6.262: Discrete Stochastic Processes 5/9/11

L24: Martingales: stopping and converging

Outline:

- **Review of martingales**
- **Stopped martingales**
- **The Kolmogorov submartingale inequality**
- **SLLN for IID rv's with a variance**
- **The martingale convergence theorem**

A sequence $\{Z_n;\, n \geq 1\}$ of rv's is a martingale if

$$\mathsf{E}\left[Z_n \mid Z_{n-1}, Z_{n-2}, \dots, Z_1\right] = Z_{n-1}; \qquad \mathsf{E}\left[|Z_n|\right] < \infty \qquad (1)$$

for all $n \geq 1$.

Recall that $\mathsf{E}\left[Z_n \mid Z_{n-1}, \dots, Z_1\right]$ is a rv that maps each sample point ω to the conditional expectation of Z_n conditional on $Z_1(\omega), \dots, Z_{n-1}(\omega)$. For a martingale, this expectation must be the rv Z_{n-1}.

Lemma: For a martingale, $\{Z_n;\, n \geq 1\}$, and for $n > i \geq 1$,

$$\mathsf{E}\left[Z_n \mid Z_i, Z_{i-1} \dots, Z_1\right] = Z_i : \qquad \mathsf{E}\left[Z_n\right] = \mathsf{E}\left[Z_i\right]$$

Can you figure out why $\mathsf{E}\left[Z_n \mid Z_m, \dots Z_1\right] = Z_n$ for $m \geq n$?

Note that $\mathsf{E}\left[Z_n \mid Z_1\right] = Z_1$ and $\mathsf{E}\left[Z_n\right] = \mathsf{E}\left[Z_1\right]$ for all $n > 1$.

Simple Examples of martingales

1) Zero-mean RW: If $Z_n = \sum_{i=1}^{n} X_i$ **where** $\{X_i; i \geq 1\}$ **are IID and zero mean, then** $\{Z_n;\, n \geq 1\}$ **is a martingale.**

2) If $Z_n = \sum_{i=1}^{n} X_i$ **where** $\mathsf{E}[X_i \mid X_{i-1}, \dots, X_1] = 0$ **for each** $i \geq 1$**, then** $\{Z_n;\, n \geq 1\}$ **is a martingale.**

3) Let $X_i = U_i Y_i$ **where** $\{U_i;\, i \geq 1\}$ **are IID, equiprobable** ± 1**. The** Y_i **are independent of the** U_i**. Then** $\{Z_n;\, n \geq 1\}$**, where** $Z_n = X_1 + \cdots X_n$**, is a martingale.**

4) Product form martingales: Let $\{X_i;\, i \geq 1\}$ **be a sequence of IID unit-mean rv's. Then** $\{Z_n;\, n \geq 1\}$**, where** $Z_n = X_1 X_2 \cdots X_n$**, is a martingale.**

If $\mathsf{p}_X(0) = \mathsf{p}_X(2) = 1/2$**, then** $\mathsf{p}_{Z_n}(2^n) = 2^{-n}$ **and** $\mathsf{p}_{Z_n}(0) = 1 - 2^{-n}$**. Thus,** $\lim_{n\to\infty} Z_n = 0$ **WP1 and** $\lim_{n\to\infty} \mathsf{E}[Z_n] = 1$**.**

Sub- and supermartingales are sequences $\{Z_n; n \geq 1\}$ **with** $\mathsf{E}[|Z_n|] < \infty$ **for which inequalities replace the equalities of martingales. For all** $n \geq 1$**,**

$$\begin{aligned} \mathsf{E}[Z_n \mid Z_{n-1}, \dots, Z_1] &\geq Z_{n-1} && \textbf{submartingale} \\ \mathsf{E}[Z_n \mid Z_{n-1}, \dots, Z_1] &\leq Z_{n-1} && \textbf{supermartingale} \end{aligned}$$

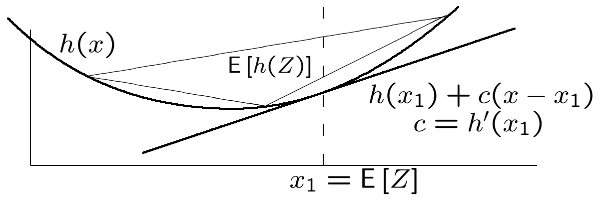
If $\{Z_n;\, n \geq 1\}$ **is a submartingale, then** $\{-Z_n;\, n \geq 1\}$ **is a supermartingale and vice-versa, so we consider only submartingales. For submartingales,**

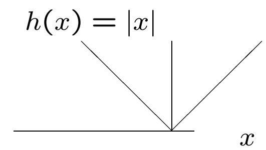
$$\mathsf{E}[Z_n \mid Z_i, \dots, Z_1] \geq Z_i \qquad \textbf{for all } n > i > 0$$

$$\mathsf{E}[Z_n] \geq \mathsf{E}[Z_i] \qquad \textbf{for all } n > i > 0$$

Convex functions

A function $h(x)$, $\mathbb{R} \to \mathbb{R}$, is convex if each tangent to the curve lies on or below the curve. The condition $h''(x) \geq 0$ is sufficient but not necessary.





Lemma (Jensen's inequality): If h is convex and Z is a rv with finite expectation, then

$$h(\mathsf{E}[Z]) \leq \mathsf{E}[h(Z)]$$

Jensen's inequality leads to the following theorem. See proof in text.

Thm 7.8.1: If $\{Z_n;\, n \geq 1\}$ is a martingale or submartingale, if h is convex, and if $\mathsf{E}[|h(Z_n)|] < \infty$ for all n, then $\{h(Z_n);\, n \geq 1\}$ is a submartingale.

For example, if $\{Z_n;\, n \geq 1\}$ is a martingale, then $\{|Z_n|;\, n \geq 1\}$, $\{Z_n^2;\, n \geq 1\}$ and $\{e^{rZ_n};\, n \geq 1$ are submartingales if the marginal expected values exist.

Stopped martingales

The definition of a stopping trial J for a stochastic process $\{Z_n; n \geq 1\}$ applies to any process. That is, J must be a rv and $\{J = n\}$ must be specified by $\{Z_1, \ldots, Z_n\}$.

A possibly defective rv J is a mapping from Ω to the extended reals $\mathbb{R}^+$ where $\{J = \infty\}$ and $\{J = -\infty\}$ might have positive probability. The other provisos of rv's still hold.

A possibly defective stopping trial is thus a stopping rule in which stopping may never happen (such as RW's with a single threshold).

A stopped process $\{Z_n^*; n \geq 1\}$ for a possibly defective stopping time J on a process $\{Z_n; n \geq 1\}$ satisfies $Z_n^* = Z_n$ if $n \leq J$ and $Z_n^* = Z_J$ if $n > J$.

For example, a given gambling strategy, where Z_n is the net worth at time n, could be modified to stop when Z_n reaches some given value. Then Z_n^* would remain at that value forever after, while Z_n follows the original strategy.

Theorem: If J is a possibly defective stopping rule for a martingale (submartingale), $\{Z_n; n \geq 1\}$, then the stopped process $\{Z_n^*; n \geq 1\}$ is a martingale (submartingale).

Pf: Obvious??? The intuition here is that before stopping occurs, $Z_n^* = Z_n$, so Z_n^* satisfies the martingale (subm.) condition. Afterwards, Z_n^* is constant, so it again satisfies the martingale (subm) condition.

Proof that $\{Z_n^*;\, n \geq 1\}$ is a martingale: Note that

$$Z_n^* = \sum_{m=1}^{n-1} Z_m \mathbb{I}_{\{J=m\}} + Z_n \mathbb{I}_{\{J \geq n\}}$$

Thus $|Z_n^*| \leq \sum_{m<n} |Z_m| + |Z_n|$. Thus means that $\mathsf{E}[|Z_n^*|] < \infty$ since it is bounded by the sum of n finite numbers.

Next, let $\vec{Z}^{(n-1)}$ denote $Z_{n-1}, \ldots, Z_1$ and consider

$$\mathsf{E}\left[Z_n^* \mid \vec{Z}^{(n-1)}\right] = \sum_{m<n} \mathsf{E}\left[Z_m \mathbb{I}_{\{J=m\}} \mid \vec{Z}^{(n-1)}\right] + \mathsf{E}\left[Z_n \mathbb{I}_{\{J \geq n\}} \mid \vec{Z}^{(n-1)}\right]$$

$$\mathsf{E}\left[Z_m \mathbb{I}_{\{J=m\}} \mid \vec{Z}^{(n-1)} = \vec{z}^{(n-1)}\right] = \begin{cases} z_m; & \textbf{if } J = m \\ 0; & \textbf{if } J \neq m. \end{cases} \qquad \textbf{for } m < n$$

$$\begin{aligned}
\mathsf{E}\left[Z_m \mathbb{I}_{\{J=m\}} \mid \vec{Z}^{(n-1)}\right] &= Z_m \mathbb{I}_{J=m}. \\
\mathsf{E}\left[Z_n \mathbb{I}_{\{J=n\}} \mid \vec{Z}^{(n-1)}\right] &= Z_{n-1} \mathbb{I}_{\{J \geq n\}} \\
\mathsf{E}\left[Z_n^* \mid \vec{Z}^{(n-1)}\right] &= \sum_{m<n} Z_m \mathbb{I}_{\{J=m\}} + Z_{n-1} \mathbb{I}_{\{J \geq n\}} \\
&= \sum_{m<n-1} Z_m \mathbb{I}_{\{J=m\}} + Z_{n-1}[\mathbb{I}_{\{J=n-1\}} + \mathbb{I}_{\{J \geq n\}}] \\
&= Z_{n-1}^*
\end{aligned}$$

This shows that $\mathsf{E}\left[Z_n^* \mid \vec{Z}^{(n-1)}\right] = Z_{n-1}^*$. To show that $\{Z_n^*;\, n \geq 1\}$ is a martingale, though, we must show that $\mathsf{E}\left[Z_n^* \mid \vec{Z}^{*(n-1)}\right] = Z_{n-1}^*$. However, $\vec{Z}^{*(n-1)}$ is a function of $\vec{Z}^{(n-1)}$.

For every sample point $\vec{z}^{(n-1)}$ of $\vec{Z}^{(n-1)}$ leading to a given $\vec{z}^{*(n-1)}$ of $\vec{Z}^{*(n-1)}$, we have

$$\mathsf{E}\left[Z_n^* \mid \vec{Z}^{(n-1)} = \vec{z}^{(n-1)}\right] = z_{n-1}^*$$

and thus

$$\mathsf{E}\left[Z_n^* \mid \vec{Z}^{*(n-1)} = \vec{z}^{*(n-1)}\right] = z_{n-1}^*.$$

QED

A consequence of the theorem, under the same assumptions, is that

$$\begin{array}{lllll} \mathsf{E}[Z_1] & \leq & \mathsf{E}[Z_n^*] & \leq & \mathsf{E}[Z_n] \qquad \textbf{(submartingale)} \\ \mathsf{E}[Z_1] & = & \mathsf{E}[Z_n^*] & = & \mathsf{E}[Z_n] \qquad \textbf{(martingale)} \end{array}$$

This is also almost intuitively obvious and proved in Section 7.8.

Recall the generating function product martingale for a random walk. That is, let $\{X_n;\, n \geq 1\}$ be IID and $\{S_n; n \geq 1\}$ be a random walk where $S_n = X_1 + \cdots + X_n$.

Then for r such that $\gamma(r)$ exists, let $Z_n = \exp[rS_n - n\gamma(r)]$. Then $\{Z_n;\, n \geq 1\}$ is a martingale and $\mathsf{E}[Z_n] = 1$ for all $n \geq 1$.

For r such that $\gamma(r)$ exists, let $Z_n = \exp[rS_n - n\gamma(r)]$. Then $\{Z_n;\, n \geq 1\}$ is a martingale and $\mathsf{E}[Z_n] = 1$ for all $n \geq 1$.

Let J be the nondefective stopping time that stops on crossing either $\alpha > 0$ or $\beta < 0$. Then $\mathsf{E}[Z_n^*] = 1$ for all $n \geq 1$.

Also, $\lim_{n\to\infty} Z_n^* = Z_J$ WP1 and

$$\mathsf{E}[Z_J] = \mathsf{E}[\exp[rS_J - J\gamma(r)] = 1$$

This is Wald's identity in a more general form. The connection of $\lim_n Z_n^*$ to Z_J needs more care (see Section 7.8), but this shows the power of martingales.

Kolmgorov's submartingale inequality

Thm: Let $\{Z_n; n \geq 1\}$ be a non-negative submartingale. Then for any positive integer m and any $a > 0$,

$$\Pr\left\{\max_{1\leq i\leq m} Z_i \geq a\right\} \leq \frac{\mathsf{E}[Z_m]}{a}. \qquad (2)$$

If we replace the max with Z_m, this is the lowly but useful Markov inequality.

Proof: Let J be the stopping time defined as the smallest $n \leq m$ such that $Z_n \geq a$.

If $Z_n \geq a$ for some $n \leq m$, then J is the smallest n for which $Z_n \geq a$.

If $Z_n < a$ for all $n \leq m$, then $J = m$. Thus the process must stop by time m, and $Z_J \geq a$ iff $Z_n \geq a$ for some $n \leq m$. Thus

$$\Pr\left\{\max_{1\leq n\leq m} Z_n \geq a\right\} = \Pr\{Z_J \geq a\} \leq \frac{\mathsf{E}[Z_J]}{a}.$$

Since the process must be stopped by time m, we have $Z_J = Z_m^*$.

$\mathsf{E}[Z_m^*] \leq \mathsf{E}[Z_m]$, so the right hand side above is less than or equal to $\mathsf{E}[Z_m]/a$, completing the proof.

The Kolmogorove submartingale inequality is a strengthening of the Markov inequality. The Chebyshev inequality is strengthened in the same way.

Let $\{Z_n;\, n \geq 1\}$ be a martingale with $\mathsf{E}\left[Z_n^2\right] < \infty$ for all $n \geq 1$. Then

$$\Pr\left\{\max_{1\leq n\leq m} |Z_n| \geq b\right\} \leq \frac{\mathsf{E}\left[Z_m^2\right]}{b^2}; \quad \text{for all integer } m \geq 2, \text{all } b > 0.$$

Let $\{S_n;\, n \geq 1\}$ be a RW with $S_n = X_1 + \cdots + X_n$ where each X_i has mean $\overline{X}$ and variance σ^2. Then for any positive integer m and any $\epsilon > 0$,

$$\Pr\left\{\max_{1\leq n\leq m} |S_n - n\overline{X}| \geq m\epsilon\right\} \leq \frac{\sigma^2}{m\epsilon^2}.$$

SLLN for IID rv's with a variance

Thm: Let $\{X_i;\, i \geq 1\}$ be a sequence of IID random variables with mean $\overline{X}$ and standard deviation $\sigma < \infty$. Let $S_n = X_1 + \cdots + X_n$. Then for any $\epsilon > 0$,

$$\Pr\left\{\lim_{n\to\infty} \frac{S_n}{n} = \overline{X}\right\} = 1$$

Idea of proof:

$$\Pr\left\{\bigcup_{m=k}^{\infty}\left\{\max_{1\leq n\leq 2^m} |S_n - n\overline{X}| \geq 2^m\epsilon\right\}\right\} \leq \sum_{m=k}^{\infty} \frac{\sigma^2}{2^m\epsilon^2} = \frac{2\sigma^2}{2^k\epsilon^2}$$

Then lower bound the left term to

$$\Pr\left\{\bigcup_{m=k}^{\infty}\left\{\max_{2^{m-1}\leq n\leq 2^m} |S_n - n\overline{X}| \geq 2n\epsilon\right\}\right\}$$

The martingale convergence theorem

Thm: Let $\{Z_n;\, n \geq 1\}$ be a martingale and assume that there is some finite M such that $\mathsf{E}\left[|Z_n|\right] \leq M$ for all n. Then there is a random variable Z such that, $\lim_{n\to\infty} Z_n = Z$ WP1.

The text proves the theorem with the additional constraint that $\mathsf{E}\left[Z_n^2\right]$ is bounded. Either bounded $\mathsf{E}\left[Z_n^2\right]$ or bounded $\mathsf{E}\left[|Z_n|\right]$ is a very strong constraint, but the theorem is still very powerful.

For a branching process $\{X_n;\, n \geq 1\}$ where the number Y of offspring of an element has $\overline{Y} > 1$, we saw that $\{X_n/\overline{Y}^n; n \geq 1\}$ is a martingale satisfying the constraint, so $X^n/\overline{Y}^n \to Z$ WP1.