njb2154
Noah J. X. Bergam

Advanced Probability

MATH4156
December 28, 2023

1 Week 1: Review of Conditional Expectation

Definition 1. *Let μ, ν be two measures on the same probability space $(\Sigma, \mathcal{F})$.*
We say $\mu \ll \nu$ (absolutely continuous) if $\nu(A) = 0 \implies \mu(A) = 0$.
We say $\mu \perp \nu$ (singular) if $\exists A \in \mathcal{F}$ such that $\nu(A) = 0$ and $\mu(A^C) = 0$.
We say $\mu \sim \nu$ (equivalent) if $\mu \ll \nu \ll \mu$.

Theorem 1 (Radon-Nikodym)**.** *Let μ be a σ-finite measure and ν a finite measure on $(\Omega, \mathcal{F})$. Then there exists a unique (up to μ-almost everywhere equivalence) function $h : \Omega \to [0, \infty)$ integrable with respect to μ such that:*

$$\nu(A) = \int_A h(\omega)\ d\mu(\omega)$$

Theorem 2 (Pinsker-Csiszar Inequality)**.**

$$2\|\mu - \nu\|_{TV}^2 \leq D(\nu|\mu)$$

Theorem 3 (Existence of Conditional Expectation)**.** *On $(\Sigma, \mathcal{F}, \mathbb{P})$ let X be an integrable random variable and $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$. Then there exists a $\mathbb{P}$-almost everywhere unique random variable $H : \Omega \to \mathbb{R}$, denoted $H = \mathbb{E}(X \mid \mathcal{G})$ such that:*

$$\int_G H d\mathbb{P} = \int_G H d\mathbb{P} \qquad \forall\ G \in \mathcal{G}$$

In equivalent notation:

$$\mathbb{E}^{\mathbb{P}}[H \cdot \mathbf{1}_G] = \mathbb{E}^{\mathbb{P}}[X \cdot \mathbf{1}_G] \qquad \forall\ G \in \mathcal{G}$$

Proof. Suppose $X \geq 0$. Then $G \to \nu(G) = \int_G X\ d\mathbb{P}$ is a measure, and finite by integrability of X ($\nu(\Omega) = \mathbb{E}(X) < \infty$). By Radon-Nikodym Theorem, there exists $H : \Omega \to [0, \infty)$ such that $\nu(G) = \int_G H\ d\mathbb{P}$ which proves the claim. □

2 Week 2: Stopping Times and Doop Decomposition

Definition 2. *A **filtration** $\mathbb{F} = (F_n)_{n\in\mathbb{N}}$ is an increasing sequence of σ-algebras.*
*A **filtered probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{F})$.*
A sequence of random variables $(Y_n)_{n\in\mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathbb{F})$ is:

- ***Adapted*** *if $\sigma(Y_n) \subset F_n$ for all n.*
- ***Predictable*** *if $\sigma(Y_n) \subset F_{n-1}$ for all n.*

*A **stopping time** $\tau : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ is a measurable map s.t. $\{\omega : \tau(\omega) \leq n\} \in \mathcal{F}_n \ \forall n$.*

Definition 3. *To any τ a stopping time we can associate a σ-algebra of events generated up to that stopping time. This is the subset of measurable sets such that their intersection with $\{\tau \leq n\}$ if $\mathcal{F}_n$ measurable for all n, i.e.*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq n\} \in \mathcal{F}_n \ \forall n\}$$

Definition 4. *A **martingale** on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ is a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ such that:*

$$\text{(martingale)} \qquad \mathbb{E}(X_n \mid \mathcal{F}_m) = X_m \qquad \forall m \leq n$$

We also define increasing and decreasing counterparts.

$$\text{(supermartingale)} \qquad \mathbb{E}(X_n \mid \mathcal{F}_m) \leq X_m \qquad \forall m \leq n$$

$$\text{(submartingale)} \qquad \mathbb{E}(X_n \mid \mathcal{F}_m) \geq X_m \qquad \forall m \leq n$$

We can also define martingales in continuous time, with a filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$ and with $s \leq t$ real numbers.

Theorem 4 (Doob Decomposition)**.** *Every submartingale X can be rewritten as $X_n = M_n + A_n$ with M a martingale and A non-decreasing. If A is chosen to be predictable, this decomposition is unique.*

Theorem 5 (Doob's Optional Sampling)**.** *On a filtered probability space consider a supermartingale X and a stopping time τ. Then we have $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$ provided that any of the following hold:*

- *τ is bounded.*
- *X is bounded (exists a uniform upper bound for all n).*
- *X is finite in expectation and X has bounded increments.*

3 Week 3: Uniformly Integrable, Square-Integrable

Theorem 6 (Doob Martingale Convergence)**.** *If a supermartingale X is suitably lower bounded, i.e. $\sup_{n\in\mathbb{N}} X_n^- < \infty$, then $X_\infty = \lim_{n\to\infty} X_n$ exists (almost everywhere) and is integrable, $\mathbb{E}(|X_\infty|) < \infty$.*

3.1 Uniformly Integrable Martingales

Definition 5. *A family $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ s called **uniformly integrable** if:*

$$\lim_{\lambda\to\infty}\sup_{\alpha\in\mathcal{A}}\mathbb{E}[|X_\alpha|\cdot \mathbf{1}_{\{|X_\alpha|>\lambda\}}]=0$$

Bounded in L^p for $p>1 \implies$ Uniformly Integrable $\implies$ Bounded in L^1.

Theorem 7 (Generalized DCT)**.** *Let (X_n) converge in probability to X. TFAE:*

- *(X_n) are uniformly integrable.*
- *$X_n \to X$ in L^1, i.e. $\mathbb{E}(|X_n|)\to\mathbb{E}(|X|)$ or $\mathbb{E}(|X_n-X|)\to 0$.*

Definition 6 (Levy Martingale)**.** *Let X be an integrable random variable and $\mathcal{F}_n$ an arbitrary filtration. Then $X_n=\mathbb{E}(X\mid\mathcal{F}_n)$ is a martingale.*

Remark 1. *Uniformly integrable martingales are Levy martingales!*

Theorem 8. *Let X be a martingale. Then the following are equivalent:*

- *X is uniformly integrable.*
- *X converges in L^1 to some $X_\infty\in L^1$.*
- *X converges a.e. to some X_∞ and becomes a martingale with last element.*
- *There exists integrable Z such that $\mathbb{E}(Z\mid\mathcal{F}_n)=X_n$.*

Theorem 9. *The origin is absorbing for a non-negative supermartingale, i.e. for $\tau=\min\{n\geq 0: X_n=0\}$ we have $X_{\tau+k}=0$ for all $k\in\mathbb{N}$ for X a martingale*

Proof. We use the following important fact: for stopping times $\sigma\leq\tau$, $\mathbb{E}(X_\tau\mid\mathcal{F}_\sigma)=X_\sigma$ (mutatis mutandis with super and sub). For us, $0\leq\mathbb{E}(X_{\tau+k})\leq\mathbb{E}(X_\tau)\leq 0$. □

3.2 Square-Integrable Martingales

Definition 7. *For M such that $\mathbb{E}(M_n^2)<\infty$ for all n (square integrable) define the **quadratic variation** or **bracket** $\langle M\rangle$ to be the unique predictable sequence that makes $M_n^2-\langle M\rangle$ a martingale (by Doob decomposition). More explicitly, we may write: $\langle M\rangle_0=0$ and*

$$\langle M\rangle_n=\sum_{k=1}^n\Big[\mathbb{E}(M_k^2\mid F_{k-1})-M_{k-1}^2\Big]=\sum_{k=1}^n\mathbb{E}\Big[(M_k-M_{k-1})^2\mid\mathcal{F}_{k-1}\Big]$$

Theorem 10 (Pythagorean Relationship)**.** *For M a square integrable martingale, non-overlapping intervals are orthogonal, so:*

$$\mathbb{E}\Big[(M_{n+j} - M_n)^2\Big] = \sum_{k=n+1}^{n+j} \mathbb{E}\Big[(M_k - M_{k-1})^2\Big]$$

Definition 8. *For M, N square integrable martingales define the* ***cross-variation*** *or* ***cross-bracket*** *with $\langle M, N\rangle_0 = 0$ and:*

$$\langle M, N\rangle_n = \sum_{k=1}^{n} \mathbb{E}\Big[(M_k - M_{k-1})(N_k - N_{k-1}) \mid \mathcal{F}_{k-1}\Big]$$

Lemma 1. *$MN - \langle M, N\rangle$ is a martingale if M, N are square integrable martingales.*

Proof. Take $j \geq 0$. Expand definition and apply martingale property.

$$\begin{aligned}
\mathbb{E}\Big[\langle M, N\rangle_{n+j} - \langle M, N\rangle_n \,\Big|\, \mathcal{F}_n\Big] &= \mathbb{E}\Big[\sum_{k=n}^{n+j} \mathbb{E}\Big[(M_k - M_{k-1})(N_k - N_{k-1}) \mid \mathcal{F}_{k-1}\Big] \,\Big|\, \mathcal{F}_n\Big] \\
&= \sum_{k=n}^{n+j} \mathbb{E}\Big[(M_k - M_{k-1})(N_k - N_{k-1}) \mid \mathcal{F}_n\Big] \\
&= \mathbb{E}\Big[M_{n+j}N_{n+j} - M_n N_n \,\Big|\, \mathcal{F}_n\Big]
\end{aligned}$$

□

Definition 9. *M, N are* ***orthogonal*** *if $\langle M, N\rangle = 0$.*

Theorem 11 (Convergence)**.** *For M a square-integrable martingale, $\lim_{n\to\infty} M_n$ exists almost everywhere on the event $\{\langle M\rangle_\infty < \infty\}$.*

Theorem 12 (SLLN)**.** *For M a square-integrable martingale:*

$$\lim_{n\to\infty} \frac{M_n}{1 + \langle M\rangle_n} = 0 \qquad a.e. \text{ on } \{\langle M\rangle_\infty = \infty\}$$

Theorem 13 (Kolmogorov Three-Series)**.** *For $(\xi_n)_{n\in\mathbb{N}}$ the series $\sum_n \xi_n$ converges in the reals if and only if the following hold for some $K \in (0, \infty)$.*

- $\sum_n \mathbb{P}(|\xi_n| > K) < \infty$
- $\sum_n \mathbb{E}(\xi_n \cdot \mathbf{1}_{\{|\xi_n| \leq K\}})$ *converges in* $\mathbb{R}$.
- $\sum_n Var(\xi_n \cdot \mathbf{1}_{\{|\xi_n| \leq K\}}) < \infty$.

3.3 Markov Chains

Definition 10. *For $g : S \to \mathbb{R}$ a numerical characteristic of some Markov Chain with state space S and transition probabilities $\{p_{ij}\}_{i,j \in S}$, define:*

$$(\Pi g)(i) = \sum_{k \in S} p_{ik} \; g(k)$$

- *g is harmonic if $\Pi g = g$.*
- *g is super-harmonic if $\Pi g \leq g$.*
- *g is sub-harmonic if $\Pi g \geq g$.*

Remark 2. *The following is a martingale, for X_n a Markov Chain:*

$$M_0^g = 0 \qquad M_n^g = g(X_n) - g(X_0) - \sum_{i=1}^{n} \Big[(\Pi g)(X_n) - g(X_n) \Big]$$

Theorem 14. *Every non-negative superharmonic function on an irreducible, recurrent Markov Chain is constant.*

4 Week 4: Some Optimization

4.1 Discrete Time Optimal Stopping

Let S_m denote the set of stopping times $\geq m$.

Optimal Stopping Problem: Take Y a sequence of non-negative, integrable random variables. Find $\tau^* \in S_0$ which maximizes $\mathbb{E}(Y_\tau)$.

Trivial Case: Consider a deterministic process $\{Y_n\}_{n \in \mathbb{N}}$, with $\mathcal{F} = \{\Omega, \phi\}$. We want p such that $Y_p = \sup_{n \in \mathbb{N}} Y_n$. A sophisticated way to study this: check that:

$$p^* = \min\{p \in \mathbb{N} : \sup_{n \geq p} Y_n = Y_p\}$$

$$\text{satisfies} \quad Y_{p^*} = \sup_{n \geq p^*} Y_n = \sup_{n \geq 1} Y_n$$

In general, $Z_n = \sup_{m \geq n} Y_n$ is a supermartingale (rather than decreasing), and we find that:

$$\tau^* = \inf\{n \geq : Y_n = Z_n\}$$

It turns out that this supermartingale is of the form:

$$Z_n = \operatorname{ess\,sup}_{\tau \in S_n} \mathbb{E}(Y_\tau \mid \mathcal{F}_n) \tag{1}$$

Definition 11 (essential supremum existence). *For every family F of random variables, there exists a unique (a.e.) random variable $g : \Omega \to \mathbb{R} \cup \{\pm\infty\}$ such that:*

- *$g \geq f$ for all $f \in F$.*
- *If $h : \Omega \to \mathbb{R} \cup \{\pm\infty\}$ is another random variable with property (i), then $h \geq g$.*

We denote $g = ess\sup(F)$

Lemma 2. *For every adapted sequence $\{Y_n\}$ of integrable random variables satisfying $\mathbb{E}(\sup_{n\in\mathbb{N}_0} Y_n^+) < \infty$ the random variables $\{Z_n\}$ as defined in (1) form an adapted integrable sequence satisfying:*

$$Z_n = \max\Big\{Y_n, \mathbb{E}(Z_{n+1} \mid \mathcal{F}_n)\Big\}$$

$$\mathbb{E}(Z_n) = \sup_{\tau\in S_n} \mathbb{E}(Y_\tau)$$

*Indeed, Z_n is the smallest nonnegative supermartingale that dominates Y_n. We call it the **Snell Envelope** of Y_n.*

4.2 Martingale Inequalities

Theorem 15 (Doob's Submartingale Inequality). *For a submartingale $\{X_n\}$ we have:*

$$\mathbb{P}\Big(\max_{0\leq n\leq N} X_n \geq \lambda\Big) \leq \frac{\mathbb{E}(X_N^+)}{\lambda} \qquad \forall\lambda > 0, N \in \mathbb{N}$$

Theorem 16 (Kolmogorov's Inequality). *For independent $\{\eta_n\}$ with mean zero and finite variance, we have:*

$$\mathbb{P}\Big(\max_{1\leq n\leq N} \Big|\sum_{j=1}^{n} \eta_j\Big|\Big) \leq \frac{1}{\lambda^2} \sum_{j=1}^{n} \mathbb{E}(\eta_j^2)$$

Proof. $X_n = \sum_{j=1}^n \eta_j$ is a martingale and by Jensen's X_n^2 is a submartingale, so apply Doob's and use the cancellation from independence. □

Theorem 17 (Azuma-Hoeffding). *Let M_n be a martingale, with $M_0 = 0$ and $\mathbb{P}(|M_{n+1} - M_n| \leq r_n) = 1$ for some sequence $\{r_n\}$. Then, for some universal $C > 0$, we have*

$$\mathbb{P}(|X_n| > \lambda) \leq 2\exp\Big(-\frac{\lambda^2}{2\sum_{k=1}^n r_k^2}\Big) \tag{2}$$

$$\|X_n\|_p \leq C\sqrt{p\sum_{k=1}^{n} r_k^2} \tag{3}$$

4.3 Stochastic Approximation

Root-Finding Problem, with Noise: Suppose $h : \mathbb{R} \to \mathbb{R}$ is continuously differentiable. Not known globally but can be measured locally, and we know it has one root θ. Newton-Raphson method solves this problem under suitable conditions. But once we add noise to our measurements the premise falls apart.

Theorem 18. *Under suitable conditions, wherein a function $h : \mathbb{R} \to \mathbb{R}$ has a unique root $h(\theta) = 0$. Then, for any real-valued gains sequence $\{\gamma_n\}$ with*

$$\sum_n \gamma_n = \infty \qquad \sum_n \gamma_n^2 < \infty$$

the following stochastic approximation algorithm converges $\mathbb{P}$-a.e. to θ.

$$\theta_{n+1} = \theta_n - \gamma_{n+1} \cdot y_{n+1} = \theta_n - \gamma_{n+1}\Big[h(\theta_n) + \epsilon_{n+1}\Big]$$

Remark 3. *Robbins and Siegmund use the following to make the above argument work.*

Lemma 3 (Almost-Supermartingale Convergence)**.** *On a filtered probability space, consider non-negative adapted sequences $\{Z_n\}$ and $\{D_n\}$ which satisfy:*

$$\mathbb{E}(Z_{n+1}|\mathcal{F}_n) \le (1 + b_n)Z_n + c_n - D_n$$

for sequences of non-negative constants $\{b_n\}$ and $\{c_n\}$ with $\sum_{n\in\mathbb{N}}(b_n+c_n) < \infty$. Then almost everywhere we have:

$$\sum_n D_n < \infty \qquad \lim_{n\to\infty} Z_n \text{ exists in } \mathbb{R}$$

5 Week 5: Processes with Independent Increments

Definition 12 (Poisson Process)**.** *Let $\{\eta_n\}$ be a sequence of independent exponential random variables with parameter λ. The Poisson process is:*

$$N(t) = \max\{n : t \ge \sum_{i=1}^n \eta_i\}$$

Theorem 19 (Properties of Poisson Process)**.** *Let $(N_t)_{t\ge 0}$ be a Poisson process.*

- *(Poisson Distribution)* $\mathbb{P}(N(t) = n) = \exp(-\lambda t)\frac{(\lambda t)^n}{n!}$
- *(Independent increments)* $N(t + s) - N(t)$ *is independent of* $N(s)$.

Definition 13. *The Wiener Process is a stochastic process $W(t)$ such that:*

- $W(0) = 0$
- $t \to W(t)$ *are almost surely continuous.*
- *For any finite sequence* $0 = t_0 < t_1 < ... < t_n$ *and Borel Sets* $A_1, ..., A_n$

$$\mathbb{P}(W(t_1) \in A_1, ..., W(t_n) \in A_n) = \int_{A_1} \cdots \int_{A_n} \prod_{i=1}^{n} p(t_i - t_{i-1}, x_{i-1}, x_i)\ dx_1 \cdots dx_n$$

$$p(t, x, y) = \frac{1}{\sqrt{2\pi t}} \exp\Big(\frac{-(x-y)^2}{2t}\Big)$$

- *(Can replace 3 with: independent, stationary, and Gaussian increments)*
- *(Can replace 3 with:* W_t *and* $W_t^2 - t$ *are martingales, by Levy)*

Remark 4. $W(t)$ *is Gaussian distributed with variance* t*.*

Theorem 20. $\mathbb{E}[W_s W_t] = \min\{s, t\}$

Proof. Let $t \geq s$ and write $W_t = W_s + (W_t - W_s)$. Then:

$$\mathbb{E}[W_s W_t] = \mathbb{E}(W_s^2) + \mathbb{E}(W_t - W_s)\mathbb{W}(W_s) = s$$

□

Theorem 21 (Construction of Brownian Motion)**.** *The Haar functions form an orthonormal basis for the Hilbert space* $L^2([0,1])$*. They are defined as follows:*

$$h_{00}(t) = 1 \qquad h_{01}(t) = \mathbf{1}\{t < 1/2\} - \mathbf{1}\{t \geq 1/2\}$$

and for $i \in \mathbb{N}$ *and* $j = 1, 2, ..., 2^i$ *define:*

$$h_{ij}(t) = 2^{i/2}\mathbf{1}\Big\{t \in \big(\frac{2j-2}{2^{i+1}}, \frac{2j-1}{2^{i+1}}\big)\Big\} - 2^{i/2}\mathbf{1}\Big\{t \in \big(\frac{2j-1}{2^{i+1}}, \frac{2j}{2^{i+1}}\big)\Big\}$$

On a complete probability space construct $\{Z_{ij}\}$ *an infinite array of independent copies of the standard normal. Then the following is Brownian motion:*

$$W_t = Z_{00}\int_0^t h_{00}(s)ds + \sum_{i \in \mathbb{N}} \sum_{j=1}^{2^i} Z_{ij} \int_0^t h_{ij}(s)ds$$

It is easy to check that $\mathbb{E}(W_t^2) = t$*.*

$$\mathbb{E}(W_t^2) = (\langle \chi_{[0,t]}, h_{00}\rangle)^2 + \sum_{i \in \mathbb{N}_0} \sum_{j \in [2^i]} (\langle \chi_{[0,t]}, h_{ij}\rangle)^2 = \|\chi_{[0,t]}\|^2 = t$$

Here we recall Parseval's indentity, $\langle f, g\rangle = \sum_{i,j}\langle f, h_{ij}\rangle\langle g, h_{ij}\rangle$*.*
The hard part of this construction is showing continuity.

6 Week 6: Path Properties of Brownian Motion

Definition 14. *For $f : [0, T] \to \mathbb{R}$ the first variation is:*

$$V_1(f) = \limsup_{\|\pi\| \to 0} \sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)|$$

where $\pi = (t_0 = 0, t_1, ..., t_n = T)$ and $\|\pi\| = \max_i |t_{i+1} - t_i|$.

Lemma 4. *Let t_i^n partition $[0, T]$ into equal parts.*

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} [W(t_{i+1}^n) - W(t_i)]^2 = T$$

Proof. Show $\lim_{n\to\infty}[\sum_{i=0}^{n-1}[W(t_{i+1}^n) - W(t_i)]^2 - T]^2 = 0$. Expand and use independence of increments and the fact that $\mathbb{E}(W_t^4) = 3t^2$. □

Theorem 22. *For almost everywhere ω, $f(t) = W(t, \omega)$ has infinite first variation.*

Remark 5. *This makes it such that $\int_0^T f(t)dW(t)$ not well-defined by ordinary Riemann-Stieltjes integration, since the paths have infinite variation.*

Theorem 23. *With probability 1, $W(t)$ is non-differentiable for all $t \geq 0$.*

Remark 6. *For W_t Brownian motion, $\exp(W_t - t/2)$ is a martingale.*

Proof. $W_t - W_s$ is a normal random variable with mean 0 and variance $t - s$. Hence:

$$\mathbb{E}(\exp(W_t - W_s)) = \int_{-\infty}^{\infty} e^x \underbrace{p(t-s, 0, x)}_{\frac{1}{\sqrt{2\pi(t-s)}}\exp(\frac{-x^2}{2(t-s)})} dx = e^{(t-s)/2} \underbrace{\int_{-\infty}^{\infty} p(t-s, 0, x-t)dx}_{=1}$$

□

Definition 15 (Progressively Measurable)**.** *$\{X_t\}$ is progressively measurable with respect to a filtration $\{F_t\}$ if for all $t \geq 0$ and $A \in \mathcal{F}$*

$$\{(s, \omega) : s \in [0, t], \omega \in \Omega, X_s(\omega) \in A\} \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t$$

Definition 16 (Strong Markov Process)**.** *A progressively measurable $\{X_t\}$ with filtration $\{F_t\}$ on a space $(\Omega, \mathcal{F})$ is a strong Markov Process with initial distribution μ if:*

- $\mathbb{P}^\pi(X_0 \in A) = \mu(A)$ *for all $A \in \mathcal{F}$.*
- *For any stopping time S on $\{F_t\}$, $t \geq 0$, and $A \in \mathcal{F}$*

$$\mathbb{P}^\pi(X_{S+t} \in A | X_s) = \mathbb{P}^\pi(X_{S+t} \in A | F_s^+) \text{ a.e. on } \{S < \infty\}$$

Theorem 24. *Brownian Motion is a martingale and a strong Markov process.*

7 Week 7: Foundations of Ito Integration

We want to define $I_t(X) = \int_0^t X_s \, dW_s$.

Definition 17 (Bracket, Continuous Doob Decomposition). *For every nonconstant square-integrable (local) martingale M with continuous sample paths, let $(t_k^{(n)})_{k\in 2^n}$ denote the dyadic partition of M's support. Then we have:*

$$\langle M \rangle = \lim_{n\to\infty} \sum_k \Big(M(t_{k+1}^{(n)}) - M(t_k^{(n)}) \Big)^2$$

and $\langle M \rangle$ is the unique process with continuous and non-decreasing paths such that $M^2 - \langle M \rangle$ is a (local) martingale.

Definition 18. *A process X is **simple** if there exists a partition $0 = t_0 < t_1 < ... < t_r < t_{r+1} = T$ such that $X_s(\omega) = \theta_j(\omega)$ for $s \in (t_j, t_{j+1}]$ where θ_j is an $\mathcal{F}_{t_j}$-measurable r.v. Naturally, for a simple process X*

$$I_t(X) = \int_0^t X_s \, dW_s = \sum_{j=0}^r \theta_j (W_{t\wedge t_{j+1}} - W_{t\wedge t_j})$$

Remark 7. *This integral on simple functions is clearly a martingale with continuous sample paths and $\mathbb{E}(I_t(X)) = 0$. For simple processes X and Y, it's square integrable with:*

$$\langle I(X) \rangle_t = \int_0^t X_u^2 \, du \qquad \langle I(X), I(Y) \rangle_t = \int_0^t X_u \cdot Y_u \, du$$

Theorem 25 (Characterization of the Stochastic Integral). *Suppose some continuous local martingale H satisfies:*

$$\langle H, N \rangle_t = \int_0^t X_u \, d\langle M, N \rangle_u$$

for every continuous local martingale N. Then $H = I^M(X)$.

Theorem 26 (Ito Isometry). *For f a bounded simple process*

$$\mathbb{E}\Big[\Big(I^W(f)\Big)^2\Big] = \mathbb{E}\Big[\Big(\int_0^T f(t,\omega) \, dW_t(\omega)\Big)^2\Big] = \mathbb{E}\Big[\int_0^T f(t,\omega)^2 \, dt\Big]$$

Remark 8. *The isometry allows us to define the integral. We create a sequence of simple processes to approximate a given continuous process. In the end, the limit exists, because we can relate it the RHS and a Cauchy sequence in L2 which is famously a Hilbert space.*

8 Week 8: Basics of Ito Calculus

Definition 19 (local martingale). *A process $(M_t)_{t\geq 0}$ is a local martingale if there exists an non-decreasing sequence of stopping times $(\tau_n)_{n\geq 1}$ such that*

- $\{M_{t\wedge\tau_n}\}_{t\geq 0}$ *is martingale for all n.*
- $\mathbb{P}(\lim_{n\to\infty}\tau_n = \infty) = 1$.

Remark 9 (local implies super). *A local martingale bounded from below is a supermartingale. Let $M_t \geq 0$ be such a local martingale.*

$$\mathbb{E}(M_t|\mathcal{F}_s) = \mathbb{E}(\liminf_{n\to\infty} M_{\tau_n\wedge t}|\mathcal{F}_s) \leq \liminf_{n\to\infty}\mathbb{E}(M_{\tau_n\wedge t}|\mathcal{F}_s) = \liminf_{n\to\infty}\mathbb{E}(M_s) = \mathbb{E}(M_s)$$

To apply Fatou we needed bounded from below.

This observation leads us also to notice that any bounded local martingale is fully a martingale (establish submartingality using the upper bound and you are done).

Theorem 27 (General Ito's Rule). *For $f \in C^2(\mathbb{R})$ we have:*

$$f(M_t) = f(M_0) + \int_0^t f'(M_s)dM_s + \frac{1}{2}\int_0^t f''(M_s)d\langle M\rangle_s$$

Definition 20 (Notion of Solution). *A general first-order stochastic differential equation (SDE) can be written in the following form:*

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

A ***strong solution*** *to SDE on a given probability space $(\Sigma, \mathcal{F}, \mathbb{P})$ with respect to the fixed Brownian motion W and initial condition ξ if X is adapted to the filtration generated by the initial condition and the Brownian filtration, it starts at the initial condition $\mathbb{P}$-almost surely, the β and σ are $\mathbb{P}$-square integrable, and it indeed satisfies the equation on the space.*

A ***weak solution*** *is a triple $(X, W), (\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)$ where X satisfies the equation on this space according to that Brownian filtration.*

Definition 21 (Tanaka Equation). *Canonical example of a stochastic differential equation which has a weak solution but no strong solution.*

$$X_0 = 0 \qquad dX_t = sgn(X_t)dB_t$$

Definition 22 (Bessel Equation). *Note that $R(t) = \sqrt{\sum_{i=1}^n W_i^2(t)}$ for n independent Brownian motions satisfies the following:*

$$R(t) = r + B(t) + \frac{n-1}{2}\int_0^t \frac{ds}{R(s)}$$

Note that $\frac{1}{R(t)}$ is the classic example of a local martingale that is not a martingale.

Definition 23 (Ornstein-Uhlenbeck). *The following equation:*

$$dX_t = -\alpha X_t dt + \sigma dW_t$$

is satisfied by:

$$X_t = X_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s$$

Definition 24 (Brownian Bridge). *Note that $W_t = B_t - tB_1$ where B_t is Brownian motion uniquely satisfies the following:*

$$X_t = W_t - \int_0^t \frac{X_s}{1-s} ds$$

9 Week 9: Properties of Diffusion Processes

Definition 25 (Ito Diffusion). *An **Ito diffusion** is of the form:*

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

or equivalently, in integral form:

$$X_t = X_0 + \int_0^t b(t, X_s)ds + \int_0^t \sigma(t, X_s)dB_s$$

where B_t is Brownian motion, $b \in \mathbb{R}^n$, and $\sigma \in \mathbb{R}^{n\times m}$.
If b and σ do not depend on t we say the diffusion is time-homogenous.

There a number of interesting aspects of Ito diffusions, including the Markov Property, the strong Markov Property, the existence of an infinitesimal generator, the Dynkin formula, and the characteristic operator.

Definition 26 (Markov Property). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a bounded Borel measurable function. Let $\{\mathcal{F}_t^{(m)}\}_{t\geq 0}$ be the filtration generated by the Brownian motion. Then for X an Ito diffusion we have for any $t, h \geq 0$.*

$$\text{(Markov)} \qquad \mathbb{E}^x[f(X_{t+h})|\mathcal{F}_t^{(m)}](\omega) = \mathbb{E}^{X_t(\omega)}[f(X_h)]$$

Let τ be a stopping time. For $h \geq 0$ we have:

$$\text{(Strong Markov)} \qquad \mathbb{E}^x[f(X_{\tau+h})|\mathcal{F}_\tau^{(m)}](\omega) = \mathbb{E}^{X_\tau(\omega)}[f(X_h)]$$

Definition 27 (Generator of an Ito Diffusion)**.** *Let X_t be an Ito diffusion. The infinitesimal generator A of X_t is:*

$$Af(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t}$$

Let $D_A(x)$ denote the set of functions for which the above limit exists.

By applying Ito's formula and some linear algebra, we have, for a time-homogenous diffusion:

$$Af(x) = \sum_{i \in [n]} b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j \in [n]} (\sigma \sigma^T)_{ij}(x) \frac{\partial f}{\partial x_i \partial x_j}$$

Remark 10. *For X an n-dimensional Brownian motion, the generator is Laplacian:*

$$A_X f = \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

For X the graph of Brownian motion ($dX_1 = dt$ and $dX_2 = dB_t$), the generator is the so-called heat operator:

$$A_X f = \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \qquad f = f(t, x)$$

The following is sort of a fundamental theorem of calculus for diffusions.

Definition 28 (Dynkin's Formula)**.** *Let f be a C^2 function $\mathbb{R}^n \to \mathbb{R}$ with ocompact support. let τ be a stopping time $\mathbb{E}^x(\tau) < \infty$.*

$$\mathbb{E}^x[f(X_\tau)] = f(x) + \mathbb{E}^x\Big[\int_0^\tau Af(X_s)ds\Big]$$

Remark 11 (On Brownian Hitting Times)**.** *Using Dynkin's formula, we can establish some interesting facts about the behavior of Brownian motion, e.g. it is recurrent in two dimensions but transient in three.*

Consider n-dimensional Brownian motion starting at $a \in \mathbb{R}^n$ with $\|a\| < R$. What is the expected time it takes for Brownian motion to exist the R-radius ball about the origin? Let τ_k be the stopping time of interest. Applying Dynkin's formula with $f(x) = x^2$, we have:

$$R^2 = \mathbb{E}^a[f(B_{\tau_k})] = f(a) + \mathbb{E}^a[\int_0^{\tau_k} Af(B_s)ds] = \|a\|^2 + n\mathbb{E}^a[\tau_k]$$

$$\implies \mathbb{E}^a[\tau_k] = \frac{1}{n}(R^2 - \|a\|^2)$$

Now suppose we are in dimension ≥ 2 and we start at b with $\|b\| > R$. What is the probability that we enter the ball? Let α_k be the first time we either enter the inner circle or exit the circle of radius $2^k R$.

Fact: *$\Delta f = 0$ for*

$$f(x) = \begin{cases} -\log(\|x\|) & n = 2 \\ \|x\|^{2-n} & n > 2 \end{cases}$$

Hence, by Dynkin, $\mathbb{E}^b[f(B_{\alpha_k})] = f(b)$ for all $k \geq 1$.
But then, for $n = 2$ we have:

$$(-\log(R))p_k + (-\log(R2^k))q_k = -\log\|b\|$$

While for $n \geq 3$ we have:

$$\|R\|^{2-n}p_k + \|R2^k\|^{2-n}q_k = -\|b\|^{2-n}$$

where p_k is the probability of hitting the inner circle and q_k for outer.
By analyzing limits as $k \to \infty$ we can establish recurrence in 2d and transience in 3d.

Remark 12. *(Original Approach, Bessel Equation) Let $R(t) = \sum_{i=1}^n W_i^2(t)$ for W_i independent Brownian motions. Recall that for $f(x_1, ..., x_n) = \sqrt{\sum_{i=1}^n x_i^2}$ we have:*

$$\partial f/\partial x_i = \frac{x_i}{f(x)} \qquad \partial f/(\partial x_i \partial x_j) = \frac{\delta_{ij}}{f(x)} - \frac{x_i x_j}{f(x)^3}$$

By Ito's Rule, we have:

$$dR(t) = \sum_{i=1}^n \frac{W_i(t)}{R(t)} dW_i(t) + \frac{1}{2}\Big(\frac{n}{R(t)} + \underbrace{\frac{\sum_i W_i^2(t)}{R(t)^3}}_{1/R(t)}\Big)$$

$$dR_t = \frac{n-1}{2R_t} dt + d\beta_t$$

where $\beta_t = \sum_i \int_0^t \frac{W_i(\theta)}{R(\theta)} dW_i(\theta)$ is a BM, we can check this by taking quadratic variation.
By Ito's Lemma and setting $n = 2$ we have:

$$\log(R_t) = \log(R_0) + \underbrace{\int_0^t \frac{dR_s}{R_s}}_{\int_0^t \frac{dt}{4R_s^2} + \frac{d\beta_t}{R_s}} - \frac{1}{4}\int_0^t \frac{ds}{R_s^2}$$

$$\log(R_t) = \log(r) + \int_0^t \frac{dB_s}{R_s}$$

We can do similar to establish the following for $n = 3$:

$$\frac{1}{R(t)} = \frac{1}{r} + \int_0^t \frac{dB_s}{R_s^2}$$

Using the following fact about Bessel process hitting times (reminiscent of the hitting time equation for Brownian motion) with $0 < a < r < b < \infty$*,*

$$\mathbb{P}_r(T_a < T_b) = \frac{f(b) - f(r)}{f(b) - f(a)}$$

where f *is chosen such that* $f(R_t)$ *is an Ito integral with respect to BM.*

$$n = 2 \qquad \mathbb{P}_r(T_a < T_b) = \frac{\log(b/r)}{\log(b/a)} \to 1 \text{ as } b \to \infty$$

$$n = 3 \qquad \mathbb{P}_r(T_a < T_b) = \frac{1/b - 1/r}{1/b - 1/a} \to \frac{a}{r} \text{ as } b \to \infty$$

Note also that $k = \min_{t \geq 0} R_t$ *has uniform distribution on* $(0, r)$*. For* $l \in (0, r)$*,*

$$\mathbb{P}(\min_{t \geq 0} R_t < l) = \mathbb{P}(\exists t \text{ s.t. } R_t < l) = \mathbb{P}(\exists t \text{ s.t. } R_t = l) = l/r$$

10 Week 10: Stochastic Control

Definition 29 (Goal Problem of Heath-Kulldorff)**.** *Consider a process* X_t *with* $x \in (0, 1)$ *of the form*

$$X_t = x_0 + \int_0^t \pi_s dW_s + b \int_0^t \pi_s ds$$

with the property that $\mathbb{P}(X_t \in [0, 1] \ \forall t \in [0, T]) = 1$*. With this we can show that the endpoints are absorbing (Cameron-Martin). Interested in* $G(x_0) = \sup_{\pi \in \Pi(x_0)} \mathbb{P}(X^{x_0, \pi}(T) = 1)$*, where* $\Pi(x_0)$ *is the set of controls which keep the process in the interval and* $\mathbb{P}(\int_0^T \pi_t^2 dt < \infty) = 1$*.*
The trick in this problem is to rewrite with $\tilde{W}_t = W_t + bt$ *as*

$$X_s = x_0 + \int_0^1 \pi_s d\tilde{W}_s$$

Then under $\mathbb{Q}$ *with* $d\mathbb{Q}/d\mathbb{P} = \exp(-bW_t - \frac{b^2}{2} t)$*.*
Analyze $A_{x_0} = \{d\mathbb{Q}/d\mathbb{P} \geq k_{x_0}\}$*. If it has* $\mathbb{Q}(A_{x_0}) = x_0$ *then*

$$\mathbb{P}(A_{x_0}) = \sup_{B : \mathbb{Q}(A_{x_0}) \leq k_{x_0}} \mathbb{P}(B) \geq G(x_0)$$

with equality if we can find $\hat{\pi}$ *such that* $\mathbb{P}(X^{x_0, \pi^*}(T) = 1) = G^*(x_0)$*. Use Brownian properties to get the following explicit form for* G^**.*

$$G^*(x_0) = \Phi(\Phi^{-1}(x_0) + b\sqrt{T})$$

To find $\hat{\pi}$, observe the following $\mathbb{Q}$-Levy-martingale.

$$F(t, B(t); x_0) = \hat{X}(t) = \mathbb{Q}(A_{x_0} | \mathcal{F}_t)$$

Notice that F solves the backwards heat equation $\partial F + \frac{1}{2} D^2 F = 0$. More concretely:

$$\hat{X}(t) = x + \int_0^t \underbrace{DF(s, B(s); x_0)}_{\hat{\pi}_s} \underbrace{dB_s}_{dW_s + bds}$$

By uniqueness, solving for DF gives us the optimal control.

Theorem 28 (Neyman-Pearson Lemma)**.** *Fix a probability space and on it two measures $\mathbb{P} \ll \mathbb{Q}$. Fix $x \in (0,1)$. Let $Z = \frac{d\mathbb{P}}{d\mathbb{Q}}$. Suppose there is $\kappa = k(x) > 0$ such that $A_x = \{Z \geq \kappa_x\}$ satisfies $\mathbb{Q}(A_x) = x$. Then:*

$$\sup\{\mathbb{P}(B) : \mathbb{Q}(B) \leq x\} = \mathbb{P}(A_x)$$

11 Week 11: Portfolio Theory

A basic model for an asset's price fluctuation over time is as follows:

$$X_t = x_0 \exp\Big(\int_0^t \underbrace{(\alpha_s - \frac{\sigma_s^2}{2})}_{\gamma_s} ds + \int_0^t \sigma_s dW_s \Big)$$

where γ_s is the (local) rate of growth, σ_s is the local dispersion, and α_s is the (local) mean rate of return. The local rate of growth emerges here from Ito's lemma: this is perhaps more evident when we write the model in its simpler differential form.

$$\frac{dX_t}{X_t} = \alpha_t dt + \sigma_t dW_t$$

12 Week 12: Representation Theorems

12.1 Stochastic Exponential and Logarithm

Definition 30. *Let M be a continuous local martingale.*

The stochastic exponential $\epsilon(M)$ is given by

$$\epsilon(M)_t = \exp\Big(M_t - \frac{1}{2}\langle M \rangle_t \Big)$$

The stochastic logartihm $\mathcal{L}(M)$ is given by

$$\mathcal{L}(M)_t = \int_0^t \frac{dZ_s}{Z_s}$$

Theorem 29 (Novikov's Condition)**.** *Let W be a d-dimensional Brownian motion and let X be a d-dimensional process that satisfies:*

$$\mathbb{P}\Big[\int_0^T (X_t^{(i)})^2 < \infty\Big] = 1 \qquad \forall i \in [d], T \in [0, \infty)$$

If the following holds, then $\epsilon(X)$ is a martingale.

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \|X_s\|^2 ds\Big)\Big] < \infty \qquad \forall T \in [0, \infty)$$

Remark 13. *The solution to the stochastic differential equation*

$$Z_t = 1 + \int_0^t Z_s dM_s$$

is the stochastic exponential of M, i.e. $Z = \exp(M - \langle M \rangle)$,

Proof. Let $X = M - \frac{1}{2}\langle M \rangle$. Let $f(x) = e^x$. Let $Z = f(X)$. Apply Ito's.

$$\begin{aligned} dZ_t = d(f(X_t)) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X \rangle_t \\ &= Z_t(dM_t - \frac{1}{2}d\langle M \rangle_t) + \frac{1}{2}Z_t d\langle M \rangle_t = Z_t dM_t \end{aligned}$$

□

Remark 14 (Yor's Formula)**.** $\epsilon(L)\epsilon(M) = \epsilon(L + M + \langle L, M \rangle)$

Remark 15. $\epsilon(\mathcal{L}(M)) = \mathcal{L}(\epsilon(M)) = M$*.*

Theorem 30 (Van Schuppen-Wong)**.** *Consider a positive continuous martingale Z_t with $Z_0 = 1$. Fix $T \in [0, \infty)$. Define:*

$$\mathbb{Q}_T(A) = \mathbb{E}^{\mathbb{P}}[Z_T \cdot \boldsymbol{1}_A]$$

Then, for any process M that is a continuous local martingale under $\mathbb{P}$, the following process is a continuous local martingale under $\mathbb{Q}_T$ for all $t \in [0, T]$.

$$M'_t = M_t - \langle L, M \rangle_t = M_t - \int_0^1 \frac{d\langle M, Z \rangle_s}{Z_s}$$

Furthermore, $\langle M \rangle_t = \langle M' \rangle_t$ for $t \in [0, T]$.

Theorem 31 (Girsanov)**.** *Take* $L_T = \int_0^T \theta_t dW_t$*. Let* $Z_T = \epsilon(L_T) = \exp(L_T - \frac{1}{2}\langle L\rangle_T) = \frac{d\mathbb{Q}_T}{d\mathbb{P}}$*. Then if* W_t *is Brownian motion under* $\mathbb{P}$*, then the following is Brownian motion under* $\mathbb{Q}$*.*

$$W_t' = W_t - \int_0^t \theta_s ds$$

For θ_t *constant, and hence* $L_t = \theta W_t$*, this reduces to the Cameron-Martin Theorem, i.e.*

$$W_t' = W_t - \theta t \qquad Z_T = \exp\left(\theta W_T - \frac{\theta^2}{2}T\right)$$

Proof. (Using Van Schuppen-Wong) Start with L_T. Take $Z_T = \epsilon(L_T)$ and $M_T = W_T$ Brownian motion in the notation of Van Schuppen-Wong. Then the following is a continuous local martingale under the Z_T-changed measure.

$$W_t' = W_t - \int_0^1 \frac{\overbrace{Z_s\theta_s ds}^{d\langle Z_s, W_s\rangle_s}}{Z_s} = W_t - \int_0^t \theta_s ds$$

□

Furthermore, $\langle W'\rangle_t = \langle W\rangle_t = t$ so by Levy characterization it is Brownian motion.

12.2 DDS and others

Theorem 32 (Dambis-Dubins-Schwarz or DDS)**.** *Let* M *be a continuous local martingale. Then there exists a Brownian motion* B *such that:*

$$M(t) = B(\langle M\rangle(t))$$

Theorem 33 (Knight)**.** *Multivariate Extension of DDS. See Karatzas and Shreve p. 179.*

Theorem 34 (Doob representation theorem)**.** *Consider a continuous local martingale* M *starting at the origin, with* $\langle M\rangle_t = \int_0^t \lambda_s ds$ *with* $\lambda : [0,\infty] \times \Omega \to \mathbb{R}$ *progressively measurable and locally integrable. Then there exists* W *such that:*

$$M_t = \int_0^t \sqrt{\lambda_s} dW_s$$

(This is converse to the idea that if $M_t = \int_0^t H_s dW_s$ *then* $\langle M\rangle_t = \int_0^t H_s^2 ds$*).*