Lecture 19: Martingale Difference Sequence & Azuma-Hoeffding Inequality

Martingale Difference Sequence

- Let $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ be a filtration
- Let $(\mathbb{F}_0, \mathbb{F}_1, \ldots, \mathbb{F}_n)$ be a martingale sequence with respect to the filtration above
- Let $\mathbb{Y}_0 = \mathbb{F}_0$, and $\mathbb{Y}_{t+1} = \mathbb{F}_{t+1} - \mathbb{F}_t$, for $0 \leqslant t < n$
- Intuition: $\mathbb{Y}_{t+1}$ measures the increase in $\mathbb{Y}_{t+1}$ from $\mathbb{Y}_t$. If $\mathbb{Y}_{t+1}$ is negative then it implies that $\mathbb{Y}_{t+1}$ is smaller than $\mathbb{Y}_t$
- Note that $\mathbb{E}\left[\mathbb{Y}_{t+1} | \mathcal{F}_t\right] = 0$, because we have $\mathbb{E}\left[\mathbb{F}_{t+1} | \mathcal{F}_t\right] = \mathbb{F}_t$

Theorem (Azuma's Inequality)

Suppose $(\mathbb{Y}_0, \ldots, \mathbb{Y}_n)$ *be a martingale difference sequence with respect to the filtration* $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$. *Assume that the following condition holds for all* $x \in \Omega$ *and* $0 \leqslant t < n$.

$$\max_{y \in \mathcal{F}(x)} \mathbb{Y}_{t+1}(y) - \min_{y \in \mathcal{F}(x)} \mathbb{Y}_{t+1}(y) \leqslant c_{t+1}$$

Then the following large deviation bound holds

$$\mathbb{P}\left[\sum_{i=1}^{n} \mathbb{Y}_i \geqslant E\right] \leqslant \exp\left(-2E^2 / \sum_{i=1}^{n} c_i^2\right)$$

Subtlety. Fix t. For different $x \in \Omega$, it is possible that $\max_{y \in \mathcal{F}(x)} \mathbb{Y}_{t+1}(y)$ is different from $\min_{y \in \mathcal{F}(x)} \mathbb{Y}_{t+1}(y)$. All that matters is that their difference is bounded by c_{t+1}.

Proof Outline

- The proof outline is identical to the Hoeffding bound proof.
- If we prove the following bound, then we are done. For any $h > 0$, we have

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

  This form of the inequality should remind us that we should be using the Hoeffding's Lemma in our proof.

Differences from Hoeffding's Bound

- The distribution $\mathbb{Y}_{t+1}$ can depend on the previous outcomes $(\omega_1, \ldots, \omega_t)$
- For different $x \in \Omega$, it is possible that $\max_{y \in \mathcal{F}(x)} \mathbb{Y}_{t+1}(y)$ is different from $\min_{y \in \mathcal{F}(x)} \mathbb{Y}_{t+1}(y)$. All that matters is that their difference is bounded by c_{t+1}