AI Institute for Advances in Optimization **Georgia Institute of Technology**

Lecture 11: Martingale Theory

Lecturer: ***Víctor H. de la Peña*** *Scribes: Heyuan Yao*

11.1 Discrete-time Martingale

11.1.1 Background and Definitions

Given a stochastic process $\{X_n\}_{n\in\mathbb{N}_0}$, as the time n increases, so does our knowledge about what has happened in the past, which can be modelled using filtration.

Definition 11.1 (Filtration (Discrete)) *Given the measurable space* $(\Omega, \mathcal{F})$*, a sequence of* σ*-algebras* $\mathcal{F}_1$*,* $\mathcal{F}_2$*, ..., on* Ω *such that*

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq ... \subseteq \mathcal{F}$$

is called a filtration.

Here $\mathcal{F}_n$ represents our knowledge at time n. It contains all events such that at time n it is possible to decide whether A has occurred or not. As n increases, there will be more such events A, i.e. the family $\mathcal{F}_n$ representing our knowledge will become larger.

Based on the definition of filtration, we can further define the notion of martingale.

Definition 11.2 (Discrete-time Martingales) *Given the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, a sequence* M_1*,* M_2*, ..., of random variables is called a martingale (resp, supermartingale, submartingale) with respect to a filtration* $\mathcal{F}_1$*,* $\mathcal{F}_2$*,..., if*

1) M_n *is integrable for each n = 1,2, ... ;*

2) M_1*,* M_2*, ... is adapted to* $\mathcal{F}_1$*,* $\mathcal{F}_2$*,...;*

3) $\mathbb{E}(M_{n+1}|\mathcal{F}_n) =$ *(resp,* $\leq$*,* $\geq$*)*M_n *a.s. for each n =1,2, ...*

The concept of a martingale has its origin in gambling, namely, it describes a fair game of chance. Similarly, the notions of submartingale and supermartingale defined below are related to favourable and unfavourable games of chance.

11.1.2 Examples

We now introduce some widely-used examples of martingales.

Mean-corrected Sums of i.i.d. r.v.s & Simple Symmetric Random Walk

Let $X_1, X_2, ...$ be independent random variables each with mean μ. Let $S_0 = 0$ and for $n > 0$ let S_n be the partial sum $S_n = X_1 + ... + X_n$. Then $M_n = S_n - n\mu$ is a martingale with respect to $\mathcal{F}_n$, the information contained in $X_1, ..., X_n$.

If we now assume that X_i is the Rademacher variable, then this stochastic process is called a **simple symmetric random walk**.

Harmonic Functions of a Markov Chain

Consider a discrete-time Markov chain $\{X_n\}_{n\in\mathbb{N}_0}$ with finite sample space χ and transition matrix $\mathcal{P}$. We say a bounded function $f : \chi \to \mathbb{R}$ is **harmonic** if $\mathcal{P}f = f$, i.e.,

$$\mathbb{E}(f(X_1)|X_0 = i) = \sum_{j\in\chi} f(j)\mathcal{P}_{ij} = f(i),$$

then $M_n^f := f(X_n)$ is a martingale w.r.t. $\mathcal{F}_n := \sigma(\{X_0, ..., X_n\})$.

Pólya Urn At time t = 0, an urn contains b black balls and w white balls. At each time t = n (n =1, 2, ...) a ball is chosen at random from the urn, and is replaced together with another ball of the same color. After time t=n, there are thus n+b+w balls in the urn, of which $b + B_n$ are black, where B_n is the number of black balls chosen by time n. Then $M_n := \frac{b+B_n}{b+w+n}$ the proportion of black balls in the urn right after t =n, is a positive martingale.

Remark: the limit of M_n, $M_\infty := \lim_{n\to\infty} M_n$ is $\mathcal{B}(b, w)$-distributed.

Martingale transform

Let $\{X_n\}_{n\in\mathbb{N}_0}$ be a martingale (w.r.t. to the filtration $\mathcal{F}_n$), and let $\{Y_n\}_{n\in\mathbb{N}_0}$ be a predictable sequence. The **martingale transform** $\{M_n\}_{n\in\mathbb{N}_0}$ is defined as follows:

$$M_n = (Y \cdot X)_n = X_0 + \sum_{t=0}^{n-1} Y_t(\Delta X)_{t+1},$$

where $(\Delta X)_{t+1} = X_{t+1} - X_t$ is the martingale difference. And $\{M_n\}_{n\in\mathbb{N}_0}$ is a martingale w.r.t. $\mathcal{F}_n$.

Wald's Martingale

Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. r.v.s with MGF $M(t) = \mathbb{E}(e^{tX_1})$ is finite for some $t > 0$, and set $S_n = X_1 + ... + X_n$ with $S_0 = 1$. The the process $\{W_n\}_{n\in\mathbb{N}}$ defined by

$$W_n = \frac{e^{tS_n}}{M^n(t)}$$

is a martingale known as **Wald's martingale**.

Likelihood-ratio Testing (Simple Hypotheses)

A random variable X is thought to be distributed according either to probability density f (null hypothesis)or to a different probability density g (alternative hypothesis). A random i.i.d. sample $X_1, ..., X_n$ is taken. Let L_n be the "likelihood ratio"

$$L_n = \frac{\mathcal{L}(g|x)}{\mathcal{L}(f|x)} = \prod_{i=1}^{n} \frac{g(X_i)}{f(X_i)},$$

and $\{L_n\}_{n\in\mathbb{N}}$ is a martingale under H_0.

11.2 Continuous-time Martingale

11.2.1 Notions

Let us now, extend the above notions to general cases, which contain continuous-time situations.

Definition 11.3 (Filtration) *Given the measurable space $(\Omega, \mathcal{F})$, a family of σ-algebras $\mathcal{F}_t$ on Ω, parametrized by $t \in T \subset \mathbb{R}$, is called a filtration if*

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$$

for any $s, t \in T$ such that $s \leq t$.

Definition 11.4 (Continuous-time Martingales) *Given the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a stochastic process ξ_t $(t \in T)$ is called a martingale (resp, supermartingale, submartingale) with respect to a filtration $\mathcal{F}_t$, if*

1) ξ_t is integrable for each $t \in T$;

2) ξ_t is adapted to $\mathcal{F}_t$ for each $t \in T$;

3) $\mathbb{E}(\xi_t|\mathcal{F}_s) = (resp,\ \leq,\ \geq) \xi_s$ a.s. for every $s, t \in T$ and $s \leq t$.

11.3 Examples

A few examples of continuous-time martingale will appear in our lectures later. And I just provide some frequently-used processes here.

Time-homogeneous Poisson Process, Subtracted by Its Mean

Let us consider a time-homogeneous Poisson process N(t), which may be the first continuous process you encountered in your probability study, with rate parameter λ. Then we directly find that the processes

$$M_t = N(t) - \lambda t$$

is a martingale, w.r.t. the filtration generated by $\{N(s)\}_{0 \leq s \leq t}$.

Brownian Motion

The **Wiener process** (or **Brownian motion**) is a stochastic process $W(t)$ with values in $\mathbb{R}$ defined for $t \in [0, \infty)$ such that

1) $W(0) = 0$ a.s.;

2) the sample paths $t \mapsto W(t)$ are a.s. continuous;

3) for any finite sequence of times $0 < t_1 < ... < t_n$ and Borel sets $A_1, ..., A_n \subset \mathbb{R}$

$$\mathbb{P}(W(t_1) \in A_1, ..., W(t_n) \in A_n) = \int_{A_1} ... \int_{A_n} p(t_1, 0, x_1) p(t_2 - t_1, x_1, x_2) ... p(t_n - t_{n-1}, x_{n-1}, x_n) dx_1 ... dx_n,$$

where

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

defined for any $x, y \in \mathbb{R}$ and $t > 0$, called the transition density.