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 \((\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_n)\) be a martingale sequence with respect to the filtration above
- Let \( Y_0 = \mathcal{F}_0 \), and \( Y_{t+1} = \mathcal{F}_{t+1} - \mathcal{F}_t \), for \( 0 \leq t < n \)
- Intuition: \( Y_{t+1} \) measures the increase in \( Y_{t+1} \) from \( Y_t \). If \( Y_{t+1} \) is negative then it implies that \( Y_{t+1} \) is smaller than \( Y_t \)
- Note that \( \mathbb{E} [Y_{t+1} | \mathcal{F}_t] = 0 \), because we have \( \mathbb{E} [\mathcal{F}_{t+1} | \mathcal{F}_t] = \mathcal{F}_t \)
Suppose \((Y_0, \ldots, Y_n)\) 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 \(0 \leq t < n\).

\[
\max_{y \in \mathcal{F}(x)} Y_{t+1}(y) - \min_{y \in \mathcal{F}(x)} Y_{t+1}(y) \leq c_{t+1}
\]

Then the following large deviation bound holds

\[
P \left[ \sum_{i=1}^{n} Y_i \geq E \right] \leq \exp \left( \frac{-2E^2}{\sum_{i=1}^{n} c_i^2} \right)
\]

**Subtlety.** Fix \(t\). For different \(x \in \Omega\), it is possible that \(\max_{y \in \mathcal{F}(x)} Y_{t+1}(y)\) is different from \(\min_{y \in \mathcal{F}(x)} Y_{t+1}(y)\). 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

\[
\mathbb{E} \left[ \exp \left( h \sum_{i=1}^{n} Y_i \right) \right] \leq \exp \left( \frac{h^2}{8} \sum_{i=1}^{n} c_i^2 \right)
\]

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 $\max_{y \in F(x)} \mathbb{Y}_{t+1}(y)$ is different from $\min_{y \in F(x)} \mathbb{Y}_{t+1}(y)$. All that matters is that their difference is bounded by $c_{t+1}$. 

Azuma’s Inequality