Martingales & Azuma-Hoeffding Inequality

Yung Yi

Martingales

Definition 12.1: *A sequence of random variables $Z_0, Z_1, \ldots$ is a* martingale *with respect to the sequence $X_0, X_1, \ldots$ if, for all $n \geq 0$, the following conditions hold:*

- *Z_n is a function of $X_0, X_1, \ldots, X_n$;*
- $\mathbf{E}[|Z_n|] < \infty$;
- $\mathbf{E}[Z_{n+1} \mid X_0, \ldots, X_n] = Z_n$.

A sequence of random variables $Z_0, Z_1, \ldots$ is called martingale when it is a martingale with respect to itself. That is, $\mathbf{E}[|Z_n|] < \infty$, and $\mathbf{E}[Z_{n+1} \mid Z_0, \ldots, Z_n] = Z_n$.

- Conditional expected value of next observation, given all past observations, is equal to the last observation

- Submartingale (>=), Supermartingale (<=)

Example: Sequential fair games, martingale

- X_i be the amount the gambler wins on the ith game
- Fair game: $E[X_i] = 0$

- Z_i be the gambler's total winnings at the end of the ith game
- $Z_i = X_0 + X_1 + X_2 + \cdots + X_i$

- $E[Z_{i+1}|X_1, X_2, \dots, X_i] = E[X_{i+1} + Z_i|X_1, X_2, \dots, X_i] =$
  $E[X_{i+1}|X_1, X_2, \dots, X_i] + E[Z_i|X_1, X_2, \dots, X_i]$
  $= E[X_{i+1}] + Z_i = Z_i$

- Thus, Z_i is martingale w.r.t. the sequence (X_i)

Example: Sequential fair games, submartingale

- X_i be the amount the gambler wins on the ith game
- Fair game: $E[X_i] = 0$

- Z_i be the gambler's total winnings at the end of the ith game
- $Z_i = X_0 + X_1 + X_2 + \cdots + X_i, T_i = Zi^2$
- Show that T_i is a submartingale

Doob's inequality

Proposition 5.3 (Doob's inequality) *Let* $(X(t))_{t\geq 0}$ *be a càdlàg (i.e. right-continuous with left limits) submartingale. Then for every* $x \geq 0$ *and* $T \geq 0$, *one has*

$$\mathbf{P}\Big(\sup_{0\leq t\leq T} |X(t)| \geq x\Big) \leq \frac{1}{x}\mathbf{E}(|X(T)|)\,. \tag{5.1}$$

- Analogy to Markov inequality for a single random variable
- Used for proving Proposition 5.2, but used in many context
- Submartingale은 "커지는 경향"이 있다는 것을 intuition으로.

Doob martingale

- A special martingale

- $X_0, X_1, \ldots, X_n$ be a sequence of RVs
- Y be a RV with $E[|Y|] < \infty$ (Y depends on $X_0, \ldots, X_n$)
- $Z_i = E[Y|X_0, \ldots, X_i], i = 0,1, \ldots, n$

- Z_i is indeed martingale since

Using the fact that
$E[V|W] = E[E[V|U, W]|W]$

$$
\begin{aligned}
\mathbf{E}[Z_{i+1} \mid X_0, \ldots, X_i] &= \mathbf{E}[\mathbf{E}[Y \mid X_0, \ldots, X_{i+1}] \mid X_0, \ldots, X_i] \\
&= \mathbf{E}[Y \mid X_0, \ldots, X_i] \\
&= Z_i .
\end{aligned}
$$

Doob martingale

- Concept of the Doob martingale
  - Expectation on a certain target r.v. Y given observation upon ith period
  - Predict the value of Y, with gradually revealing X_i's to collect information incrementally
  - Z_i represents refined estimates of Y

- In most applications,
  - The first element Z_0 is just $E[Y]$, where Y is independent of "trivial" X_0
  - Gradually knowing the exact value…
  - Finally, $Z_n = Y$ (deterministic)

Doob martingale

- Examples on random graph $G = G_{n,p}$
- X_i be edge indicator on particular edge index $i = 1, \dots, \binom{n}{2}$

- F be any function on graph G
- $Z_i = E[F(G)|X_1, \dots, X_i]$ be expected function value given observations
- This process is called edge exposure martingale

- *Vertex exposure martingale
  - $Z_i = E[F(G)|G_1, \dots, G_i]$ where G_i represents the subgraph of G induced by first i vertices observed

Tail Inequalities for Maringales

Azuma-Hoeffding inequality

- Chernoff-like tail bounds of martingales
- Even when the underlying random variables are not independent
  - NOTE: Chernoff bound for Poisson trials: independent, but not indentical

Azuma-Hoeffding inequality

- Chernoff-like tail bounds of martingales

- **Theorem 12.4 [Azuma-Hoeffding Inequality]:**
  Let $X_0, \ldots, X_n$ *be a martingale such that* $|X_k - X_{k-1}| \le c_k$
  Then, for all $t \ge 0$ *and any* $\lambda > 0$*,*

$$\Pr(|X_t - X_0|) \ge \lambda) \le 2e^{-\lambda^2/(2\sum_{k=1}^{t} c_k^2)}$$

Azuma-Hoeffding inequality

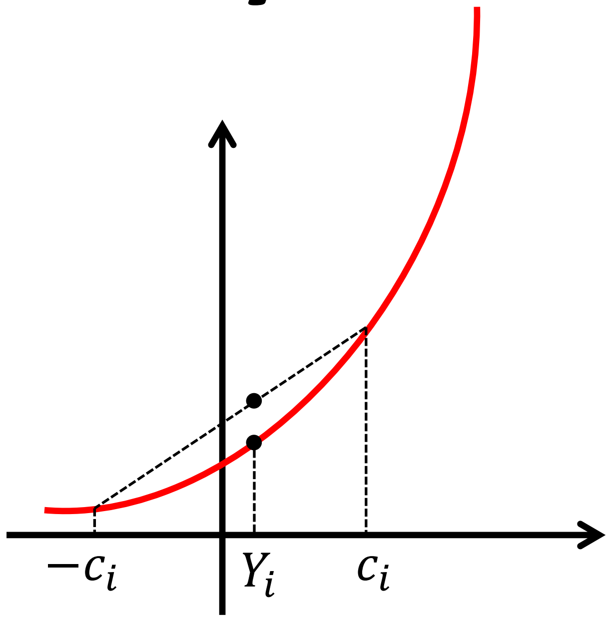
- Proof) Similarly to Chernoff bounds, first derive an upper bound for $E\left[e^{\alpha(X_t - X_0)}\right]$

  Define $Y_i \triangleq X_i - X_{i-1}, i = 1, \dots, t$, since $X_0, \dots$ is martingale,

$$
\begin{aligned}
E[Y_i | X_0, \dots, X_{i-1}] &= E[X_i - X_{i-1} | X_0, \dots, X_{i-1}] \\
&= E[X_i | X_0, \dots, X_{i-1}] - X_{i-1} = 0
\end{aligned}
$$

Azuma-Hoeffding inequality

- Now consider $E[e^{\alpha Y_i}|X_0, \dots, X_{i-1}]$ by writing $Y_i = -c_i \frac{1-Y_i/c_i}{2} + c_i \frac{1+Y_i/c_i}{2}$



- Using convexity of $e^{\alpha Y_i}$, we have

$$e^{\alpha Y_i} \leq \frac{1-Y_i/c_i}{2} e^{-\alpha c_i} + \frac{1+Y_i/c_i}{2} e^{\alpha c_i}$$

$$= \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}\left(e^{\alpha c_i} - e^{-\alpha c_i}\right)$$

- Therefore, we have

$$E[e^{\alpha Y_i}|X_0, \dots, X_{i-1}] \leq E\left[\frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}\left(e^{\alpha c_i} - e^{-\alpha c_i}\right)\middle| X_0, \dots, X_{i-1}\right]$$

$$= \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} \leq e^{\frac{(\alpha c_i)^2}{2}}$$

Derived by Taylor series expansion of e^x

Azuma-Hoeffding inequality

- It follows that

$$E\left[e^{\alpha(X_t - X_0)}\right] = E\left[\prod_{i=1}^{t} e^{\alpha Y_i}\right]$$

$$= E\left[\prod_{i=1}^{t-1} e^{\alpha Y_i}\right] E[e^{\alpha Y_t} | X_0, \dots, X_{t-1}] \le E\left[\prod_{i=1}^{t-1} e^{\alpha Y_i}\right] e^{\frac{(\alpha c_t)^2}{2}}$$

$$\le \cdots \le e^{\frac{\alpha^2 \sum_{i=1}^{t} c_i^2}{2}}$$

- Hence, $\Pr(X_t - X_0 \ge \lambda) = \Pr\left(e^{\alpha(X_t - X_0)} \ge e^{\alpha\lambda}\right) \le$

$$\frac{E\left[e^{\alpha(X_t - X_0)}\right]}{e^{\alpha\lambda}} \le e^{\alpha^2 \sum_{k=1}^{t} \frac{c_i^2}{2} - \alpha\lambda} \le e^{-\frac{\lambda^2}{2 \sum_{k=1}^{t} c_k^2}}$$

Markov inequality

Applications of the Azuma-Hoeffding inequality

- General technique of applications

- Say that a function $f(\bar{X}) = f(X_1, \ldots, X_n)$ satisfies Lipschitz condition with bound c if for any i, x, y_i

$$|f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, y_i, \ldots, x_n)| \le c$$

Changing single coordinate can change the function value by at most c

- Let $Z_k = E[f(\bar{X})|X_1, \ldots, X_k]$ be estimate of $f(\bar{X})$ upon kth observation ($Z_0 = E[f(\bar{X})]$) → Z_k is clearly Doob martingale

- Since $Z_k - Z_{k-1}$ is bounded within interval whose width is less than c from Lipschitz condition (proof omitted), we can apply Azuma-Hoeffding inequality (Theorem 12.6) to derive the bound of $f(\bar{X}) - E[f(\bar{X})]$

Applications: Balls and bins

- Throwing m balls independently and uniformly into n bins
- X_i: RV representing the bin into which ith ball falls

- F: # of empty bins after m balls are thrown
- $E[F] = n\left(1 - \frac{1}{n}\right)^m$
- The sequence $Z_i = E[F|X_1, \dots, X_i]$ is a Doob martingale
- $F = f(X_1, \dots, X_n)$ satisfies Lipschitz condition with bound 1

- Therefore from Azuma-Hoeffding inequality (Theorem 12.6)

$$\Pr(|F - E[F]| \geq \epsilon) \leq 2e^{-\frac{2\epsilon^2}{m}}$$

- We can derive the bound even without knowing $E[F]$