Lecture 13: Martingales and Azuma's Inequality (Few Details)

Doob's Martingale I

- We prove that Doob's construction yields a martingale
- Suppose $\mathbb{X}_1, \dots, \mathbb{X}_n$ are random variables over the sample space $\Omega_1, \dots, \Omega_n$ respectively
- Let $\mathbb{X} = (\mathbb{X}_1, \dots, \mathbb{X}_n)$ be the random variable over $\Omega = \Omega_1 \times \cdots \times \Omega_n$
- Let $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ be the *natural filtration* associated with $(\mathbb{X}_1, \dots, \mathbb{X}_n)$
- Suppose $f \colon \Omega \to \mathbb{R}$ be a function
- For $0 \leqslant i \leqslant n$, consider the function $g_i \colon \Omega \to \mathbb{R}$ defined as follows

$$g_i(x) = \mathbb{E}\left[f(\mathbb{X}_1, \dots, \mathbb{X}_n) | \mathcal{F}_i\right](x)$$

Suppose $(x_1, \dots, x_i) = (\omega_1, \dots, \omega_i)$. Then, the function $g_i(x)$ is the conditional expectation of $f(y)$, for all y such that $(y_1, \dots, y_i) = (\omega_1, \dots, \omega_i)$.

Doob's Martingale II

- First observation

Observation

For $0 \leqslant i \leqslant n$, the function g_i is $\mathcal{F}_i$-measurable.

This is easy to see because if $\mathcal{F}_i(x) = \mathcal{F}_i(y)$, i.e., the first i outcomes of x and y match, then we have $g_i(x) = g_i(y)$.

- Define the random variable $\mathbb{G}_i = g_i(\mathbb{X})$.

Observations.

- Observe that the random variable $\mathbb{G}_i$ is $\mathcal{F}_i$-measurable
- Note that $\mathbb{G}_0 = \mathbb{E}\left[f(\mathbb{X})\right]$

Doob's Martingale IV

- Crucial lemma

Lemma

$$\mathbb{E}\left[\mathbb{G}_{i+1}|\mathcal{F}_i\right](x) = (\mathbb{G}_i|\mathcal{F}_i)(x)$$

The proof is on the next slide. Note that this result suffices to show that $(\mathbb{G}_0, \ldots, \mathbb{G}_n)$ is a martingale with respect to the natural filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1, \subset \cdots \subset \mathcal{F}_n$

Doob's Martingale V

Proof.

- Suppose $(x_1, \ldots, x_i) = (\omega_1, \ldots, \omega_i)$
- Note that the RHS is $(\mathbb{G}_i|\mathcal{F}_i)(x) = \mathbb{E}\left[f(\omega_1, \ldots, \omega_i, \mathbb{X}_{i+1}, \ldots, \mathbb{X}_n)\right]$
- Note that the LHS is

$$
\begin{aligned}
&\sum_{y\in\Omega} \mathbb{P}\left[\mathbb{X} = y|\mathbb{X}_1 = \omega_1, \ldots, \mathbb{X}_i = \omega_i\right] g_{i+1}(y) \\
=&\sum_{y\in\Omega} \sum_{\omega_{i+1}\in\Omega_{i+1}} \mathbb{P}\left[\mathbb{X} = y, \mathbb{X}_{i+1} = \omega_{i+1}|\mathbb{X}_1 = \omega_1, \ldots, \mathbb{X}_i = \omega_i\right] g_{i+1}(y) \\
=&\sum_{\omega_{i+1}\in\Omega_{i+1}} \mathbb{P}\left[\mathbb{X}_{i+1} = \omega_{i+1}|\mathbb{X}_1 = \omega_1, \ldots, \mathbb{X}_i = \omega_i\right] \\
&\qquad \sum_{y\in\Omega} \mathbb{P}\left[\mathbb{X} = y|\mathbb{X}_1 = \omega_1, \ldots, \mathbb{X}_i = \omega_i, \mathbb{X}_{i+1} = \omega_{i+1}\right] g_{i+1}(y) \\
=&\sum_{\omega_{i+1}\in\Omega} \mathbb{P}\left[\mathbb{X}_{i+1} = \omega_{i+1}|\mathbb{X}_1 = \omega_1, \ldots, \mathbb{X}_i = \omega_i\right] \\
&\qquad \mathbb{E}\left[f(\omega_1, \ldots, \omega_{i+1}, \mathbb{X}_{i+2}, \ldots, \mathbb{X}_n)\right] \\
=&\mathbb{E}\left[f(\omega_1, \ldots, \omega_i, \mathbb{X}_{i+1}, \ldots, \mathbb{X}_n)\right]
\end{aligned}
$$

□

Application of Hoeffding's Lemma in Azuma's InequalityI

- Let $(\Delta\mathbb{G}_1, \ldots, \Delta\mathbb{G}_n)$ be a martingale difference sequence with respect to a filtration $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1, \subset \cdots \subset \mathcal{F}_n$
- For $1 \leqslant i \leqslant n$ and $x \in \Omega$ let $S_{i,x}$ be the support of the conditional distribution $(\Delta\mathbb{G}_i | \mathcal{F}_{i-1})(x)$. Let $a_{i,x}$ and $b_{i,x}$ be the infimum and the supremum of the elements in $S_{i,x}$. Suppose, there exists c_i such that $b_{i,x} - a_{i,x} \leqslant c_i$.
- Our goal is to prove a crucial step in the proof of Azuma's inequality that shows

$$\mathbb{P}\left[\sum_{i=1}^{n} \Delta\mathbb{G}_i \geqslant t\right] \leqslant \exp\left(\frac{2t^2}{\sum_{i=1}^{n} c_i^2}\right)$$

- The proof is similar to the proof of the Hoeffding's bound, except a crucial step. Our focus is that particular step. We want to claim the following

$$\mathbb{E}\left[\exp\sum_{i=1}^{n} h\Delta\mathbb{G}_i\right] \leqslant \exp\left(\frac{h^2\sum_{i=1}^{n} c_i^2}{8}\right)$$

For Hoeffding's bound, this was easy, because $\Delta\mathbb{G}_i$ variables were independent. So, we did the following manipulation in the Hoeffding's bound

$$\mathbb{E}\left[\exp\sum_{i=1}^{n} h\Delta\mathbb{G}_i\right] = \prod_{i=1}^{n}\mathbb{E}\left[\exp h\Delta\mathbb{G}_i\right]$$
$$\leqslant \prod_{i=1}^{n}\exp\left(\frac{h^2 c_i^2}{8}\right) = \exp\left(\frac{h^2\sum_{i=1}^{n} c_i^2}{8}\right)$$

- However, we do not have the independence guarantee in martingale difference sequences. We need to proceed in an alternate manner. In the sequel, we prove the result for martingale difference sequences.
- Our goal is to upper-bound the quantity

$$\mathbb{E}\left[\exp h \sum_{i=1}^{n} \Delta\mathbb{G}_i\right]$$

- This expression is equivalent to

$$\sum_{\omega_1,\dots,\omega_n} \mathbb{P}\left[\Delta\mathbb{G}_1 = \omega_1, \dots, \Delta\mathbb{G}_n = \omega_n\right] \exp(h(\omega_1 + \cdots + \omega_n))$$

- By the chain rule, we can express it as

$$\sum_{\omega_1,\ldots,\omega_{n-1}} \mathbb{P}\left[\Delta\mathbb{G}_1 = \omega_1, \ldots, \Delta\mathbb{G}_{n-1} = \omega_{n-1}\right] \exp(h(\omega_1 + \cdots + \omega_{n-1}))$$

$$\sum_{\omega_n} \mathbb{P}\left[\Delta\mathbb{G}_n = \omega_n | \Delta\mathbb{G}_1 = \omega_1, \ldots, \Delta\mathbb{G}_{n-1} = \omega_{n-1}\right] \exp(h\omega_n)$$

- Note that the random variable $\left(\Delta\mathbb{G}_n = \omega_n | \Delta\mathbb{G}_1 = \omega_1, \ldots, \Delta\mathbb{G}_{n-1} = \omega_{n-1}\right)$ has mean 0 (because it is a martingale difference sequence) and the difference between the maximum and minimum values this random variable achieves is c_n (irrespective of the values of $\omega_1, \ldots, \omega_{n-1}$). We can apply Hoeffding's lemma on <u>this</u> variable. So, we get that the previous expression is

$$\leqslant \sum_{\omega_1,\ldots,\omega_{n-1}} \mathbb{P}\left[\Delta\mathbb{G}_1 = \omega_1, \ldots, \Delta\mathbb{G}_{n-1} = \omega_{n-1}\right] \exp(h(\omega_1 + \cdots + \omega_{n-1})) \exp\left(\frac{h^2 c_n^2}{8}\right)$$

- Now, we rearrange this expression to get $\exp\left(-\frac{h^2 c_n^2}{8}\right)$ out of the summation. And, we can use induction on the remaining expression.
- As a consequence, we get the upper-bound

$$\leqslant \prod_{i=1}^{n} \exp(h^2 c_i^2/8)$$

This is exactly what we set out to prove initially.