Lecture 4

Recap:

Hoeffding's inequality

$X_1, \dots, X_n$ are indep. with $\mathbb{E}[X_i] = 0$

$X_i \in [a_i, b_i]$

$$\mathbb{P}\left(\sum_{i=1}^n X_i > t\right) \le \exp\left(\frac{-2t^2}{\sum_{i=1}^n (a_i - b_i)^2}\right).$$

Bennett's inequality

Same assumptions as above: $|X_i| < c$

$$\sigma^2 = \frac{1}{n} \sum_{t=1}^n Var(X_i)$$

$$\mathbb{P}\left(\sum_{i=1}^n X_i > t\right) \le \exp\left(\frac{-t^2}{2n\sigma^2 + \frac{2ct}{3}}\right).$$

Agenda

$\boxed{A}$ Proof of Bennett's ineq. (previous lecture notes)

$\boxed{B}$ Azuma's inequality

$\boxed{C}$ McDiarmid's inequality

Beyond indep. rvs and sums

[B] Azuma's inequality

We now prove a concentration bound for $\sum_{i=1}^{n} X_i$ when $X_1, \dots, X_n$ are not independent, but satisfy a more relaxed condition.

We assume throughout $\mathbb{E}[X_i] = 0$.

__Definition__ (Multiplicative family)

A collection of rvs $X_1, \dots, X_n$ s.t. $\mathbb{E}[X_i] = 0$, for all $1 \le i \le n$, constitutes a multiplicative family if for every $i_1, \dots, i_k$, $k \le n$,

$$\mathbb{E}[X_{i_1} X_{i_2} \dots X_{i_k}] = 0.$$

__Examples of multiplicative family__

(a) indep. $X_1, \dots, X_n$

(b) $\{X_i\}_{1 \le i \le n}$ is a martingale difference seq.

i.e., $\exists$ a martingale $Z_1, \dots, Z_n$ s.t. $X_i = Z_i - Z_{i-1}$.

Then, $\mathbb{E}[X_{i_1} \dots X_{i_k}] = \mathbb{E}[(Z_{i_k} - Z_{i_k - 1}) . X_{i_1} \dots X_{i_{k-1}}]$

$$= \mathbb{E}\left[\mathbb{E}\left[(Z_{i_k} - Z_{i_{k-1}}) . X_{i_1} \ldots X_{i_{k-1}} \,\middle|\, \mathcal{F}_{i_{k-1}}\right]\right]$$

$$= \mathbb{E}\left[X_{i_1} \ldots X_{i_{k-1}}\, \mathbb{E}\left[Z_{i_k} - Z_{i_{k-1}} \,\middle|\, \mathcal{F}_{i_{k-1}}\right]\right]$$

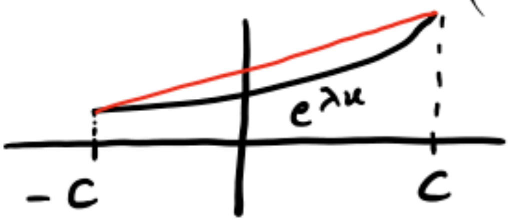
$$= 0$$

<u>Lemma</u> (Azuma's inequality)

For a multiplicative family $X_1, \ldots, X_n$ with $|X_i| \le c_i$,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i > t\right) \le \exp\left(-t^2 \Big/ 2\sum_{i=1}^{n} c_i^2\right)$$

Proof.



$$e^{\lambda x} \le ax + b \quad \forall\, x \in [-c, c] \text{ with}$$
$$e^{-\lambda c} = -ac + b \quad \text{and} \quad e^{\lambda c} = +ac + b$$
$$\Rightarrow \quad a = \frac{e^{\lambda c} - e^{-\lambda c}}{2c}, \quad b = \frac{e^{-\lambda c} + e^{\lambda c}}{2}$$

Thus,
$$\mathbb{E}\left[e^{\lambda \sum_{i=1}^{n} X_i}\right] = \mathbb{E}\left[\prod_{i=1}^{n} e^{\lambda X_i}\right]$$
$$\le \mathbb{E}\left[\prod_{i=1}^{n} (a_i X_i + b_i)\right]$$
$$= \mathbb{E}\left[\prod_{i=1}^{n} b_i\right] \quad \text{(by the multiplicative prop)}$$

$$= \prod_{i=1}^{n} \left(\frac{e^{-\lambda c_i} + e^{\lambda c_i}}{2} \right)$$

$$= \prod_{i=1}^{n} \left(1 + \frac{\lambda^2 c_i^2}{2!} + \frac{\lambda^4 c_i^4}{4!} + \cdots \right)$$

$$\leq \prod_{i=1}^{n} e^{\lambda^2 c_i^2 / 2}, \quad \forall \lambda > 0.$$

Therefore, $\psi_{\sum_{i=1}^{n} X_i}(\lambda) \leq \frac{\lambda^2}{2} \sum_{i=1}^{n} c_i^2$. ∎

Remark. (Application for Martingales)

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} (Z_i - Z_{i-1}) = Z_n - Z_0 .$$

Therefore, Azuma's inequality implies that for a martingale $\{Z_n\}$ such that $|\Delta_i| \leq c$,

$$\mathbb{P}\left(Z_n - Z_0 > c\sqrt{2n \log \frac{1}{\delta}} \right) \leq \delta .$$

In fact, using Doob's maximal inequality, we can get the same bound for $\max_{1 \leq i \leq n} (Z_i - Z_0)$.

* Martingale : $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n$

subMartingale : $\mathbb{E}[X_{n+1} | \mathcal{F}_n] \geq X_n$

Doob's maximal inequality

Let $\{X_n\}_{n\in\mathbb{N}}$ be a submartingale. Then,

$$\mathbb{P}\left(\sup_{0\le i\le n} X_i > t\right) \le \frac{\mathbb{E}[|X_i|]}{t}.$$

How to use this for martingales?

If $\{Z_n\}$ is a martingale $\Rightarrow \{Z_n - Z_0\}$ martingale

$\Rightarrow e^{\lambda(Z_n - Z_0)}$ submartingale (why?)

$$\Rightarrow \mathbb{P}\left(\sup_{0\le i\le n}(Z_i - Z_0) > t\right)$$

$$\le \mathbb{P}\left(\sup_{0\le i\le n} e^{\lambda(Z_i - Z_0)} > e^{\lambda t}\right)$$

$$\le \frac{\mathbb{E}\left[e^{\lambda(Z_n - Z_0)}\right]}{e^{\lambda t}} \quad \text{(by Doob's ineq.)}$$

Now, we can bound $\psi_{Z_n - Z_0}(\lambda)$ as in Azuma's inequality to get

$$\mathbb{P}\left(\sup_{0\le i\le n} Z_i - Z_0 > t\right) \le \exp\left(-\frac{t^2}{2nc^2}\right).$$

This tells us how martingale noise works:

Let $Y_n = f(n) + M_n$, $n \in \mathbb{N}$. Then,

$$\mathbb{P}\left(\sup_{0 \le n \le T} |Y_n - f(n)| > t \right) \le \exp\left(-\frac{T^2}{2nc}\right)$$

where we have assumed $|M_n - M_{n-1}| \le c$ a.s.

$\rightarrow$ no more that $O(\sqrt{n})$

— x —

[C] McDiarmid's inequality

We now move beyond "sum" function and present the first instantiation of Talagrand's principle.

Definition (Bounded difference property)

A function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies BDP with $\underline{c} = (c_1, \dots, c_n) \in \mathbb{R}_+^n$ if

$$f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n) < c_i, \quad \forall\, x_i, x_i', \; 1 \le i \le n.$$

Another interpretation:

Consider a rescaled version of Hamming distance: $d_c(x,y) = \sum_{i=1}^{n} c_i \mathbb{1}(x_i \neq y_i)$.

Then, f satisfies BDP with c iff

$$|f(x) - f(y)| \leq d_c(x,y),$$

namely f is 1-Lipchitz in $\underbrace{d_c(\cdot,\cdot)}_{\text{semi-metric}}$.

<u>Lemma</u> (McDiarmid)

Consider a function f satisfying BDP with $\underline{c}$.

Then, for $Z = f(X_1, \dots, X_n)$ with indep. $X_1, \dots, X_n$,

$$\mathbb{P}(Z - \mathbb{E}Z > t) \leq \exp\left(-\frac{t^2}{2\|c\|_2^2}\right).$$

(and similarly,

$$\mathbb{P}(-Z + \mathbb{E}Z > t) \leq \exp\left(\frac{-t^2}{2\|c\|_2^2}\right).)$$

<u>Proof</u>. Let $X_i = \mathbb{E}[Z \mid X_1, \dots, X_i]$.

Then, Y_i is a martingale. Also, ⑧

$$
\begin{aligned}
Y_i - Y_{i-1} &= \mathbb{E}[Z \mid X^i = x^i] - \mathbb{E}[Z \mid X^{i-1} = x^{i-1}] \\
&= \mathbb{E}[Z \mid X^i = x^i] \; \mathbb{E}\big[\mathbb{E}[Z \mid X^i] \mid X^{i-1} = x^{i-1}\big] \\
&= \mathbb{E}\Big[\, \mathbb{E}[Z \mid X^i = x^i] - \mathbb{E}[Z \mid X^{i-1} = x^{i-1}, X_i] \;\Big|\; X^{i-1} = x^{i-1} \Big] \\
&\leq c_i .
\end{aligned}
$$

Therefore, by Azuma's inequality,

$$
\mathbb{P}\left(\sum_{i=1}^{n} (Y_i - Y_{i-1}) > t \right) = \mathbb{P}(Z - \mathbb{E}Z > t)
$$

$$
\leq \exp\left(- \frac{t^2}{2\|c\|_2^2} \right).
$$

$\square$