MATH 526: Discrete State Stochastic Processes
Lecture 20

Chapter 5: Martingales

Prakash Chakraborty

University of Michigan

Random Walk: $(X_n)_{n=0}^{\infty}$ are i.i.d. r.v.'s, with mean μ, and $M_n = \sum_{i=0}^{n} X_i$. Notice that M_n is completely determined by $X_0, \ldots, X_n$.

Definition (Adaptness)

discrete time Stochastic Process

We say that the sequence of r.v.'s (M_n) is **adapted** to the sequence of r.v.'s (X_n) if

Model of randomness $$M_n = f_n(X_0, \ldots, X_n), \quad \text{for all } n,$$

where each f_n is a deterministic function.

- Any sequence is always adapted to itself.
- Assume (M_n) is adapted to (X_n). $f_n = X_n$
  - The r.v.'s $X_0, \ldots, X_n$ play the role of **information available at time** n: their values are observed at time n and they determine the value of M_n.
  - $X_0, \ldots, X_n$ contain **equal, or greater, amount of information** than $M_0, \ldots, M_n$.
    Mn是一个函数、减少信息量
  - $M_0, \ldots, M_n$ may contain **strictly less information** than $X_0, \ldots, X_n$: when it is impossible to recover X_n from $M_1, \ldots, M_n$. For example, $M_n = X_n^2$.
  - If X_n can be recovered from $M_0, \ldots, M_n$, then, $M_0, \ldots, M_n$ contain **the same amount of information** as $X_0, \ldots, X_n$. In the random walk example, we have: $X_n = M_n - M_{n-1}$.

Let us go back to the example of random walk. Because the random walk (M_n) is adapted to (X_n), we have:

$$\Rightarrow \quad \mathbb{E}(M_n \,|\, X_0, \ldots, X_n) = M_n.$$

And

$$\mathbb{E}(M_{n+1} \,|\, X_0, \ldots, X_n) = M_n + \mu.$$

Notation: for brevity, we will denote the family of variables on which we condition, $(X_0, \ldots, X_n)$, by $\mathcal{F}_n^X$.

Thus,

$$\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) = M_n + \mu.$$

If $\mu = 0$

$$\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) = M_n,$$

and (M_n) becomes a **martingale**.

Definition (Martingale)

A discrete time stochastic process $(M_n)_{n=0}^{\infty}$ is a **martingale** with respect to $(X_n)_{n=0}^{\infty}$, if

1. $(M_n)_{n=0}^{\infty}$ is adapted to $(X_n)_{n=0}^{\infty}$,
2. $\mathbb{E}|M_n| < \infty$, for all $n \geq 0$,
3. and $\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) = M_n$, for all $n \geq 0$.

The condition $\mathbb{E}|M_n| < \infty$ is needed to ensure that the conditional expectation is well defined.

Martingale property means that the conditional expectation of the next value of the process, given all currently available information, is simply the current value of the process. Statistically speaking, the *increments of a martingale process are pure noise*: **the best prediction for the next value of a martingale is simply its current value**.

Remark

If it is not specified with respect to what the process is a martingale, then, either the available information is clear from the context, or the specific structure of available information is not important, or the process is a martingale with respect to itself:

$$\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^M) = M_n.$$

Martingale should be understood as something that **doesn't change on average**. It is the stochastic analogue of a constant process. Analogously, one can introduce processes that increase or decrease on average.

non decreasing

Definition (Submartingale)

(M_n) is a **submartingale** with respect to (X_n), if

1. (M_n) is adapted to (X_n),
2. $\mathbb{E}|M_n| < \infty$, for all $n \geq 0$,
3. and $\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) \geq M_n$, for all $n \geq 0$.

R.W w/ positive mean,

non-increasing

Definition (Supermartingale)

(M_n) is a **supermartingale** with respect to (X_n), if

1. (M_n) is adapted to (X_n),
2. $\mathbb{E}|M_n| < \infty$, for all $n \geq 0$,
3. and $\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) \leq M_n$, for all $n \geq 0$.

In the random walk example, (M_n) is a submartingale if $\mu \geq 0$:

$$\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) = M_n + \mu \geq M_n.$$

Similarly, (M_n) is a supermartingale if $\mu \leq 0$:

$$\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n^X) = M_n + \mu \leq M_n.$$

Remark

Notice that martingale can be defined as a process which is a submartingale and a supermartingale at the same time.

We can, sometimes, modify (*compensate*) a sub- or supermartingale, to make it a martingale.

Example **(Asymmetric simple random walk).** Assume that $\{X_i\}$ are i.i.d. random variables, taking value 1 with probability p, and -1 with probability $q = 1 - p$.

$$S_n = \sum_{i=1}^{n} X_i.$$

And

$$\mathbb{E}(S_{n+1} \mid \mathcal{F}_n^X) = S_n + \underbrace{p - q}_{\mu}.$$

→ 非 mG.

Thus, (S_n) is a **submartingale** if $p \geq q$, and it is a **supermartingale** if $p \leq q$.

(1) Define $M_n := S_n - (p - q)n$. Then, $M_n \sim \{X_1 \cdots X_n\}$.

$$\begin{aligned}
\mathbb{E}(M_{n+1} \mid \mathcal{F}_n^X) &= \mathbb{E}(S_{n+1} \mid \mathcal{F}_n^X) - (p - q)(n + 1) \\
&= S_n + (p - q) - (p - q)(n + 1) \\
&= S_n - (p - q)n = M_n.
\end{aligned}$$

Hence, (M_n) is a **martingale**.

Martingales do not necessarily have to arise as sums of independent r.v.'s (random walks). In particular, there is another way to construct a martingale from (S_n):

(2)

$$M_n := (q/p)^{S_n}.$$

$\overset{1}{p} \times \overset{-1}{p-1}$

Let us check that (M_n) is a martingale with respect to (X_n):

- Adaptedness with respect to (X_n) is immediate.
-
$$\mathbb{E}|M_n| < \max\left\{ \left(\frac{q}{p}\right)^{X_0+n}, \left(\frac{p}{q}\right)^{X_0+n} \right\} < \infty.$$

-
$$\begin{aligned}\mathbb{E}(M_{n+1}\,|\,\mathcal{F}_n^X) &= M_n \cdot \mathbb{E}\left[\left(\frac{q}{p}\right)^{X_{n+1}}\right]\\ &= M_n \cdot \left(\frac{q}{p}\cdot p + \frac{p}{q}\, q\right)\\ &= M_n \cdot 1 \qquad c\cdot p + \frac{1}{c}\, q = 1\\ &= M_n.\end{aligned}$$

In the above, we used the following **property of conditional expectation**: if X is a deterministic function of $(Y_1, \ldots, Y_n)$, then

$$\mathbb{E}\left(XZ|\,Y_1, \ldots, Y_n\right) = X\mathbb{E}\left(Z|\,Y_1, \ldots, Y_n\right),$$

for any r.v. Z, for which the conditional expectation is well defined.

Basic properties of martingales.

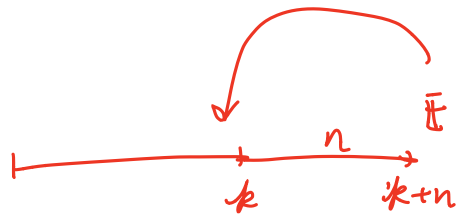
Assume that (M_n) is a **martingale** with respect to (X_n).

- $\mathbb{E}(M_{k+n} \,|\, \mathcal{F}_k^X) = M_k$, for all $n, k \geq 0$.

  Proof. $\mathbb{E}(M_{k+2} \,|\, \mathcal{F}_k^X)$

  $\overset{\text{Tower prop}}{=} \mathbb{E}\left(\mathbb{E}(M_{k+2} \,|\, \mathcal{F}_{k+1}^X) \,|\, \mathcal{F}_k^X \right).$

  $= \mathbb{E}(M_{k+1} \,|\, \mathcal{F}_k^X) = M_k$



- $\mathbb{E}M_n = \mathbb{E}M_0$, for all $n \geq 0$.

  Proof. $\mathbb{E}(\mathbb{E}(M_n \,|\, \mathcal{F}_0^X)) \overset{\text{Prop 1}}{=} \mathbb{E}(M_0)$

  $\mathbb{E}(M_n) =$

Assume (M_n) is a **submartingale** (**supermartingale**) with respect to (X_n):

- $\mathbb{E}[M_{k+n} \,|\, \mathcal{F}_k^X] \geq (\leq)\ M_k$, for all $n, k \geq 0$.
- $\mathbb{E}[M_n] \geq (\leq)\ \mathbb{E}M_0$, for all $n \geq 0$.

Functions of stochastic processes as martingales (submartingales, supermartingales).

Let us make a connection between martingales and Markov chains.

Question

How to determine **if a given function of a Markov chain is a martingale**?

Theorem

Let $(X_n)_{n=0}^{\infty}$ be a **Markov chain**, *with state space $\mathbb{S}$ and transition matrix p, and assume that function f satisfies:* $\mathbb{E}|M_n| < \infty$ $\quad$ $\mathbb{E}|M_n| = \Sigma$

- $\sum_{y \in \mathbb{S}} p^n(x, y)|f(y, n)| < \infty$, *for all $x \in \mathbb{S}$ and $n \geq 0$;* $\longrightarrow \mathbb{E}|M_n|$
- $\sum_y p(x, y) f(y, n+1) = f(x, n)$, *for all $x \in \mathbb{S}$ and $n \geq 0$.* $\rightarrow \mathbb{E}(M_{n+1} | \mathcal{F}_n^X)$.

Then, $(M_n = f(X_n, n))$ is a **martingale** *with respect to (X_n).*

Remark

If $f(x)$ is independent of n, then the second condition in the theorem becomes: $\sum_y p(x, y) f(y) = f(x)$. This equation is something we have seen before, when computing the exit probabilities. This is not a coincidence: we will see that, in fact, the **martingale theory allows for a much simpler computation of the exit probabilities**.

Proof. ☼

$$\mathbb{E}|M_n| = \sum_{y \in S} p^n(x,y) \cdot |f(y,n)| < \infty \quad \text{for all } x \in S$$

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n^X) = \sum p(x_n,y) \cdot f(y,n+1) = f(x,n)$$

Remark

In fact, we have shown that, for a Markov chain (X_n), with $\mathbb{E}|X_n| < \infty$, **the following useful (and intuitively obvious) identity holds**:

$$\mathbb{E}\Big[f(X_{n+1}, n+1) \,|\, \mathcal{F}_n^X \Big] = \sum_{y \in \mathbb{S}} f(y, n+1) p\,(X_n, y)\,.$$

Example (Branching process). Denote by X_n the number of members in the n-th generation of a population. Each member gives birth to a random number of children. These random numbers $\{Y_i\}$ are independent across members and across generations, and they are identically distributed with mean μ.

(X_n) is a MC with the state space $\mathbb{S} = \{0, 1, 2, \ldots\}$ and a transition matrix p.

Notice that $\mathbb{E}(X_{n+1} \mid X_n = x) = \sum_{y \in \mathbb{S}} p(x, y) y = \mathbb{E} \sum_{i=1}^{x} Y_i = \boxed{\mu x}$.

a ftn of MC that is MG

Introduce $Z_n = f(X_n, n) = X_n / \mu^n$. We show that (Z_n) is a martingale:

① Z_n adapted to $\mathcal{F}_n^X$. ✓

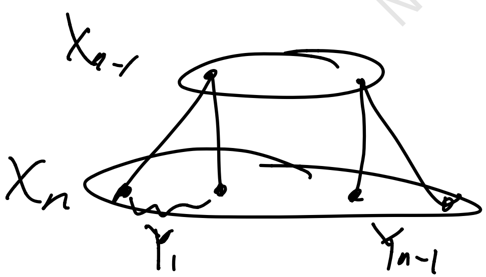
② $\mathbb{E}(|Z_n|) = \mathbb{E}(X_n/\mu^n) = \frac{1}{\mu^n} \mathbb{E}(X_n)$

$= \frac{1}{\mu^n} \mathbb{E}(\mathbb{E}(X_n \mid X_{n-1}))$

$= \frac{1}{\mu^n} \cdot \mu \cdot \mathbb{E}(X_{n-1})$.

$= \frac{\mathbb{E}(X_{n-1})}{\mu^{n-1}} = \mathbb{E}(|Z_{n-1}|)$

$= \mathbb{E}(|Z_0|)$



② $\mathbb{E}|Z_n| = \sum_{y=0}^{\infty} p^n(x,y)\,|f(y,n)|$

$= \sum_{y=0}^{\infty} p^n(x,y) \frac{|y|}{\mu^n}$

$= \frac{1}{\mu^n} \sum_{y=0}^{\infty} \sum_{z=0}^{\infty} p^{n-1}(x,z)\, p(z,y)\, |y|$

$\mathbb{E}(X_1 \mid X_0 = z).$

$= \frac{1}{\mu^n} \sum_{z=0}^{\infty} p^{n-1}(x,z) \boxed{\sum_{y=0}^{\infty} p(z,y)\,|y|} = \mu \cdot |z|$

$= \sum_{z=0}^{\infty} p^{n-1}(x,z) \frac{|z|}{\mu^{n-1}} = \cdots < \infty$

③ $\mathbb{E}(Z_{n+1} \mid \mathcal{F}_n^X) = \frac{1}{\mu^{n+1}} \cdot \mu X_n = \frac{X_n}{\mu^n} = Z_n$

Theorem

- *If (M_n) is a* **martingale** *and f is a* **convex** *function, such that $\mathbb{E}|f(M_n)| < \infty$, then, $(f(M_n))$ is a* **submartingale**. *In particular, if (M_n) is a* **martingale**, *and $\mathbb{E}M_n^2 < \infty$, then, (M_n^2) is a* **submartingale**.
- *If (M_n) is a* **submartingale** *and f is a* **nondecreasing convex** *function, such that $\mathbb{E}|f(M_n)| < \infty$, then, $(f(M_n))$ is a* **submartingale**.

Proof. Using Jensen's inequality, we obtain:

$$\mathbb{E}(f(M_{n+1}) \,|\, \mathcal{F}_n) \geq f\left(\underline{\mathbb{E}(M_{n+1} \,|\, \mathcal{F}_n)}\right) \geq f(M_n).$$

$\Downarrow$

Jansen

$p_1 f(x_1) + p_2 f(x_2) \geq f(p_1 x_1 + p_2 x_2)$

∎

mG + convex $\Rightarrow$ sub mG.

The increments of a martingale are orthogonal.

Theorem

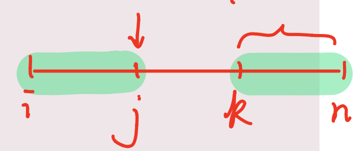
If (M_n) is a martingale, such that $\mathbb{E}M_n^2 < \infty$, for all $n \geq 0$, then, for any time indices $0 \leq i \leq j \leq k \leq n$, we have:

$\mathbb{E}XY \leq \sqrt{\mathbb{E}X^2 \cdot \mathbb{E}Y^2}$

$\Rightarrow \mathbb{E}XY$ exists.

$$\mathbb{E}((M_n - M_k)M_j) = 0,$$

$$\mathbb{E}((M_n - M_k)(M_j - M_i)) = 0,$$



provided the above expectations are well defined.

Proof.

$$\mathbb{E}((M_n - M_k)M_j) = \mathbb{E}\left(\mathbb{E}((M_n - M_k)M_j \,|\, \mathcal{F}_k)\right) = \mathbb{E}\left(\boxed{M_j}\,\mathbb{E}(M_n - M_k \,|\, \mathcal{F}_k)\right) = 0. \blacksquare$$

$(\mathbb{E}(M_n | \mathcal{F}_k) - M_k) = 0$

$\mathbb{E}(M_n - M_k)$

$= \mathbb{E}(\mathbb{E}(M_n - M_k | \mathcal{F}_k))$

$= \mathbb{E}(\mathbb{E}(M_n | \mathcal{F}_k) - M_k)$

We also have this useful result.

Theorem

If (M_n) is a martingale, such that $\mathbb{E}M_n^2 < \infty$, for all $n \geq 0$, then (Y_n) is a martingale, where

$M_{n+1}^2 = (M_{n+1} - M_n + M_n)^2 = (M_{n+1} - M_n)^2 + M_n^2 + 2(M_{n+1} - M_n) \cdot M_n$

↓ 0

$$Y_n = M_n^2 - \sum_{k=1}^{n} \mathbb{E}((M_k - M_{k-1})^2 \,|\, \mathcal{F}_{k-1})$$

convex Jensen.

The process (Y_n) is called the **compensated square of a martingale**.

$S_n^2 - \sum_{k=1}^{n} \mathbb{E}((S_k - S_{k-1})^2 \,|\, \mathcal{F}_{k-1}) = S_n^2 - \sum_{k=1}^{n} \mathbb{E}(X_k^2) = S_n^2 - \sigma^2 \cdot n$

Example. In particular, for a sequence of i.i.d. r.v.'s (X_n), with $\mathbb{E}X_n = 0$ and $\mathbb{E}X_n^2 = \sigma^2$, we have: $S_n = \sum_{i=1}^n X_i$ and $S_n^2 - \sigma^2 n$ are both martingales.

Suggested Exercises (Durrett, 3rd ed.), not for submission

Exercises 5.1, 5.2.

5.1. *Brother–sister mating.* Consider the six state chain defined in Exercise 1.66. Show that the total number of A's is a martingale and use this to compute the probability of getting absorbed into the 2,2 (i.e., all A's state) starting from each initial state.

5.2. *Lognormal stock prices.* Consider the special case of Example 5.4 in which $X_i = e^{\eta_i}$ where $\eta_i = \text{normal}(\mu, \sigma^2)$. For what values of μ and σ is $M_n = M_0 \cdot X_1 \cdots X_n$ a martingale?

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n^X) = \mathbb{E}(M_0 \cdots X_{n+1} \mid \mathcal{F}_n)$$

$$= M_0 \cdot X_1 \cdots X_n \cdot \mathbb{E}(X_{n+1})$$

$$= \mu \cdot M_n \Rightarrow \mu = 1$$

$$E|M_n| < \infty$$

$$\underbrace{M_0 \cdot E(X_i)^n}$$