langle dM, dD\rangle_t, \end{aligned}$$

because D_t is a martingale and so its variation with respect to a finite variation process is $\langle\langle M, L\rangle, D\rangle = 0$.

We can also compute

$$dD_t = d(e^{L_t - \frac{1}{2}\langle L, L\rangle_t}) = D\Big(dL_t - \frac{1}{2}d\langle L, L\rangle_t + \frac{1}{2}d\langle L, L\rangle_t\Big) = D(dL_t).$$

Therefore

$$d(\tilde{M}_t D_t) = D_t dM_t - Dd\langle M, L\rangle_t + \tilde{M}_t D_t + D_t\langle dM, dL\rangle_t = D_t dM_t + \tilde{M}_t dD_t.$$

Because M_t and D_t are martingales, $\tilde{M}_t D_t$ is a martingale with respect to P. $\square$

26.1 Cameron–Martin formula

For example, take $L_t \int_0^t b(B_s, s)dB_s$ by setting

$$Q_t = P_t e^{\int_0^t b(B_s,s)dB_s - \frac{1}{2}\int_0^t b^2(B_s,s)ds}.$$

If $M_t = B_t$, then

$$\tilde{M} = B_t - \Big\langle B_t, \int_0^t b(B_s, s)dB_s\Big\rangle = B_t - \int_0^t b(B_s, s)ds$$

is a Brownian motion with respect to Q.

If $b(B_s, s) = g(s)$, then

$$D_\infty = e^{\int_0^\infty g(s)dB_s - \frac{1}{2}\int_0^\infty g^2(s)ds}.$$

If we write $h(t) = \int_0^t g(s)ds$, then for any function Φ on $C(\mathbb{R}_+, \mathbb{R})$,

$$\mathbb{E}^W[\Phi(B + h)] = \mathbb{E}^Q[\Phi((B - h) + h)] = \mathbb{E}^Q[\Phi(B)] = \mathbb{E}^W[D_\infty \Phi(B)]$$

where W is the Wiener measure with respect to P. This is the **Cameron–Martin formula**.

Intuitively, this is true because, in the case when Φ is given by the simple discrete form $F(B(t_1))G(B(t_2))$,

$$\begin{aligned}
&\mathbb{E}^W[F(B(t_1)+h(t_1))G(B(t_2)+h(t_2))] \\
&\quad = \int e^{-\frac{x^2}{2t_1}} e^{-\frac{(x-y)^2}{2(t_2-t_1)}} F(x+h(t_1))G(y+h(t_2)) \\
&\quad = \int F(x)G(y)e^{-\frac{x^2}{2t_1}-\frac{(x-y)^2}{2(t_2-t_1)}+\frac{h(t_1)x}{t_1}+\frac{(h(t_1)-h(t_2))(x-y)}{(t_2-t_1)}-\frac{1}{2}(\frac{h(t_1)^2}{t_1}+\frac{(h(t_2)-h(t_1))^2}{t_2-t_1})^2}.
\end{aligned}$$

Here the last three terms correspond to D_∞.

There are some applications of Itô calculus in solving differential equations involving processes.

1. Suppose $\partial_t u = \frac{\Delta}{2}u$ with the initial condition $u_0(x) = \phi(x)$. Then
$$u(x,t) = \mathbb{E}_x[\phi(x_t)]$$
for some x_t.
2. Let $\Delta u(x) = 0$ for $x \in D$ and $u(x) = g(x)$ on $x \in \partial D$. Then
$$u(x) = \mathbb{E}_x[\phi(x_T)], \quad T = \inf\{t : x_t \notin D\}.$$
3. Feynman–Kac formula.

Let us prove the first statement. Let $v(s,x) = u(t-s,x)$. Then

$$\partial_s v + \frac{\Delta}{2}v = 0.$$

Now Itô's formula tells us

$$\mathbb{E}_x[v(t,X(t))] = \mathbb{E}_x[v(0,x_0)]+\mathbb{E}\int_0^t \Big(\frac{\partial v}{\partial s}+\frac{\Delta v}{2}\Big)ds+\mathbb{E}[\text{martingale}] = v(0,x) = u(t,x).$$

So $\mathbb{E}_x[\phi(X(t))] = u(t,x)$.

27 April 5, 2017

27.1 Application to the Dirichlet problem

Let us look at the Dirichlet problem, given by

$$\Delta u(x) = 0 \text{ for } x \in D, \quad u(x) = g(x) \text{ for } x \in \partial D.$$

Consider the Brownian motion B_t starting from x. Then

$$u(B_t) = u(B_0) + \int_0^t \frac{\Delta}{2} u(B_s) ds + (\text{martingale term}) = u(B_0) + (\text{martingale}).$$

If we take $T = \inf\{t : B_t \notin D\}$ then stopping here and taking expectation gives, by the optional stopping theorem,

$$\mathbb{E}_x[u(B_{t\wedge T})] = u(x).$$

Let us assume that $P(T < \infty) = 1$. This is true if, in particular, D is bounded. Then

$$u(x) = \lim_{t\to\infty} \mathbb{E}_x[B_{t\wedge T}] = \mathbb{E}_x[B_T] = \mathbb{E}_x[g(B_T)].$$

Let $\mathbb{T}(x, dy)$ be the probability measure of the Brownian motion starting from x to exit at y. Then

$$u(x) = \int g(y)\mathbb{T}(x, dy). \tag{$*$}$$

This arguments shows that if u solves the equation, it has to be given by this integral. But when does $(*)$ actually solve the problem? You need some condition on D. There is something called the exterior cone condition that implies the existence of the Dirichlet problem. The same condition implies that $(*)$ solves the equation.

Consider the Dirichlet problem given by

$$g(x) = \begin{cases} 0 & \text{if } |x| = R, \\ 1 & \text{if } |x| = \epsilon. \end{cases}$$

The solution is going to be, in $d = 2$,

$$u_\epsilon^R(x) = \frac{\log R - \log|x|}{\log R - \log \epsilon},$$

and for $d = 3$,

$$u_\epsilon^R(x) = \frac{\frac{1}{R} - \frac{1}{|x|}}{\frac{1}{R} - \frac{1}{\epsilon}}.$$

This has the interpretation that $u(x)$ is the probability of the Brownian motion starting from x exits $\epsilon \le |x| \le R$ from S_ϵ.

Taking $R \to \infty$, we get

$$u_\epsilon^R(x) \to \begin{cases} 1 & \text{if } d = 2 \\ < 1 & \text{if } d \geq 3. \end{cases}$$

That is, for any $\epsilon > 0$ the Brownian motion starting from x will visit B_ϵ before going to infinity, in the case $d = 2$. For $d \geq 3$ this is not true.

27.2 Stochastic differential equations

We want to look at equations of the form

$$dX_t = \sigma(t, X_t)dB_t + b(t, X_t)dt,$$

with respect to a filtration $\mathcal{F}_t$. Something being a solution to the equation $E(\sigma, b)$ means

(i) $(\Omega, \mathcal{F}_t, P)$ is a probability space,

(ii) B_t is a Brownian motion with respect to $\mathcal{F}_t$,

(iii) there exists X_t a continuous progressively measurable with respect to $\mathcal{F}_t$ such that
$$X_t = \int_0^t \sigma(s, X_s)dB_s + \int_0^t b(s, X_s)ds + X_0.$$

Definition 27.1. A solution X is called a **strong solution** if $\mathcal{F}_t$ is adapted with respect to the canonical filtration of the Brownian motion.

Definition 27.2. A equation is **pathwise unique** if any two solutions X and X' for the same B_t has $X = X'$ almost surely.

Theorem 27.3. *If σ and b are uniformly Lipschitz then the SDE has a strong solution and is pathwise unique.*

Proof. Iterate this process, i.e., let $X_t^0 = X$ and

$$X_t^n = \int_0^t \sigma(s, X_s^{n-1})dB_s + \int_0^t b(s, X_s^{n-1})ds + X.$$

We want to check that this iteration converges. To show this, define

$$g_n(t) = \mathbb{E}\Big[\sup_{0\leq s\leq t} |X_s^n - X_s^{n-1}|^2\Big].$$

We claim that

$$g_n(t) \leq C_T'(C_T)^{n-1}\frac{t^{n-1}}{(n-1)!}.$$

We have

$$\mathbb{E}\Big[\sup_{0\le s\le t}|X_s^{n+1}-X_s^n|^2\Big] \le 2\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^t(\sigma(s,X_s^n)-\sigma(s,X_s^{n-1}))dB_s\right|^2\right] + 2\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^t(b(s,X_s^n)-b(s,X_s^{n-1}))ds\right|^2\right].$$

Using the Lipschitz condition, we are going to bound this, next time. $\square$

28 April 7, 2017

We want to prove the existence and uniqueness of stochastic PDEs.

28.1 Existence and uniqueness in the Lipschitz case

Theorem 28.1. *Consider the stochastic differential equation*

$$dX = \sigma(t, X_t)dB_t + b(t, X_t)dt, \qquad (*)$$

where $\sigma(t,x)$ and $b(t,x)$ are uniformly Lipschitz with respect to x. Then there exists a pathwise unique solution to $(*)$, i.e., there exists a $(\Omega, \mathcal{F}, P)$ with continuous sample paths such that

$$X_t = X + \int_0^t \sigma(s, X_s)dB_s + \int_0^t b(s, X_s)ds$$

almost surely.

Proof of existence. Let us define, $X_t^0 = X$ and

$$X_t^n = X + \int_0^t \sigma(s, X_s^{n-1})dB_s + \int_0^t b(s, X_s^{n-1})ds.$$

Let us define

$$g_n(T) = \mathbb{E}\Big[\sup_{0\le s\le t} |X_s^n - X_s^{n-1}|^2\Big].$$

For fixed T and $t \in [0, T]$, we have

$$\begin{aligned}
g_{n+1}(t) = \mathbb{E}\Big[\sup_{0\le s\le t} |X_s^{n+1} - X_s^n|^2\Big] &\le 2\mathbb{E}\Big[\sup_{0\le s\le t}\Big(\int_0^s (\sigma(X_u^n) - \sigma(X_u^{n-1}))dB_u\Big)^2\Big] \\
&\quad + 2\mathbb{E}\Big[\sup_{0\le s\le t}\Big(\int_0^s (b(X_u^n) - b(X_u^{n-1}))du\Big)^2\Big] \\
&\le C\mathbb{E}\Big[\int_0^t (\sigma(X_u^n)^2 - \sigma(X_u^{n-1}))^2 du + t\int_0^t (b(X_u^n) - b(X_u^{n-1}))^2 du\Big] \\
&\le C_T \mathbb{E}\Big[\int_0^t \sup_{0\le s\le u} |X_s^n - X_s^{n-1}|^2 du\Big] \le C_T \int_0^t g_n(u)du,
\end{aligned}$$

by the Burkholder–Davis–Gundy inequality and the Lipschitz condition. From this inequality, you will inductively get

$$g_{n+1}(t) \le (C_T)^n \frac{t^n}{n!}.$$

That is, for any T fixed, $\sum_{n=1}^\infty g_n(T)^{1/2} < \infty$. Hence the process converges and it is a solution. ☐

Proof of uniqueness. This actually just follows from the same arguments. If we define

$$g(t) = \mathbb{E}\Big[\sup_{0\le s\le t} |X_s - Y_s|^2\Big],$$

then

$$g(t) \le C\int_0^t g(s)ds.$$

Then by Gronwall's inequality, $g = 0$. $\square$

Let us apply Itô's formula to a solution X_t. We get

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t \nabla f_s dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X, X\rangle_s \\ &= f(X_0) + \int_0^t (\nabla f_s)\sigma(s, X_s)dB_s + \int_0^t (\nabla f_s)^2 b(s, X_s)ds \\ &\quad + \frac{1}{2}\int_0^t f''(X_s)\sigma^2(s, X_s)ds \\ &= f(X_0) + \int_0^t \Big(\frac{1}{2}\sigma^2\Delta + b\nabla\Big) f(X_s)ds + (\text{martingale}). \end{aligned}$$

We call the operator

$$L = \frac{1}{2}\sigma^2\Delta + b\nabla$$

the **generator** of the process.

The Brownian motion is a solution to $\partial_t u = \frac{1}{2}\Delta u$. This is illustrated by Itô's formula

$$f(B_t) = f(B_0) + \int_0^t \frac{1}{2}\Delta f(B_s)ds + M.$$

Likewise, to solve the equation $\partial_t u = Lu$ for $L = \frac{1}{2}\sigma^2\Delta + b\nabla$, you can solve the equation

$$ds = \sigma dB + ddt.$$

28.2 Kolomogorov forward and backward equations

Theorem 28.2. *Let $P(s, x; t, y)$ be a transition probability density such that*

$$\mathbb{E}_{x,s}[h(x(t))] = \int p(s, x; t, y)h(y)dy.$$

Then

$$\begin{aligned} &\partial_s p(s, x; t, y) + L_x p(s, x; t, y) = 0, \\ &\partial_t p(s, x; t, y) = L_y^* p(s, x; t, y). \end{aligned}$$

Proof. Let $f(t,x)$ be the probability density of $x(t)$. Denote $f_0(x) = f(0,x)$. Then formally

$$f(x,t) = \int p(0,z;t,x) f_0(z) dz.$$

Then we get

$$\int f(t,x)h(x)dx = \int f_0(x)\mathbb{E}_x[h(x(t))]dx.$$

The derivative with respect to t gives

$$\begin{aligned}
\partial_t \int f(t,x)h(x)dx &= \partial_t \int f_0(x)\mathbb{E}_x[h(x(t))]dx \\
&= \partial_t \int f_0(x)\mathbb{E}_x\Big[h(x(0)) + \int_0^t (Lh)(x(s))ds\Big] \\
&= \int f_0(x)\mathbb{E}_x(Lh)(x(t)) = \int (Lh)(x)f(t,x)dx \\
&= \int h(x)(L^*f)(t,x)dx.
\end{aligned}$$

This gives the forward equation, and its dual is the backward equation. □

29 April 10, 2017

We have looked at the Kolmogorov's equation. Given a differential equation

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt,$$

let $f(t,x)$ be the probability density of the solution X_t. Then

$$\partial_t f(t,x) = L^* f(t,x), \quad \text{where } L = \frac{\sigma(x)^2}{2}\Delta + b(x)\nabla.$$

The Kolmogorov equations say that, if the transition probability is $p(s,x;t,y)$,

$$\mathbb{E}_{x;s}[h(X_t)] = \int p(s,x;t,y)h(y)dy,$$

then

$$(\partial_s + L_x)p(s,x;t,y) = 0, \quad (\partial_t - L_y^*)p(s,x;t,y) = 0.$$

Proof. Recall that by Itô's formula,

$$h(X_t) = h(X_s) + \int_0^t (Lh)(X_s)ds + (\text{martingale}).$$

So

$$\begin{aligned}\partial_t \int f(t,x)h(x) = \partial_t \int f_0(x)\mathbb{E}_x(h(X_t)) = \int f_0(x)\mathbb{E}_x[(Lh)(X_t)] \\ = \int f(t,x)(Lh)(x)dx = \int (L^*f)(t,x)h(x)dx.\end{aligned}$$

□

More generally, if $f(t,x)$ is the density relative to μ, then

$$\partial_t \int f(t,x)h(x)\mu(dx) = \int f(t,x)(Lh)(x)\mu(dx) = \int (L_\mu^*)f(t,x)h(x)\mu(dx)$$

and so

$$\partial_t f = L_\mu^* f. \tag{$*$}$$

Definition 29.1. μ is called an **invariant measure** of the process if $L_\mu^* 1 = 0$, i.e., 1 is a solution to $(*)$.

In other words, μ is invariant if and only if

$$\int (Lf)d\mu = 0$$

for any f with compact support.

Definition 29.2. μ is called a **reversible measure** of the process if $L_\mu^* = L$, i.e.,

$$\int f(Lh)d\mu = \int (Lf)hd\mu.$$

In this case,

$$-\int h(Lh)d\mu = D_\mu(h)$$

is called the **Dirichlet form** of the process.

29.1 Girsanov's formula as an SDE

Girsanov's formula says that if

$$\frac{dQ}{dP}\bigg|_{[0,t]} = e^{\int_0^t b(X_s)dB_s - \frac{1}{2}\int_0^t b^2(X_s)ds}$$

and M is a local martingale with respect to P, then

$$\tilde{M} = M - \left\langle M, \int_0^t b(X_s)dB_s \right\rangle$$

is a local martingale with respect to Q.

Take $M = B$ with respect to P. Then this implies that

$$\beta_t = \tilde{M} = B_t - \left\langle B_t, \int_0^t b(B_s)dB_s \right\rangle = B_t - \int_0^t b(B_s)ds$$

is a martingale with respect to Q. We note that $\langle \tilde{M}, \tilde{M} \rangle_t = t$. Therefore $\tilde{M}$ is in fact a Brownian motion with respect to Q. But B_t is no longer a Brownian motion in Q. So we change notation and write X_t instead of B_t. Then

$$dX_t = d\beta_t + b(X_t)dt, \tag{†}$$

i.e., X_t is the solution to the SDE (†), i.e., the probability measure Q is the solution to (†).

This means that when you have a SDE that looks like $dX_t = \sigma dB_t + bdt$, the really important term is σdB_t, since the bdt comes from Girsanov's formula by changing the probability measure.

We will look at three examples of SDEs:

(1) Ornsetin–Uhlenbeck process

(2) Geometric Brownian motion - Black–Scholes formula

(3) Bessel process

29.2 Ornstein–Uhlenbeck process

Let us first look at the **Ornstein–Uhlenbeck process**:

$$dX_t = dB_t - \lambda X_t dt,$$

where the negative coefficient means that we are pushing back the process into the origin. If we define $X_t = e^{-\lambda t} Y_t$, then

$$dY_t = d(e^{\lambda t} X_t) = \lambda Y_t dt + e^{\lambda t} dB_t - e^{\lambda t} \lambda X_t dt$$

and so $dY_t = e^{\lambda t} dB_t$. That is,

$$X_t = e^{-\lambda t} Y_t = e^{-\lambda t} X_0 + \int_0^t e^{-\lambda(t-s)} dB_s,$$

and this is the exact solution.

The generator is

$$L = \frac{1}{2}\frac{\partial^2}{\partial x^2} - \lambda x \frac{\partial}{\partial x},$$

and you can check that the Gaussian measure $\mu = d_\lambda e^{-c_\lambda x^2}$ is an invariant measure.

29.3 Geometric Brownian motion

Let us now look at the **geometric Brownian motion**:

$$dX_t = \sigma X_t dB_t + rX_t dt.$$

If we again make a substitution $X_t = e^{rt}Y_t$, then

$$dY_t = -rY_t dt + e^{-rt} rX_t dt + e^{-rt}\sigma X_t dB_t = \sigma Y_t dB_t.$$

If we simply guess $Y_t = f(B_t, t)$, then

$$dY_t = f'(B_t, t)dB_t + \frac{f''(B_t, t)}{2}dt + (\partial_t f)(B_t, t)dt.$$

So we want $\partial_t f + f''/2 = 0$ and $f'(B_t, t) = \sigma f$. You can solve this, and you get

$$Y_t = f(B_t, t) = e^{\sigma B_t} - \frac{\sigma^2}{2}t.$$

Therefore the exact solution is given by

$$X_t = e^{rt}e^{\sigma B_t - \frac{\sigma^2}{2}t} + X_0.$$

30 April 12, 2017

Last time we looked at the geometric Brownian motion, given by

$$dX_t = \sigma X_t dB_t + \gamma X_t dt.$$

There is something called a **Black–Scholes model** for stock prices. Let S_t be the stock price at t and let β_t be the bond price at t. The value of investment is given by $V_t = a_t S_t + b_t \beta_t$. The stock dynamics is

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad d\beta_t = r\beta_t dt, \quad dV_t = a_t dS_t + b_t d\beta_t.$$

The constants μ, σ, r are given, and we want to find a_t, b_t such that $V_T = h(S_T)$ is maximal. We have

$$\begin{aligned} dV_t &= a_t dS_t + bd\beta_t = a_t(\mu S_t dt + \sigma S_t dB_t) + b_t r\beta_t dt \\ &= (a_t \mu S_t + br\beta_t)dt + a_t \sigma S_t dB_t. \end{aligned}$$

Suppose $V_t = f(t, S_t)$ is a function of t and S_t. Then by Itô's formula,

$$\begin{aligned} dV_t &= \partial_t f + \frac{1}{2} f'' d\langle S, S\rangle_t + f'(t, S_t) dS_t \\ &= \Big[\partial_t f + \frac{1}{2}\sigma^2 S^2 f'' + \mu S_t f'\Big] dt + f'\sigma S_t dB_t. \end{aligned}$$

Comparing this with the other equation gives

$$a_t = f'(t, S_t), \quad b_t = \frac{f_t + \frac{1}{2} f'' \sigma^2 S_t}{r\beta_t}.$$

Using the formula $f(t, S_t) = a_t S_t + b_t \beta_t$ gives

$$f_t + \frac{1}{2}\sigma^2 x^2 f_{xx}(t, x) + f_x(t, x)x - rf(t, x) = 0.$$

This is the Black–Scholes equation.

30.1 Bessel process

Let us now look at the **Bessel process** given by

$$dX_t = 2\sqrt{X_t} dB_t + mdt, \quad m > 0.$$

If we look at the d-dimensional Brownian motion $\beta_t = (\beta^1, \dots, \beta^d)$ and define $X_t = \beta_t^2 = (\beta^1)^2 + \dots + (\beta^d)^2$, then we can check that $dX_t = 2\sqrt{X_t} dB_t + ddt$. So $m = d$ is exactly the process of $|\beta_t|^2$.

Let us define

$$T_r = \inf\{t \geq 0 : X_t = r\},$$

where we assume $x \leq r$. We now claim that $P(T_r = \infty) = 0$.

Suppose that $\epsilon < x < r$ and define

$$M_t = \begin{cases} X_t^{1-m/2} & m \neq 2 \\ \log X_t & m = 2. \end{cases}$$

Then Itô's calculus gives

$$\begin{aligned} dM_{t\wedge T_\epsilon} &= \Big(1-\frac{m}{2}\Big)X_t^{-m/2}dX_{t\wedge T_\epsilon} + \frac{1}{2}\Big(1-\frac{m}{2}\Big)\Big(-\frac{m}{2}\Big)X^{-m/2-1}d\langle X,X\rangle_{t\wedge T_\epsilon} \\ &= \Big[\Big(1-\frac{m}{2}\Big)mX_{t\wedge T_\epsilon}^{-m/2} + \frac{1}{2}\Big(1-\frac{m}{2}\Big)\Big(-\frac{m}{2}\Big)X_{t\wedge T_\epsilon}^{-m/2}\Big]dt \\ &\quad + 2\Big(1-\frac{m}{2}\Big)X_{t\wedge T_\epsilon}^{-m/2+1/2}dB_{t\wedge T_\epsilon} = 2\Big(1-\frac{m}{2}\Big)X_{t\wedge T_\epsilon}^{-m/2+1/2}dB_{t\wedge T_\epsilon}. \end{aligned}$$

Thus

$$M_{t\wedge T_\epsilon} = 2\Big(1-\frac{m}{2}\Big)\int_0^{t\wedge T_\epsilon} X_s^{-m/2+1/2}dB_s.$$

This is a martingale for ϵ fixed, and just a local martingale if $\epsilon \to 0$. By optional stopping, we have $\mathbb{E}[M_{t\wedge T_\epsilon \wedge T_A}] = x^{1-m/2}$ and if you solve the equation, we get

$$P(T_\epsilon < T_A) = \begin{cases} \dfrac{A^{1-m/2} - x^{1-m/2}}{A^{1-m/2} - \epsilon^{1-m/2}} & m \neq 2 \\ \dfrac{\log A - \log x}{\log A - \log \epsilon} & m = 2. \end{cases}$$

This shows that for $m \geq 2$, as $\epsilon \to 0$, the probability that the Bessel process reach 0 is zero.

31 April 14, 2017

Last time we talked about the Bessel process. This is given by

$$dX_t = 2\sqrt{X_t}dB_t + mdt.$$

An example, when $m = d$ is an integer, is $X_t = |B_t^d|^2$. If we define $M_t = X_t^{1-m/2}$ then

$$dM_t = \Big(1 - \frac{m}{2}\Big)X_t^{-m/2}2\sqrt{X_t}dB_t.$$

Then M_t is a local martingale.

In the case of $m = 3$, we would have $M_t = |B_t|^{-1}$, and so $d|B_t|^{-1} = C|B_t|^{-2}dB_t$. Then

$$|B_t|^{-1} = \int_0^t |B_s|^{-2}dB_s$$

is a local martingale.

Recall that $\mathbb{E}(\int_0^t |B_s|^{-2}dB_s)^2 < \infty$ then $|B_t|^{-1}$ would be a martingale. However, you can check that

$$\mathbb{E}\Big[\int_0^t |B_s|^{-4}ds\Big] = \infty.$$

Indeed, $|B_t|^{-1}$ is a local martingale but not a martingale. Likewise, in $d = 2$, $M_t = \log|B_t|$ is a local martingale but not a martingale.

31.1 Maximum principle

Theorem 31.1. *Suppose you want to solve the equation*

$$\partial_t u + \frac{1}{2}\Delta u + g = 0, \quad u(T, x) = f(x)$$

with boundary condition $u(t, x) = k(t, x)$ for $x \in \partial\Omega$ and $T > 0$. Or more generally, consider the equation $\partial_t u + Lu + g = 0$ (in this case, we use the stochastic process generated by L instead of the Brownian motion). Then

$$u(0, x) = \mathbb{E}_x\Big[f(x, T))\mathbf{1}_{(\tau\geq T)} + k(\tau, x(\tau))\mathbf{1}_{(\tau<T)} + \int_0^{\tau\wedge T} g(s, x(s))\Big].$$

Proof. From Itô's formula, for a process X,

$$u(t, X(t)) = \int_0^t \Big(\partial_t u + \frac{\Delta}{2}u\Big)ds + (\text{martingale}) + u(0, x).$$

If we define $\tilde{\tau} = \tau \wedge T$, then

$$\mathbb{E}_x[u(\tau \wedge T, X(\tau \wedge T))] + \mathbb{E}_x \int_0^{\tau\wedge T} g(s, X_s)ds = u(0, x).$$

This proves the theorem. $\square$

Corollary 31.2. *If $f_2 \geq f_1$, $g_2 \geq g_1$, and $k_2 \geq k_1$, then $u_2(s,x) \geq u_1(s,x)$, i.e., we get monotonicity. Moreover, if $g = 0$, the maximum of $u(t,x)$ in $[0,T] \times \Omega$ can be achieved on the boundary. If the maximum is achieved in the interior, then u is a constant.*

Proof. If $\sup f \vee \sup k \leq M$, then $\mathbb{E}_x \leq M$. □

31.2 2-dimensional Brownian motion

Consider a 2-dimensional Brownian motion $B_t = X_t + iY_t$.

Theorem 31.3 (Conformal invariance, Theorem 7.8)**.** *Suppose Φ is analytic. Then $\Phi(B)$ is a Brownian motion up to a time change.*

Proof. Let us compute $d\Phi(B)$. We have

$$d\Phi(B) = d(g(X_t,Y_t) + ik(X_t,Y_t)) = \frac{\partial g}{\partial x} dX_t + \frac{\partial g}{\partial y} dY_t + i\Big(\frac{\partial k}{\partial x} dX_t + \frac{\partial k}{\partial y} dY_t\Big).$$

The quadratic variation of one component is

$$\langle dg, dg \rangle = \Big[\Big(\frac{\partial g}{\partial x}\Big)^2 + \Big(\frac{\partial g}{\partial y}\Big)^2\Big] dt.$$

So $g(X_t,Y_t)$ is a Brownian motion with time change. The similar thing holds for k. Moreover,

$$\langle dg, dk \rangle = \Big\langle \frac{\partial g}{\partial x} dX + \frac{\partial g}{\partial y} dY, \frac{\partial k}{\partial x} dX + \frac{\partial k}{\partial t} dY \Big\rangle = \Big(\frac{\partial g}{\partial x}\frac{\partial k}{\partial x} + \frac{\partial g}{\partial y}\frac{\partial k}{\partial y}\Big) dt = 0$$

from the Cauchy–Riemann equations. □

32 April 17, 2017

Today we are going to look more at the 2-dimensional Brownian motion.

32.1 More on 2-dimensional Brownian motion

Theorem 32.1 (7.19)**.** *Let B_t be a Brownian motion. Then*

$$B_t = \exp(\beta_{H_t} + i\gamma_{H_t})$$

where $H_t = \int_0^t ds/|B_s|^2$ and β and γ are independent Brownian motions.

Proof. We want $\beta_{H_t} + i\gamma_{H_t} = \log B_t$. Differentiate both sides, and we get

$$d\log B_t = \frac{dX_t + idY_t}{X_t + iY_t} - \frac{1}{2}\Big(\frac{1}{X_t + iY_t}\Big)^2 (dX_t + idY_t)^2 = \frac{dX_t + idY_t}{X_t + iY_t}.$$

Then we get

$$d\beta_t = \frac{XdX + YdY}{(X^2 + Y^2)}, \quad d\gamma_t = \frac{XdY - YdX}{X^2 + Y^2}.$$

Now we have

$$\langle d\beta_t, d\beta_t\rangle = \frac{(XdX + YdY)^2}{(X^2+Y^2)^2} = \frac{(X^2+Y^2)dt}{(X^2+Y^2)^2} = \frac{dt}{(X^2+Y^2)},$$
$$\langle d\gamma_t, d\gamma_t\rangle = \cdots = \frac{dt}{(X^2+Y^2)},$$
$$\langle d\beta_t, d\gamma_t\rangle = \frac{(XdX + yDY)(XdY - YdX)}{(X^2+Y^2)^2} = 0.$$

What we want is actually β_{H_t} instead of β_t. If β is a Brownian motion, then

$$\langle d\beta_{H_t}, d\beta_{H_t}\rangle = dH_t = \frac{1}{X_t^2 + Y_t^2} dt$$

and so we see that β_{H_t} for β a Brownian motion works. □

From this we get, we get

$$dH_t = \frac{1}{|B_t|^2} dt = \frac{1}{\exp(2\beta_{H_t})} dt$$

and so

$$\exp(2\beta_{H_t}) dH_t = dt.$$

Lemma 32.2. *Let $T_1^{(\lambda)} = \inf\{t \geq 0; \beta_t^{(\lambda)} = 1\}$, where $B_t^{(\lambda)} = \lambda^{-1}\beta_{\lambda^2 t}$ is the rescaled Brownian motion. Then*

$$\frac{4}{(\log t)^2} H_t - T_1^{(\log t/2)} \to 0$$

in probability as $t \to \infty$.

Proof. From $\exp(2\beta_{H_t})dH_t = dt$, we have

$$\inf_s \left\{ \int_0^s e^{2\beta_u} du = t \right\} = H_t.$$

Let $\lambda_t = \log t/2$. Then the lemma tells us that

$$P(H_t > \lambda_t^2 T_{1+\epsilon}^{\lambda_t}) \to 0.$$

Then

$$\begin{aligned} P(H_t > \lambda_t^2 T_{1+\epsilon}\lambda_t) &= P\left[\int_0^{\lambda_t^2 T_{1+\epsilon}^{(\lambda_t)}} e^{2\beta_u} du < t \right] \\ &= P\left[\frac{\log \lambda_t}{\lambda_t} + \frac{1}{2\lambda_t} \log \int_0^{T_{1+\epsilon}^{(\lambda_t)}} e^{2\lambda_t \beta_v^{(\lambda_t)}} dv < 1 \right] \\ &\leq P\left[\frac{1}{2\lambda_t} \log \int_0^{T_{1+\epsilon}} e^{2\lambda_t \beta_v} dv < 1 \right] = P\Big(\max_{0 \leq v \leq T_{1+\epsilon}} \beta_v < 1 \Big) = 0 \end{aligned}$$

after the rescaling $u = \lambda_t^2 v$. □

Theorem 32.3. *If B_s is a 2-dimensional Brownian motion starting from any $z \neq 0$,*

$$\lim_{t\to\infty} P\Big(\inf_{0\leq s\leq t} |B_s| \leq t^{-a/2} \Big) = \frac{1}{1+a}.$$

Proof. Due to scaling, we as may as well assume that $z = 1$. We have

$$\log\Big(\inf_{0\leq s\leq t} |B_s| \Big) = \inf_{0\leq s\leq t} \beta_{H_s} = \inf_{0\leq u\leq H_t} \beta_u$$

by monotonicity. Now we have

$$\frac{2}{\log t} \log\Big(\inf_{0\leq s\leq t} |B_s| \Big) = \frac{1}{\lambda_t} \inf_{0\leq u\leq H_t} \beta_u = \inf_{0\leq s\leq \lambda_t^{-2} H_t} \beta_s^{(\lambda_t)}.$$

By the lemma, $\lambda_t^{-2} H_t$ converges to $T_1^{(\lambda_t)}$ and also

$$\inf_{0\leq s\leq T_1^{(\lambda_t)}} \beta_s^{(\lambda_t)} = \inf_{0\leq s\leq T_1} \beta_s.$$

We will continue next time. □

33 April 19, 2017

Last time we showed that a 2-dimensional Brownian motion can be expressed as

$$B_t = e^{\beta_{H_t} + i\gamma_{H_t}}, \quad H_t = \int_0^t \frac{1}{|B_s|^2} ds,$$

where β and γ are independent Brownian motions.

Lemma 33.1 (7.21)**.** *In probability,*

$$\frac{4}{(\log t)^2} H_t - T_1^{(\log t/2)} \to 0$$

as $t \to \infty$, where $T_1^{(\lambda)} = \inf\{s \geq 0, \beta_t^{(\lambda)} = 1\}$.

Proposition 33.2 (7.22)**.** $\lim_{t\to\infty} P\Big(\inf_{0\leq s\leq t} |B_s| \leq t^{-a/2}\Big) = \dfrac{1}{1+a}$.

Proof. Because H_t is monotone, we have

$$\log\Big(\inf_{0\leq s\leq t} |B_s|\Big) = \inf_{0\leq s\leq t} \beta_{H_s} = \inf_{0\leq u\leq H_t} \beta_u.$$

Then

$$\frac{2}{\log t} \log\Big(\inf_{0\leq s\leq t} |B_s|\Big) = \frac{1}{\lambda_t} \inf_{0\leq u\leq H_t} \beta_u = \inf_{0\leq v\leq H_t/\lambda_t^2} \beta_v^{(\lambda_t)} = \inf_{0\leq v\leq T_1^{(\lambda_t)}} \beta_v^{(\lambda_t)}$$

after the rescaling $u = \lambda_t^2 v$. Then the probability is

$$\begin{aligned} P\Big(\frac{1}{\lambda_t} \log \inf_{0\leq s\leq t} |B_s| \leq -a\Big) &= P\Big(\inf_{0\leq v\leq T_1^{(\lambda_t)}} \beta_v^{(\lambda_t)} \leq -a\Big) \\ &= P\Big(\inf_{0\leq v\leq T_1} \beta_v \leq -a\Big) = P(T_{-a} < T_1) = \frac{1}{1+a} \end{aligned}$$

because of the property of the 1-dimensional Brownian motion. (You can see this from solving the Laplace equation with $u(-a) = 1$ and $u(1) = 0$.) $\square$

33.1 Feynman–Kac formula

Theorem 33.3. *Consider an equation*

$$\partial_t u + \frac{1}{2}\Delta u + V(u) = 0$$

with the boundary condition $u(T, x) = f(x)$. Then

$$u(0, x) = \mathbb{E}_x\Big[\exp\Big(\int_0^t V(x(s)) ds\Big) f(x(T))\Big].$$

Proof. Take a u solving the equation. Then

$$\begin{aligned} d\Big[u(t,x(t))\exp\Big(\int_0^t V(x(s))ds\Big)\Big] &= \Big(\partial_t u + \frac{1}{2}\Delta u\Big)\exp\Big(\int_0^t V(x(s))ds\Big)dt \\ &\quad + \nabla_x u(t,x(t))dx(t) + u\exp\Big(\int_0^t V(x(s))ds\Big)V(x(t))dt \\ &\quad + du\, d\exp\Big(\int_0^t V(x(s))ds\Big) \\ &= (\partial_t u + \frac{1}{2}\Delta u + Vu)\exp\Big(\int_0^t V(x(s))ds\Big)dt + (\text{martingale}) \end{aligned}$$

is a martingale. So we get

$$u(0,x) = \mathbb{E}_x\Big[\exp\Big(\int_0^T V(x(s))ds\Big)f(x(T))\Big]. \qquad \square$$

Here is an intuitive proof. The equation can be written as $\partial_t u = (\frac{1}{2}\Delta + V)(u)$ and so we can write

$$u = e^{t(\frac{1}{2}\Delta + V)}u_0.$$

There is also the Trotter product formula

$$e^{A+B} = \lim_{n\to\infty}(e^{A/n}e^{B/n})^n.$$

Then our solution can be thought of as

$$e^{\frac{\Delta}{2}\frac{t}{n}}e^{\frac{t}{n}V}e^{\frac{\Delta}{2}\frac{t}{n}}e^{\frac{t}{n}V}\cdots = \frac{1}{\sqrt{2\pi\cdots}}e^{-\frac{(x_0-x_1)^2}{2(t/n)}}e^{\frac{t}{n}V(x_1)}e^{0\frac{(x_1-x_2)^2}{2(t/n)}}e^{\frac{t}{n}V(x_2)}\cdots.$$

Then path $x_0 \to x_1 \to x_2 \to \cdots$ becomes a discrete approximation of the Brownian motion, and the factor we pick up is

$$e^{\frac{t}{n}(V(x_1)+\cdots+V(x_n))} \to e^{\int_0^t V(x(s))ds}$$

under the limit $n \to \infty$.

33.2 Girsanov formula again

There is a nice proof of Girsanov's formula. Let $b(t, X(t))$ be a function and suppose Q solves the equation

$$dX = bdt + d\beta.$$

Let P be the law of β. Then we want to prove that

$$\frac{dQ}{dP} = \exp\Big(\int_0^t b(s,X(s))dx(s) - \frac{1}{2}\int_0^t b^2 ds\Big).$$

Let us define

$$Z(t) = \int_0^t b(s, X(s))dX(s)$$

and consider the object

$$d\exp\left(\lambda X(t) + Z(t) - \frac{1}{2}\int_0^t (\lambda^2 + 2\lambda b(s, X(s)) + b^2)ds\right).$$

This is going to be a martingale plus a dt term. Our claim is that the dt term vanish. We check that Itô calculus gives

$$\lambda dX - \frac{1}{2}(\lambda^2 + 2\lambda b + b^2)dt + dZ + \lambda dXdZ + \frac{1}{2}\lambda^2 dXdX + \frac{1}{2}dZdZ.$$

Here $\lambda dXdZ = bdt$ and $\frac{1}{2}\lambda^2 dXdX = \frac{\lambda^2}{2}dt$ and $\frac{1}{2}dZdZ = \frac{b^2}{2}dt$. So the dt term indeed vanishes. Everything follows from this.

34 April 21, 2017

The Girsanov formula says that

$$\frac{dQ}{dP} = \exp\bigg(\int_0^t b(s, X(s))dX(s) - \frac{1}{2}\int_0^t b^2(s, X(s))ds\bigg),$$

where Q solves $dX = b(t, X(t))dt = d\beta$.

The idea is that the Brownian motion measures something like

$$(\text{const})e^{-\frac{(x_2-x_1)^2}{2t/n} - \frac{(x_3-x_2)^2}{2t/n} - \cdots}.$$

Here the discrete difference $X_t - X_{t+1}$ is like $d\beta(t) = dX - b(t, X(t))dt$ and so this quantity is like

$$\exp\Big(-\frac{1}{2t/n}\Big[x_2 - x_1 - b(t_1, x_1)\frac{t}{n}\Big]^2 - \cdots\Big) = \exp\Big(-\frac{(x_1-x_2)^2}{2t/n} + b(t_1, x_1)(x_2 - x_1) - \cdots\Big).$$

Here $\sum_j b_(t_j, x_j)(x_{j+1} - x_j)$ is like $\int_0^t b(s, x(s))ds$. So this is why we get Girsanov's formula.

34.1 Kipnis–Varadhan cutoff lemma

Theorem 34.1. *Suppose μ is a reversible measure with respect to a process with generator L. Let $D(f) = -\int fLf d\mu$ be the Dirichlet form. Then*

$$P^\mu\Big(\sup_{0\le s\le t} |f(X(s))| \ge \ell\Big) \le \frac{1}{\ell}\sqrt{\|f\|^2_{L^2\mu} + D(f)}.$$

Proof. Suppose $\partial_t u + Lu + g = 0$. Then by Itô calculus,

$$du(t, x) = (\partial_t u + Lu)dt + (\text{martingale}).$$

Now consider the stopping time $\tau = \inf\{s \ge 0, f(X(s)) = \ell\}$. Then

$$u(t\wedge\tau, x) = \int_0^{t\wedge\tau} -g(s, X(s))ds + u(0, x) + (\text{martingale}).$$

Suppose a function ϕ_λ solves the equation

$$\begin{cases} (\lambda - L)\phi = 0 & x \in G = \{y : f(y) < \ell\}, \\ \phi_\lambda = 1 & x \in \partial G. \end{cases}$$

If we take $u(t, x) = \phi_\lambda(x)e^{-\lambda t}$ then

$$\partial_t u + Lu = -\lambda\phi_\lambda(x)e^{-\lambda t} + \lambda\phi_\lambda(x)e^{-\lambda t} = 0.$$

Then you can show that $\phi_\lambda = \mathbb{E}_x e^{-\lambda\tau}$. Then we get

$$P(\tau < t) \le e^{\lambda t}\mathbb{E}^\mu\phi_\lambda(x) \le e^{\lambda t}\sqrt{\mathbb{E}^\mu[\phi_\lambda(x)^2]}$$

by Chebyshev's inequality.

But recall that ϕ_λ is an eigenfunction of L. But note that for any h with $h = 1$ on ∂G, we have $\langle h, (\lambda - L)h\rangle\mu \geq \langle \phi_\lambda, (\lambda - L)\phi_\lambda\rangle_\mu$ and so

$$\|\phi_\lambda\|_{L^2}^2 \leq \|h\|_2^2 + \frac{1}{\lambda} D(h).$$

Therefore

$$P(\tau < t) \leq e^{\lambda t} \sqrt{\|h\|_2^2 + \frac{1}{\lambda} D(h)}.$$

Take $\lambda = 1/t$ and $h = \ell^{-1}(|f| \wedge \ell)$ gives the theorem. $\square$

35 April 24, 2017

Let μ be a reversible ($\int gLf d\mu = \int (Lg) f d\mu$) measure of a process with Dirichlet form $D = -\int fLf d\mu$. Then we want to show

$$P\Big(\sup_{0\le s\le t} |f(X(s))| \ge \ell\Big) \le \frac{c}{\ell}\sqrt{\|f\|^2_{L^2_\mu} + tD(f)}.$$

Proof. Let $G = \{x : |f(x)| \ge \ell\}$ and τ is the stopping time to exit the open set $\{y : f(y) < \ell\}$. Suppose ϕ_λ is the ground state such that

$$(\lambda - L)\phi_\lambda(x) = 0 \text{ for } x \in G, \quad \phi_\lambda(x) = 1 \text{ for } x \in \partial G.$$

This being a ground state means that λ is minimal. So ϕ_λ is nonnegative. (The first eigenfunction of an elliptic problem is nonnegative.)

If we define $u(t,x) = \phi_\lambda(x)e^{-\lambda t}$ then $\partial_t u + Lu = 0$. Then Itô's formula implies that

$$u(0,x) = -\mathbb{E}_x\left[\int_0^t (\partial_s + L)u(s,x)ds\right] + \mathbb{E}_x[u(t,x)] = \mathbb{E}_x[u(t,x)].$$

So

$$\begin{aligned}\phi_\lambda(x) = \mathbb{E}_x[u(t\wedge\tau,x)] &= \mathbb{E}_x[u(\tau,x)\mathbf{1}_{(\tau\le t)}] + \mathbb{E}_x[u(t,x)\mathbf{1}_{(\tau>t)}] \\ &= \mathbb{E}_x[e^{-\lambda\tau}\mathbf{1}_{(\tau\le t)} + e^{-\lambda t}\phi_\lambda(x)\mathbf{1}_{(\tau>t)}].\end{aligned}$$

Now we have, with respect to μ,

$$e^{-\lambda t}P_\mu(\tau\le t) \le \mathbb{E}^\mu[e^{-\lambda\tau}\mathbf{1}_{(\tau\le t)} \le \mathbb{E}^\mu\phi_\lambda(x) \le (\mathbb{E}^\mu(\phi_\lambda(x)^2))^{1/2}.$$

Since ϕ_λ is the minimizer, we have

$$\langle\phi_\lambda, (\lambda - L)\phi_\lambda\rangle \le \langle h, (\lambda - L)h\rangle$$

for any h such that $h(x) = 1$ on ∂G. This implies that $\lambda\|\phi_\lambda\|^2 + D(\phi_\lambda) \le \lambda\|h\|^2 + D(h)$. If we let $h = \frac{1}{\ell}\min(|f(x)|,\ell)$, we have

$$\|\phi_\lambda\|_2^2 \le \|h\|_2^2 + \frac{1}{\lambda}D(h) = \frac{1}{\ell^2}\|f\|_2^2 + \frac{1}{\ell^2\lambda}D(f).$$

Plugging this into the previous formula gives

$$P_\mu(\tau\le t) \le \frac{e^{\lambda t}}{\ell}\sqrt{\|f\|_2^2 + \frac{1}{\lambda}D(f)} \le \frac{e}{\ell}\sqrt{\|f\|_2^2 + tD(f)}.$$ □

35.1 Final review

Here are the things we have done so far:

- construction of Brownian motions (Gaussian processes, Lévy's construction)

- martingales (discrete)
- local martingales (continuous time); roughly speaking, M is a local martingale if it can be appproximated by a martingale
- a nonnegative local martingale is a supermartingale
- construction of quadratic variation $\langle M, M \rangle_t$
- if $\mathbb{E}[\langle M, M \rangle_t] < \infty$ for each t then M is a martingale (Theorem 4.13)
- stochastic integration $\int HdM$, when $\int H^s d\langle M, M \rangle < \infty$ almost surely
- if $\mathbb{E}[\int H^s d\langle M, M \rangle_\infty] < \infty$ then it is a marginale
- Itô's formula (in general, we only know that $\int \cdots dM$ is a local martingale)
- Burkholder–Davis–Gundy inequality
$$\mathbb{E}[(M_T^*)^p] \leq C_p \mathbb{E}[\langle M, M \rangle_T^{p/2}].$$
- Levi's theorem for characterizing Brownian motions, $\langle M, M \rangle_t = t$
- dominated convergence, $\int_0^t H_s^n dX_s \to \int_0^t H_s dX_s$ in probability if $H_s^n \to H_s$ and $|H_s^n| \leq K_s$ and $\int_0^t K_s^2 d\langle X, X \rangle_s < \infty$ (Proposition 5.18)
- existence and uniqueness of $dX = bdt + \sigma d\beta$ for σ and β Lipschitz
- applications of Itô's formula; $dF = \nabla F dM + \frac{1}{2} \nabla^2 F d\langle M, M \rangle$, check that $\nabla F dM$ is a martingale, and then take expectation
- Feynman–Kac formaula, Girnasov, Dirichlet problem, Kipnis–Varadhan, ...

You will not remember the details, and I will not remember the details. But it is important to have the big picture.

Index