MATH 23500 Final Review

Rohan Buluswar

December 2024

1 Introduction

These are my final review notes for MATH 23500, Markov Chains, Martingales, and Brownian Motion, at the University of Chicago with Prof. Ewain Gwynne.

2 Markov Chains

Definition 2.1 (**Stochastic processes**). *A stochastic process is a collection of random variables $(X_t)_t$, taking values in a state space S and indexed by times $t \in T$.*

Definition 2.2 (**Markov processes**). *A discrete-time stochastic process $(X_n)_{n\in\mathbb{N}}$ is called a Markov process if for all $n \in \mathbb{N}$ and $s_0, s_1, ..., s_{n-1}, s_n \in S$,*

$$\mathbb{P}(X_n = s_n | X_0 = s_0, ..., X_{n-1} = s_{n-1}) = \mathbb{P}(X_n = s_n | X_{n-1} = s_{n-1})$$

It is moreover called time-homogeneous if for all $n \in \mathbb{N}$ and $x, y \in S$,

$$\mathbb{P}(X_n = y | X_{n-1} = x) = \mathbb{P}(X_1 = y | X_0 = x)$$

We can therefore fully describe a Markov chain by an initial distribution and transition probabilities $p(x, y) = \mathbb{P}(X_1 = y|X_0 = x)$. We may further define $p^n(x, y) = \mathbb{P}(X_n = y|X_0 = x)$.

Proposition 2.1. *If $(X_n)_{n\in\mathbb{N}} \in S$ is a Markov chain, then for any $n, m \in \mathbb{N}$ and $x, y \in S$, we have*

$$p^{n+m}(x, y) = \sum_{z\in S} p^n(x, z)p^m(z, y)$$

Proposition 2.2. *We can define the transition matrix by $P[i, j] = p(i, j)$. Then for any $n \in \mathbb{N}$, $P^n[i, j] = p^n(i, j)$.*

3 Recurrence and Transience for Finite S

Definition 3.1 (**Communication classes**). *If $(X_n)_{n\in\mathbb{N}}$ takes values in a finite state space S then we may define a relation by $x \sim y$ if there exist $m, n \in \mathbb{N}$ such that $p^m(x, y) > 0$ and $p^n(x, y) > 0$. This is in fact an equivalence relation, and allows us to partition S into blocs called communication classes.*

Definition 3.2 (**Recurrence, Irreducibility**). *A communication class C is called recurrent if $p(x, y) = 0$ for any $x \in C$ and $y \notin C$. Otherwise, it is called transient. If a Markov chain has only one communication class, then it is called irreducible.*

Proposition 3.1. *Suppose S is finite, C is a recurrent communication class, and $X_0 \in C$. Then with probability 1, (X_n) visits each $x \in C$ infinitely many times. On the other hand, if C is transient, then with probability 1, (X_n) eventually leaves C and never returns.*

4 Strong Markov Property

Definition 4.1 (**Stopping time**). *A random variable $\tau \in \mathbb{N}_0 \cup \{\infty\}$ is called a stopping time if for all $n \in \mathbb{N}_0$, the event $\{\tau = n\}$ is fully determined by $X_0, X_1, ..., X_n$.*

Theorem 4.1 (**Strong Markov property**). *Let τ be a stopping time. Let $n \geq 0$, $m \geq 1$, and let $x_0, x_1, ..., x_n \in S$ such that $\mathbb{P}(X_0 = x_0, ..., X_\tau = x_n) > 0$. Then for any $y_1, ..., y_m \in S$, we have*

$$\mathbb{P}(X_{\tau+1} = y_1, ..., X_{\tau+m=y_m} | X_0 = x_0, ..., X_\tau = x_n) = \mathbb{P}(X_1 = y_1, ..., X_m = y_m | X_0 = x_n)$$

Proposition 4.1. *Suppose $X_0 = x \in S$ and with probability 1, (X_n) hits x infinitely many times. If $\tau_0, \tau_1, ...$ are successive times such that $X_n = x$, then the increments $(X_{\tau_k}, ..., X_{\tau_{k+1}})$ are i.i.d. for all $k \in \mathbb{N}_{\geq 0}$.*

Proposition 4.2. *Suppose that $X_0 = x$ and with probability 1, $X_n = x$ for infinitely many n. If $y \in S$ such that $p^m(x, y) > 0$ for some $m \in \mathbb{N}$, we can define M to be the number of times we hit x before hitting y. Then M has a geometric distribution with some probability p, i.e. $\mathbb{P}(M = m) = p(1-p)^{m-1}$ for $m \geq 1$.*

Periodicity

Definition 4.2 (**Period**). *Given a Markov chain $(X_n)_{n \in \mathbb{N}_0}$ on a countable state space S and $x \in S$, we define the period of x by*

$$gcd(J_x) := gcd(\{n \geq 1 : p^n(x, x) > 0\})$$

Proposition 4.3. *Let (X_n) be as above, and consider $x \in S$. If $d = gcd(J_x)$ then J_x contains kd for large enough k.*

Proposition 4.4. *If $x \sim y$ are in the same communication class, then they have the same period, $d(x) = d(y)$.*

Definition 4.3 (**Aperiodic**). *If all $x \in S$ have $d(x) = 1$, then we call (X_n) aperiodic.*

5 Stationary Distributions

Definition 5.1. *Let (X_n) be a discrete-time Markov chain on a finite state space S. We say that $\pi : S \to [0, 1]$ is a stationary distribution if*

$$\sum_{x \in S} \pi(x) = 1$$

$$\pi(y) = \sum_{x \in S} \pi(x)p(x,y), \forall y \in S$$

Equivalently, we have:

1. $\pi P = \pi$, *where* $S = \{1, ..., n\}$ *and* P *is the transition matrix*
2. *If* $X_0 \sim \pi$ *then* $X_n \sim \pi$ *for all* $n \in \mathbb{N}$

Theorem 5.1. *Let* (X_n) *be an irreducible and aperiodic Markov chain on a finite state space* S. *Then it has a unique stationary distribution* π *and for any* $x, y \in S$,

$$\lim_{n \to \infty} \mathbb{P}(X_n = y | X_0 = x) = \pi(y)$$

Proposition 5.1. *Let* π *be the stationary distribution for* (X_n) *and for* $x \in S$, *let* $T_x = min\{n \geq 1 : X_n = x\}$. *Then*

$$\pi(x) = \frac{1}{\mathbb{E}[T_x | X_0 = x]}$$

Recurrence for Countably Infinite S

Definition 5.2. *Let* (X_n) *be a Markov chain on a countable state space* S. *We call* (X_n) *irreducible if for all* $x, y \in S$, *there exists* $n \in N$ *such that* $p^n(x,y) > 0$.

Definition 5.3. *A state* $x \in S$ *is called recurrent if*

$$\mathbb{P}(\exists \text{ infinitely many } n \geq 1 : X_n = x | X_0 = x) = 1$$

and transient otherwise.

Proposition 5.2. *If* (X_n) *is irreducible, then either every state is recurrent or every state is transient.*

Proposition 5.3. *A state* $x \in S$ *is recurrent iff*

$$\sum_{n=0}^{\infty} p^n(x,x) = \infty$$

Proposition 5.4. *Let* (X_n) *be the biased random walk on* $\mathbb{Z}$: *the probability of taking a step forward is* p, *and the probability of taking a step backward is* $1-p$. *If* $N \geq 1$ *and* $x \in \{1, ..., N-1\}$, *then*

$$\mathbb{P}(X \text{ hits } N \text{ before } 0 | X_0 = x) = \frac{(\frac{1-p}{p})^x - 1}{(\frac{1-p}{p})^N - 1},\ p \neq \frac{1}{2}$$

$$\mathbb{P}(X \text{ hits } N \text{ before } 0 | X_0 = x) = \frac{x}{N},\ p = \frac{1}{2}$$

Taking the limit as $N \to \infty$, *we have for any* $x \geq 1$ *that*

$$\mathbb{P}(X \text{ hits } 0 | X_0 = x) = 1,\ p \leq \frac{1}{2}$$

$$\mathbb{P}(X \text{ hits } 0 | X_0 = x) = (\frac{1-p}{p})^x,\ p > \frac{1}{2}$$

Consequently, (X_n) *is recurrent iff* $p = \frac{1}{2}$, *and transient otherwise.*

Definition 5.4. *Suppose S is countable and (X_n) is irreducible and recurrent. We say that it is null-recurrent if for all $x, y \in S$,*

$$\lim_{n\to\infty} p^n(x,y) = 0$$

and positive-recurrent otherwise.

Proposition 5.5. *If (X_n) is irreducible, then the following are equivalent:*

1. *(X_n) is positive recurrent*
2. *(X_n) has a stationary distribution π*
3. *$\limsup_{n\to\infty} p^n(x,y) > 0$ for all $x, y \in S$*
4. *If $T_x = \min\{n \geq 1 : X_n = x\}$ then $\mathbb{E}(T_x|X_0 = x) < \infty$ for all $x \in S$*

Moreover, if (X_n) is aperiodic and positive recurrent then the stationary distribution π is unique and

$$\pi(x) = \frac{1}{\mathbb{E}(T_x|X_0 = x)}$$

for all $x \in S$.

Proposition 5.6. *If (X_n) is irreducible and $x \in S$ then*

1. *(X_n) is transient $\iff$ $\sum_{n=0}^{\infty} p^n(x,x) < \infty$*
2. *(X_n) is null-recurrent $\iff$ $\lim_{n\to\infty} p^n(x,x) = 0$ and $\sum_{n=0}^{\infty} p^n(x,x) = \infty$*
3. *(X_n) is positive-recurrent $\iff$ $\limsup_{n\to\infty} p^n(x,x) > 0$*

Theorem 5.2. *The simple random walk on $\mathbb{Z}^d$ is null-recurrent for $d = 1, 2$ and transient for $d \geq 3$.*

Proposition 5.7 (Stirling's formula).

$$n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$$

Proposition 5.8. *Let (X_n) be the biased random walk (with reflected boundary) on $\mathbb{N}_{\geq 0}$ with probability $p \leq \frac{1}{2}$. Then the stationary distribution is given by*

$$\pi(x) = (1 - \frac{p}{1-p})(\frac{p}{1-p})^x$$

Branching Processes

We can model population growth with a branching process as follows. First, let $(p_k)_{k\geq 0}$ with $p_k \geq 0$ and $\sum p_k = 1$. Then let Y_m be i.i.d random variables with $\mathbb{P}(Y_m = k) = p_k$. Then we define $X_0 = 1$ and $X_{n+1} = \sum_{m=1}^{X_n} Y_m$. Here, Y_m models the number of children of individual m, which dies after one generation.

Definition 5.5 (Extinction probability). *We define the extinction probability by*

$$\mathbb{P}(\exists n \geq 1 : X_n = 0 | X_0 = 1)$$

Remark: if $\mu = \mathbb{E}(Y_i)$ then $\mathbb{E}(X_n) = \mu^n \mathbb{E}(X_0)$.

Proposition 5.9. *If* $\mu < 1$ *then* $a = 1$.

Definition 5.6 (Generating function). *For any random variable* $Y \in \mathbb{N}_{\geq 0}$*, we define its generating function by*

$$\varphi(s) = \varphi_Y(s) = \mathbb{E}(s^Y) = \sum_{k=0}^{\infty} \mathbb{P}(Y = k) s^k$$

Remark: $\varphi'(1) = \mathbb{E}(Y)$ and if $Y_1, ..., Y_m$ are independent then

$$\varphi_{Y_1+Y_2+...+Y_m}(s) = \prod_{i=1}^{m} \varphi_{Y_j}(s)$$

Proposition 5.10. *If* (X_n) *is a branching process and* φ *is the generating function for* $Y_i \sim (p_k)$*, then*

$$\varphi_{X_n}(s) = \varphi^{(n)}(s)$$

Proposition 5.11. *Let* (X_n) *be a branching process with offspring distribution* (p_k)*, and let* φ *be its generating function. Then if* $0 < p_0 < 1$*, the extinction probability is the smallest positive solution to* $\varphi(s) = s$.

Proposition 5.12. *Let* (X_n) *be a branching process with offspring distribution* (p_k) *and mean* μ*. Then*

1. *If* $\mu > 1$ *then* $a < 1$
2. *If* $\mu \leq 1$ *and* $p_0 \neq 0$ *then* $a = 1$

Poisson Processes

Definition 5.7. *A random variable* $Y \sim Poiss(\lambda)$ *has the Poisson distribution with parameter* λ *if for all* $k \in \mathbb{N}_{\geq 0}$,

$$\mathbb{P}(Y = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Remark: $\mathbb{E}(Y) = \lambda$.

Proposition 5.13. *If* $Y_1 \sim Poiss(\lambda_1)$ *and* $Y_2 \sim Poiss(\lambda_2)$ *are independent then* $Y_1 + Y_2 \sim Poiss(\lambda_1 + \lambda_2)$.

Definition 5.8. *The Poisson process with rate* λ *is the continuous-time stochastic process* (X_t) *such that* $X_0 = 0$ *and for any times*

$$0 \leq s_1 \leq t_1 \leq s_2 \leq t_2 \leq ... \leq s_k \leq t_k$$

the random variables $X_{t_j} - X_{s_j}$ *are i.i.d for* $1 \leq j \leq k$ *and each has a Poisson distribution with mean* $\lambda(t_j - s_j)$.

Definition 5.9. *A random variable* $T \in [0, \infty)$ *has the exponential distribution with parameter* λ *if for each* $t \geq 0$, $\mathbb{P}(T \geq t) = e^{-\lambda t}$.

Remark: Such a distribution has mean $\frac{1}{\lambda}$.

Definition 5.10. *Let the arrival times of a Poisson process be denoted by* $\tau_j = \inf\{t \geq 0 : X_t = j\}$. *Then the increments* $\tau_j - \tau_{j-1}$ *are i.i.d and have the exponential distribution with parameter* λ.

Proposition 5.14. *If* $T \sim exp(\lambda)$ *then the distribution of* $T - t$ *conditioned on the event* $\{T \geq t\}$ *is also exponential*(λ).

Proposition 5.15. *Let* $T_1, ..., T_n$ *be independent exponential random variables with parameters* $\lambda_1, ..., \lambda_n$. *Then* $\min\{T_1, ..., T_n\} \sim \exp(\lambda_1 + ... + \lambda_n)$. *Moreover,*

$$\mathbb{P}(T_j = \min\{T_1, ..., T_n\}) = \frac{\lambda_j}{\lambda_1 + ... + \lambda_n}$$

Continuous-Time Markov Chains

Definition 5.11. *We define* (X_t), *a continuous-time Markov chain taking values in a state space* S, *as follows. It is defined by rates* $\alpha(x, y) \geq 0$ *for* $x \neq y \in S$. *Supposing we start in state* x, *then for all* y *with* $\alpha(x, y) \neq 0$, *we define independent random variables* $T_y \sim \exp(\alpha(x, y))$. *Then we set* $T = \min T_y$, *and at time* T *the Markov chain jumps from* x *to* y *such that* $T_y = T$. *Note that* $T \sim \exp(\alpha(x))$ *where* $\alpha(x) = \sum_{y \neq x} \alpha(x, y)$. *Moreover,* $\mathbb{P}(X_T = y) = \frac{\alpha(x,y)}{\alpha(x)}$. *From here, we repeat the process starting at* y.

Proposition 5.16 (Continuous-time Markov property). *Let* $t \geq 0$ *and* $x \in S$. *The conditional distribution of* $(X_{t+s})_{s \geq 0}$ *given* $X_t = x$ *is the same as the distribution of* $(X_s)_{s \geq 0}$ *given* $X_0 = x$.

Definition 5.12. *Given a continuous time Markov chain* (X_t), *we may define the associated discrete-time Markov chain* $(\tilde{X_n})$ *by* $\tilde{X_n} = X_{\tau_n}$, *where* τ_n *is the time of the n-th jump. Then the discrete-time Markov chain has transition probabilities* $p(x, y) = \frac{\alpha(x,y)}{\alpha(x)}$ *and* $p(x, x) = 0$.

Definition 5.13. *We define the time-t transition probabilities by* $p_t(x, y) = \mathbb{P}(X_t = y | X_0 = x)$. *Labeling* $S = \{1, ..., N\}$, *we define the transition matrix by* $P_t[i, j] = p_t(i, j)$ *for* $1 \leq i, j \leq N$.

Definition 5.14 (Infinitesimal generator). *We define the infinitesimal generator by* $A[i, j] = \alpha(i, j)$ *for* $i \neq j$ *and* $A[i, i] = -\alpha(i)$.

Proposition 5.17. *The transition matrices* P_t *solves* $P_0 = I$ *and* $\frac{d}{dt} P_t = P_t A$. *As a result,* $P_t = e^{tA}$. *If* $A = QDQ^{-1}$ *is diagonalizable then* $e^{tA} = Qe^{tD}Q^{-1}$.

Definition 5.15. *A continuous-time Markov chain is irreducible if* $p_t(x, y) > 0$ *for all* $x, y \in S$ *and* $t > 0$.

Proposition 5.18. *A stationary distribution for a continuous-time Markov chain is one such that* $\sum_{x \in S} \pi(x) = 1$ *and* $X_0 \sim \pi \implies X_t \sim \pi$ *for all* $t > 0$. *This is equivalent to* $\pi A = 0$.

Proposition 5.19. *If* (X_t) *is irreducible, then it admits a unique stationary distribution* π. *Moreover,* $\lim_{t \to \infty} p_t(x, y) = \pi(y)$ *for all* $x, y \in S$.

Summary of differences between discrete-time and continuous-time Markov chains:

1. In both cases, S is discrete, but t is indexed by $\mathbb{N}_{\geq 0}$ or $\mathbb{R}_{\geq 0}$

2. **Discrete time:** transition matrix P of probabilities $p(x, y)$
3. **Continuous time:** transition matrix A of rates $\alpha(x, y)$
4. **Discrete time:** $p^n(x, y) = P^n[x, y]$
5. **Continuous time:** $p_t(x, y) = e^{tA}[x, y]$
6. **Discrete time:** stationary distribution solves $\pi P = \pi$
7. **Continuous time:** stationary distribution solves $\pi A = 0$.

The midterm content ends here.

6 Conditional Expectation

For discrete random variables X and Y, the conditional expectation given $Y = y$ is defined as

$$\mathbb{E}[X|Y = y] = \frac{\sum_{x\in S} xf(x, y)}{\sum_{x\in S} f(x, y)}$$

The random variable $\mathbb{E}[X|Y]$ is thus defined as

$$\mathbb{E}[X|Y] = \frac{\sum_{x\in S} xf(x, Y)}{\sum_{x\in S} f(x, Y)}$$

For continuous random variables X and Y, the conditional expectation as a random variable is given by

$$\mathbb{E}[X|Y] = \frac{\int xf(x, Y)dx}{\int f(x, Y)}$$

where $f(x, y)$ is the joint density and

$$\mathbb{P}[X \in A, Y \in B] = \int_A \int_B f(x, y)\, dy\, dx$$

Definition 6.1 (**Conditional expectation**). *Let X and Y be random variables with X taking values in $\mathbb{R}$ and $\mathbb{E}[|X|] < \infty$. Then the conditional expectation $\mathbb{E}[X|Y]$ is the unique random variable with the following properties:*

1. *$\mathbb{E}[X|Y]$ is a function of Y*
2. *If $F(Y)$ is a function of Y taking values in $\mathbb{R}$ with $\mathbb{E}[|F(Y)|] < \infty$ then*

$$\mathbb{E}[XF(Y)] = \mathbb{E}[\mathbb{E}[X|Y]F(Y)]$$

Definition 6.2. *Given a sequence of random variables $Y_0, ..., Y_n$, we define $\mathcal{F}_n$ to be the information in $Y_0, ..., Y_n$, so that $\mathbb{E}[X|\mathcal{F}_n] = \mathbb{E}[X|(Y_0, ..., Y_n)]$.*

Proposition 6.1. *$\mathbb{E}[X|\mathcal{F}_n]$ is the unique random variable such that $\mathbb{E}[X|\mathcal{F}_n]$ is $\mathcal{F}_n$-measurable and if Z is $\mathcal{F}_n$-measurable then $\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_n]Z]$.*

Proposition 6.2. *If X_1 and X_2 are random variables in $\mathbb{R}$ and $a, b \in \mathbb{R}$ are nonrandom then $\mathbb{E}[aX_1 + bX_2|\mathcal{F}_n] = a\mathbb{E}[X_1|\mathcal{F}_n] + b\mathbb{E}[X_2|\mathcal{F}_n]$.*

Proposition 6.3. *If X is $\mathcal{F}_n$-measurable then $\mathbb{E}[X|\mathcal{F}_n] = X$.*

Proposition 6.4. *If X is independent from $\mathcal{F}_n$ then $\mathbb{E}[X|\mathcal{F}_n] = \mathbb{E}[X]$.*

Proposition 6.5. *If Z is $\mathcal{F}_n$-measurable then $\mathbb{E}[ZX|\mathcal{F}_n] = Z\mathbb{E}[X|\mathcal{F}_n]$.*

Proposition 6.6. *We have $\mathbb{E}[\mathbb{E}[X|\mathcal{F}_n]] = \mathbb{E}[X]$.*

Proposition 6.7. *If $m \leq n$ then $\mathbb{E}[\mathbb{E}[X|\mathcal{F}_n]|\mathcal{F}_m] = \mathbb{E}[X|\mathcal{F}_m]$.*

Proposition 6.8. *We have $|\mathbb{E}[X|\mathcal{F}_n]| \leq \mathbb{E}[|X|\,|\mathcal{F}_n]$.*

7 Martingales

Definition 7.1. *Let $\{M_n\}$ be a sequence of random variables and let $\{\mathcal{F}_n\}$ be a filtration. $\{M_n\}$ is called a martingale w.r.t $\{\mathcal{F}_n\}$ if it satisfies the following properties:*

1. *Each M_n is $\mathcal{F}_n$ measurable*
2. *$\mathbb{E}[|M_n|] < \infty$ for each n*
3. *For each $m \leq n$, $\mathbb{E}[M_n|\mathcal{F}_m] = M_m$*

Proposition 7.1. *If M_n is a martingale then $\mathbb{E}[M_n] = \mathbb{E}[M_0]$ for each $n \in \mathbb{N}$.*

Proposition 7.2. *To show that $\mathbb{E}[M_n|\mathcal{F}_m] = M_m$ for all $m \leq n$, it suffices to check that $\mathbb{E}[M_n|\mathcal{F}_{n-1}] = M_{n-1}$ for all $n \geq 1$.*

Example 7.1. Let $X_1, X_2, ...$ be i.i.d random variables with mean μ and variance σ^2. Moreover, let $S_n = \sum_{i=1}^{n} X_n$ and let $\mathcal{F}_n$ be the information in $(X_1, ..., X_n)$. Then $M_n = S_n - \mu n$ is a martingale w.r.t $\mathcal{F}_n$. Moreover, if $\mu = 0$ then $M_n = S_n^2 - \sigma^2 n$ is a martingale.

Example 7.2. Let $Y_0, Y_1, ...$ be random variables and let $\mathcal{F}_n$ be the information contained in $(Y_0, ..., Y_n)$. If X is a real-valued random variable satisfying $\mathbb{E}[|X|] < \infty$, then $M_n = \mathbb{E}[X|\mathcal{F}_n]$ is a martingale.

Theorem 7.1 (**Optional Stopping Theorem**)**.** *Let $\{M_n\}$ be a martingale w.r.t $\{\mathcal{F}_n\}$ and let τ be a stopping time. Assume that:*

1. $\mathbb{P}[\tau < \infty] = 1$
2. $\mathbb{E}[|M_\tau|] < \infty$
3. $\lim_{n\to\infty} \mathbb{E}[M_n 1_{\{\tau > n\}}] = 0$

Then $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.

Proposition 7.3. *To check the last condition in the optional stopping theorem, it suffices to show that one of the following holds:*

1. *There exists a non-random $C > 0$ such that $|M_n| \leq C$ for all n.*
2. *There exists $C > 0$ such that $\mathbb{E}[M_n^2] \leq C$ for all n.*

Proposition 7.4. *Let $\{M_n\}$ be a martingale w.r.t $\{\mathcal{F}_n\}$ and let τ be a stopping time. Then $\{M_{n\wedge\tau}\}$ (where $n \wedge \tau$ is the minimum of the two times) is also a martingale w.r.t $\{\mathcal{F}_n\}$.*

Proposition 7.5. *Let $\{X_n\}$ be a discrete-time Markov chain on a countable state space S, and let C be a finite transient communication class. Then there exists $c > 0$ such that for any $x \in C$ and $j \in \mathbb{N}$, $\mathbb{P}[X \, exits \; C \text{ before time } j | X_0 = x] \geq 1 - e^{-cj}$.*

Theorem 7.2 (Martingale Convergence Theorem). *Let $\{M_n\}$ be a martingale w.r.t $\{\mathcal{F}_n\}$. Assume that there exists $C > 0$ such that $\mathbb{E}[|M_n|] \leq C$ for all n. Then with probability 1, the limit*

$$M_\infty = \lim_{n\to\infty} M_n$$

exists.

Proposition 7.6. *To apply the martingale convergence theorem, it suffices to instead check one of the following:*

1. *There exists $K > 0$ such that $M_n \geq -K$ for all n. (Or M_n is bounded above.)*
2. *There exists $C > 0$ such that $\mathbb{E}[M_n^2] \leq C$ for all n.*

8 Brownian Motion

Definition 8.1 (Standard Brownian Motion). *Standard Brownian Motion is the continuous-time process $\{B_t\}_{t\geq 0}$ taking values in $\mathbb{R}$ with the following properties:*

1. *The map $t \mapsto B_t$ is continuous*
2. *For all $s_1 \leq t_1 \leq s_2 \leq t_2 \leq ... \leq s_K \leq t_K$, the increments $B_{t_j} - B_{s_j}$ for $j = 1, 2, ..., K$ are independent*
3. *For all $0 \leq s < t$, $B_t - B_s \sim N(0, t - s)$*

Proposition 8.1. *If $X \sim N(\mu, \sigma^2)$ then the density function is given by*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

Proposition 8.2. *Let B_t be a standard Brownian motion and let $c > 0$. Then $t \mapsto c^{-\frac{1}{2}} B_{ct}$ is also a standard Brownian motion.*

Proposition 8.3. *For each $t \geq 0$, it holds with probability 1 that B is not differentiable at t.*

Definition 8.2. *Let $(X_1^n, ..., X_K^n)$ be vectors of K real-valued random variables. Then we say*

$$(X_1^n, ..., X_K^n) \to (X_1, ..., X_K)$$

in distribution if for any real numbers $a_1 < b_1$,...,$a_K < b_K$ such that

$$\mathbb{P}[X_j \in \{a_j, b_j\}] = 0$$

for all $j = 1, ..., K$, we have that

$$\lim_{n\to\infty} \mathbb{P}[a_1 < X_1^n < b_1, ..., a_K < X_K^n < b_K] = \mathbb{P}[a_1 < X_1 < b_1, ..., a_K < X_K < b_K]$$

Proposition 8.4. *Convergence in distribution is equivalent to the following: for every bounded continuous function $f : \mathbb{R}^K \to \mathbb{R}$,*

$$\lim_{n\to\infty} \mathbb{E}[f(X_1^n, ..., X_K^n)] = \mathbb{E}[f(X_1, ..., X_K)]$$

Theorem 8.1. *Let $\{X_j\}$ be the unbiased random walk on $\mathbb{Z}$. Moreover, let $X_t := (t-j)X_{j+1} + (j+1-t)X_j$ for all $t \in [j, j+1)$. We can consider times $0 \le t_1 \le t_2 \le ... \le t_K$. Then the joint distribution $(n^{-\frac{1}{2}} X_{nt_1}, ..., n^{-\frac{1}{2}} X_{nt_K})$ converges to the joint distribution of $(B_{t_1}, ..., B_{t_k})$.*

We define the continuous-time filtration $\{\mathcal{F}_t\}$ as the information in $\{B_s : s \le t\}$.

Proposition 8.5. *For each $t \ge 0$, the process $\{B_{s+t} - B_t\}_{s\ge0}$ is a standard Brownian motion that is independent from $\mathcal{F}_t$.*

Proposition 8.6. *The standard Brownian motion B_t is a martingale w.r.t $\mathcal{F}_t$.*

Theorem 8.2. *Let $\{M_t\}$ be a martingale w.r.t $\{\mathcal{F}_t\}$ such that $t \mapsto M_t$ is continuous and let τ be a stopping time. Suppose that there exists $C > 0$ such that*

$$\mathbb{P}[\tau < \infty, |M_t| \le C \,\forall t \le \tau] = 1$$

Then $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$

Proposition 8.7. *Let $a, b > 0$ and let $\tau = \min\{t \ge 0 : B_t \in \{-a, b\}\}$. Then*

$$\mathbb{P}[B_\tau = b] = \frac{a}{a+b}$$

Proposition 8.8. *Brownian motion is recurrent, in the sense that for each $T > 0$,*

$$\mathbb{P}[\exists t > T : B_t = 0] = 1$$

Proposition 8.9. *Let τ be a stopping time. Then the process $\{B_{s+\tau} - B_\tau\}_{s\ge0}$ is a Brownian motion independent of $\mathcal{F}_\tau$.*

Proposition 8.10 (Reflection principle). *Let $\{B_t\}$ be a standard Brownian motion and let $a > 0$. Then*

$$\mathbb{P}[\max_{0\le s\le t} B_s \ge a] = 2\mathbb{P}[B_t \ge a]$$

Definition 8.3. *The density function of the normal distribution on* $\mathbb{R}^d$ *with mean zero and covariance matrix* $\sigma^2 I$ *is given by*

$$f(x) = \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} \exp(-\frac{||x||^2}{2\sigma^2})$$

Proposition 8.11. *Let* $\overline{X} \sim N(0, \sigma^2 I)$ *be a normal random variable in* $\mathbb{R}^d$ *and let* Q *be a* $d \times d$ *orthonormal matrix. Then* $Q\overline{X} \sim \overline{X}$*.*

Definition 8.4. *The standard Brownian motion on* $\mathbb{R}^d$ *is the process* $\{(B_t^1, ..., B_t^d)\}_{t \geq 0}$ *where the* B_t^i *are independent standard Brownian motions in* $\mathbb{R}$*.*

Proposition 8.12. *The standard Brownian motion* $\overline{B}$ *on* $\mathbb{R}^d$ *is the unique process with the following properties:*

1. *The function* $t \mapsto \overline{B_t}$ *is continuous*
2. $\overline{B_t}$ *has independent increments*
3. *For each* $0 \leq s < t$*,* $\overline{B_t} - \overline{B_s} \sim N(0, (t-s)I)$

We have that $\overline{B}$ satisfies Brownian scaling, the Markov property, and the strong Markov property.

Proposition 8.13. *Let* $\overline{B_t}$ *be a standard Brownian motion on* $\mathbb{R}^d$ *and let* Q *be a* $d \times d$ *orthonormal matrix. Then* $Q\overline{B_t}$ *is also a standard Brownian motion.*