MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Lecture 3 — Feb 20, SP17

Lecturer: Guy Bresler — Scribe notes by Linus Hamilton

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They are posted to serve class purposes.*

Martingales

Content.

1. Martingales
2. Martingale Transform
3. Doob's Upcrossing Inequality
4. Martingale Convergence Theorem

1 Martingales

A martingale is a sequence of random variables $M_0, M_1, \dots$ whose expected value "tends to stay the same." That is, if we know that the value of M_{n-1} is m_{n-1}, then $\mathbb{E}[M_n | M_{n-1} = m_{n-1}] = m_{n-1}$. They are often used to describe a process evolving over time, e.g. a random walk. Here is a formal definition:

Definition. Let $X_1, X_2, \dots$ be a sequence of independent random variables. A ***martingale*** is a sequence of random variables $M_0, M_1, \dots$ with finite mean satisfying:

(1) Each M_n is a function of $X_1, \dots, X_n$ (in particular M_0 is constant), and

(2) $\mathbb{E}[M_n | X_1, \dots, X_{n-1}] = M_{n-1}$. (In case you are unfamiliar with this notation, it is explained in the appendix.)

In general one may have M_0 with some starting distribution but you might as well just start your martingale at time index 1 in that case.

A ***sub*martingale** has the same definition as a martingale, except that the equality in (2) is replaced by $\mathbb{E}[M_n | X_1, \dots, X_{n-1}] \geq M_{n-1}$. Similarly a ***super*martingale** replaces (2) by $\mathbb{E}[M_n | X_1, \dots, X_{n-1}] \leq M_{n-1}$. Notice that the terms of a *sub*martingale tend to *increase*, while the terms of a *super*martingale tend to *decrease*. This is annoyingly backwards.

If you understand a martingale as "a sequence of random variables that, on average, stays the same," it gives intuition for the proofs in these scribe notes.

Fact. *If* (M_n) *is a martingale, then* $\mathbb{E}[M_0] = \mathbb{E}[M_1] = \cdots$.

Proof. For any n, we have $\mathbb{E}[M_n] = \mathbb{E}[\mathbb{E}[M_n|X_1, \ldots, X_{n-1}]] \mathbb{E}[M_{n-1}]$. The first step is by tower property of conditional expectation (see appendix). $\square$

Similarly, for supermartingales, $\mathbb{E}[M_0] \geq \mathbb{E}[M_1] \geq \mathbb{E}[M_2] \geq \cdots$. The proof is identical. And for submartingales, replace $\geq$ by $\leq$.

The following theorem is useful for analyzing martingales. We will prove it soon.

Theorem. *(Martingale convergence theorem) Let* (M_n) *be a submartingale with* $\sup_n \{\mathbb{E}[\max(M_n, 0)]\} < +\infty$. *Then almost surely,* (M_n) *converges, and* $\mathbb{E}[|\lim_{n\to\infty} M_n|] < \infty$.

2 Examples of martingales

1. A simple random walk on $\mathbb{Z}$. Here $M_0 = 0$, and $M_{n+1} = \begin{cases} M_n + 1 & \text{w.p. } 1/2 \\ M_n - 1 & \text{w.p. } 1/2 \end{cases}$.

2. Let M_0 be any constant, and define $M_n = X_n \cdot M_{n-1}$, where the X_i are independent and $\mathbb{E}[X_n] = 1$.

3. Let Y and $X_1, X_2, \ldots$ be random variables.[1] Then $M_n = \mathbb{E}[Y|X_1, X_2, \ldots]$ is a martingale. Why? Because the variables $X_1, X_2 \ldots$ slowly reveal information about Y. A recent real-world example should cement your intuition:

$$Y = \begin{cases} 1 & \text{Patriots win} \\ 0 & \text{Falcons win} \end{cases}, \quad X_i = \text{everything that happens in } i^{\text{th}} \text{ minute of game}$$

   This kind of martingale is called a *Doob martingale*.

4. Percolation on the infinite d-ary tree, where percolation occurs with probability p. Let X_n be the number of nodes on the n^{th} level of the tree that are percolated. Then $M_n = \frac{X_n}{(dp)^n}$ is a martingale.

[1]Technical constraint: $\mathbb{E}[|Y| < \infty]$.

3 Martingale transforms

Setting: You are investing in a certain stock, and can choose how much to sell/buy each day.

A sequence (H_n) is *predictable* if each H_n is a function of $X_1, X_2, \ldots, X_{n-1}$. Imagine that you are investing in a stock, and after day $n-1$, set your investment amount to H_n. Then your total income after day n is

$$H_1(M_1 - M_0) + H_2(M_2 - M_1) + \cdots + H_n(M_n - M_{n-1}).$$

We will denote the above quantity by $(H \cdot M)_n$. This is called a ***martingale transform***.

Fact. *If (M_n) is a martingale, then so is $(H \cdot M)_n$.*

Proof. We have

$$\begin{aligned}
\mathbb{E}\left[(H \cdot M)_n | X_1, \ldots, X_{n-1}\right] &= \mathbb{E}\left[(H \cdot M)_{n-1} + H_n(M_n - M_{n-1}) | X_1, \ldots, X_{n-1}\right] \\
&= (H \cdot M)_{n-1} + H_n \mathbb{E}\left[M_n - M_{n-1} | X_1, \ldots, X_{n-1}\right] \\
&= (H \cdot M)_{n-1}
\end{aligned}$$

as desired. □

A similar proof shows that if (M_n) is a (super/sub)martingale and $H_n \geq$, then $(H \cdot M)$ is a (super/sub)martingale.

As a consequence, we have the following "theorem":

Let (M_n) be a supermartingale. Then in finite time, in expectation, you cannot make money by buying stock in (M_n), even if you can change your investment amount every step.

4 Doob's Upcrossing Inequality

In a variety of contexts, it is impossible to make money by betting on a (super)martingale. This is unfortunate for gamblers, but we can use it to generate theorems. Just propose a gambling strategy, write down the fact "it can't win on average," and then translate that into a theorem.

Let's do that with "buy-low-sell-high." Suppose the value of a stock follows a supermartingale (M_n). When the stock value gets less than a certain number a, the gambler *buys* (i.e. sets her investment to 1.) When the stock value gets more than another number b, the gambler *sells* (i.e. sets her investment to 0.)

How much money does the gambler make? Whenever the stock crosses the interval (a, b) upward, she makes at least $(b - a)$ money. Apart from these 'upcrossings,' you can check that the gambler earns

$$\max(0, a - M_0) - \max(0, a - M_n).$$

But since (M_n) is a supermartingale, the gambler can't make money on average. This implies the following result, known as Doob's Upcrossing Inequality.

Theorem. *Let (M_n) be a supermartingale and (a, b) be an interval. Let U_n be the number of times that $M_0, \ldots, M_n$ crosses over the interval (a, b) going upward. Then*

$$(b - a)\mathbb{E}[U_n] + \mathbb{E}[\max(0, a - M_0)] - \mathbb{E}[\max(0, a - M_n)] \leq 0.$$

4.1 Proof of the Martingale Convergence Theorem

The Upcrossing Inequality allows a beautiful proof of the Martingale Convergence Theorem. Recall the theorem:

Theorem. *(Martingale convergence theorem) Let (M_n) be a submartingale with $\sup_n \{\mathbb{E}[\max(M_n, 0)]\} < +\infty$. Then almost surely, (M_n) converges, and $\mathbb{E}[|\lim_{n\to\infty} M_n|] < \infty$.*

Proof. What does it mean for (M_n) not to converge? It means that $\limsup M_n \neq \liminf M_n$. In which case, for any a, b with $\liminf M_n < a < b < \limsup M_n$, the sequence (M_n) makes infinitely many upcrossings of the interval (a, b).

Doob's upcrossing inequality tells us that since $\sup_n \{\mathbb{E}[\max(M_n, 0)]\} < +\infty$, the expected number of upcrossings of any interval is finite. Whence any particular interval almost surely has finitely many upcrossings. Take the union over all intervals with rational endpoints: we deduce that almost surely, *every* such interval has finitely many upcrossings. Therefore, almost surely $\liminf M_n = \limsup M_n$, so almost surely (M_n) converges!

Let $M = \lim M_n$. Now we only need to prove the second part of the theorem, $\mathbb{E}[|M|] < \infty$. This follows from $\mathbb{E}[\max(M, 0)] < \infty$ and $\mathbb{E}[\max(-M, 0)] < \infty$, both of which follow from Fatou's lemma. $\square$

5 Stopping Times

Setting: A gambler bets \$1 on a coin toss, pledging to stop once he has either gained \$10 or lost \$100 overall.

Let $X_1, X_2, \ldots$ be a sequence of random variables. A *stopping time* N (not a constant!) is a random variable such that the event $[N \leq k]$ is a deterministic function of $X_1, X_2, \ldots, X_k$. That is to say, after seeing $X_1, X_2, \ldots, X_k$, we must deterministically decide to either stop and set $N = k$, or keep going on to X_{k+1}.

Fact. *Let (M_n) be a martingale and N a stopping time. Then $(M_{n \wedge N})$ is a martingale. (Here $n \wedge N$ means $\min(n, N)$.)*

Proof. Set
$$H_n = \begin{cases} 1 & \text{if we have not stopped at time } n \\ 0 & \text{if we have stopped at time } n \end{cases}.$$
Then $(M_{n \wedge N})$ is the martingale transformation $(H \cdot M)$. $\square$

Of course, the above proof works with "martingale" replaced by "supermartingale" or "submartingale."

It follows that in expectation, a gambler can't make money in finite time by cleverly deciding when to stop betting. This is a special case of the result in section 3, but it is worth singling out.

6 Appendix: the notation $\mathbb{E}[X|Y]$

The notation $\mathbb{E}[X|Y]$ is the expectation of X, conditioned on the random variable Y. It is a function of Y. If Y is discrete, it has the definition

$$y \mapsto \mathbb{E}[X|Y = y].$$

If Y is a continuous random variable, then the above definition doesn't work because $\Pr[Y = y] = 0$. This is fixable by instead conditioning on $[Y \in (y - \epsilon, y + \epsilon)]$ and taking $\epsilon \to 0$. If you know about σ-fields, then you can write $\mathbb{E}[X|Y] = \mathbb{E}[X|\mathcal{F}]$ where $\mathcal{F}$ is the σ-field $\sigma(Y)$.

Conditional expectation has some useful properties:

- $\mathbb{E}[f(Y)X|Y] = f(Y)\mathbb{E}[X|Y]$
- $\mathbb{E}[X] = \mathbb{E}_Y[\mathbb{E}[X|Y]]$ (sometimes called the Tower property)
- $\mathbb{E}[\phi(X)|Y] = \phi(\mathbb{E}[X|Y])$ if ϕ is convex. (This is analogous to Jensen's inequality.)