Advanced Algorithms (V)

Shanghai Jiao Tong University

Chihao Zhang

March 30, 2020

Review

Hoeffding Inequality

Let $X = \sum_{i=1}^{n} X_i$ where each $X_i \in [a_i, b_i]$. If all X_i are independent, then

$$\Pr\left[X - \mathbf{E}[X] \geq t\right] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

Martingale

Martingale

Given a sequence of finite variables $\{Z_n\}_{n\geq 0}$, we call it a martingale w.r.t. another sequence $\{X_n\}_{n\geq 0}$ if for all $n \geq 1$:

$$\mathbf{E}[Z_n \mid X_0, X_1, \ldots, X_{n-1}] = Z_{n-1}$$

Martingale

Given a sequence of finite variables $\{Z_n\}_{n\geq 0}$, we call it a martingale w.r.t. another sequence $\{X_n\}_{n\geq 0}$ if for all $n \geq 1$:

$$\mathbf{E}[Z_n \mid X_0, X_1, \ldots, X_{n-1}] = Z_{n-1}$$

"$\leq$" - super-martingale
"$\geq$" - sub-martingale

Martingale

Given a sequence of finite variables $\{Z_n\}_{n\ge0}$, we call it a martingale w.r.t. another sequence $\{X_n\}_{n\ge0}$ if for all $n \ge 1$:

$$\mathbf{E}[Z_n \mid X_0, X_1, \ldots, X_{n-1}] = Z_{n-1}$$

"$\le$" - super-martingale
"$\ge$" - sub-martingale

- Z_n is usually a function of $X_0, X_1, \ldots, X_n$
- Variables $\{X_n\}$ are not necessarily independent

More formally, assume the probability space is $(\Omega, \mathscr{F}, \Pr)$

More formally, assume the probability space is $(\Omega, \mathscr{F}, \Pr)$

If we use $\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ to denote the σ-algebra generated by $\{X_i\}_{0 \le i \le n}$, the condition becomes to

$$\mathbf{E}[Z_n \mid \mathscr{F}_{n-1}] = Z_{n-1}$$

More formally, assume the probability space is $(\Omega, \mathscr{F}, \Pr)$

If we use $\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ to denote the σ-algebra generated by $\{X_i\}_{0 \le i \le n}$, the condition becomes to

$$\mathbf{E}[Z_n \mid \mathscr{F}_{n-1}] = Z_{n-1}$$

$\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \cdots \subseteq \mathscr{F}$ is a family of *filtrations*

Examples

Examples

- $X_1, X_2, \ldots$ independent with $\mathbf{E}[X_i] = 0$, $Z_n = \sum_{i=1}^{n} X_i$

Examples

- $X_1, X_2, \ldots$ independent with $\mathbf{E}[X_i] = 0$, $Z_n = \sum_{i=1}^{n} X_i$

- $X_1, X_2, \ldots$ independent with $\mathbf{E}[X_i] = 1$, $Z_n = \prod_{i=1}^{n} X_i$

Doob Martingale

Doob Martingale

Let $\bar{X}_n = X_0, X_1, \ldots, X_n$ be a sequence of random variables and let $f(X_1, \ldots, X_n) = f(\bar{X}_n)$ be a function

Doob Martingale

Let $\bar{X}_n = X_0, X_1, \ldots, X_n$ be a sequence of random variables and let $f(X_1, \ldots, X_n) = f(\bar{X}_n)$ be a function

Define $\forall i \geq 0, \quad Z_i = \mathbf{E}[f(\bar{X}_n) \mid X_1, \ldots, X_i]$

Doob Martingale

Let $\bar{X}_n = X_0, X_1, \ldots, X_n$ be a sequence of random variables and let $f(X_1, \ldots, X_n) = f(\bar{X}_n)$ be a function

Define $\forall i \geq 0, \quad Z_i = \mathbf{E}[f(\bar{X}_n) \mid X_1, \ldots, X_i]$

$\{Z_i\}_{i \geq 0}$ is a martingale w.r.t. $\{X_i\}_{i \geq 0}$

Doob Martingale

Let $\bar{X}_n = X_0, X_1, \ldots, X_n$ be a sequence of random variables and let $f(X_1, \ldots, X_n) = f(\bar{X}_n)$ be a function

Define $\forall i \geq 0, \quad Z_i = \mathbf{E}[f(\bar{X}_n) \mid X_1, \ldots, X_i]$

$\{Z_i\}_{i \geq 0}$ is a martingale w.r.t. $\{X_i\}_{i \geq 0}$

Proof.

$$
\begin{aligned}
\mathbf{E}[Z_i \mid \bar{X}_{i-1}] &= \mathbf{E}[\mathbf{E}[f(\bar{X}_n) \mid \bar{X}_i] \mid \bar{X}_{i-1}] \\
&= \mathbf{E}[f(\bar{X}_n) \mid \bar{X}_{i-1}] = Z_{i-1}
\end{aligned}
$$

Vertex Exposure Martingale

Vertex Exposure Martingale

Consider a random graph $G \sim G(n, p)$

Vertex Exposure Martingale

Consider a random graph $G \sim G(n, p)$

Let $F(G)$ be any function defined on G, e.g.
$F(\cdot) = \chi(\cdot)$, the chromatic number of G

Vertex Exposure Martingale

Consider a random graph $G \sim G(n, p)$

Let $F(G)$ be any function defined on G, e.g.
$F(\cdot) = \chi(\cdot)$, the chromatic number of G

For $i = 0, \ldots, n$, X_i denotes the edges between vertex i and vertices $j < i$

Vertex Exposure Martingale

Consider a random graph $G \sim G(n, p)$

Let $F(G)$ be any function defined on G, e.g. $F(\cdot) = \chi(\cdot)$, the chromatic number of G

For $i = 0, \ldots, n$, X_i denotes the edges between vertex i and vertices $j < i$

$Z_i = \mathbf{E}[F(G) \mid X_1, \ldots, X_i]$ is a Doob martingale

Azuma-Hoeffding

Azuma-Hoeffding

Let $\{X_n\}_{n \geq 1}$ be random variables such that $a_i \leq X_i \leq b_i$

Azuma-Hoeffding

Let $\{X_n\}_{n \geq 1}$ be random variables such that $a_i \leq X_i \leq b_i$

Let $S_n = \sum_{i=0}^{n} X_i$ be a martingale (or super-martingale)

Azuma-Hoeffding

Let $\{X_n\}_{n\geq 1}$ be random variables such that $a_i \leq X_i \leq b_i$

Let $S_n = \sum_{i=0}^{n} X_i$ be a martingale (or super-martingale)

For all $n \geq 1$ and $t > 0$, $\Pr[S_n - S_0 \geq t] \leq e^{-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}}$

Azuma-Hoeffding

Let $\{X_n\}_{n\geq 1}$ be random variables such that $a_i \leq X_i \leq b_i$

Let $S_n = \sum_{i=0}^{n} X_i$ be a martingale (or super-martingale)

For all $n \geq 1$ and $t > 0$, $\Pr[S_n - S_0 \geq t] \leq e^{-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}}$

The key to the proof is an upper bound on $\mathbf{E}[\exp(\delta(S_n - S_0))]$

We have

Note that $\mathbf{E}[\exp(\delta(S_n - S_0))] = \mathbf{E}\left[\exp\left(\delta \cdot \sum_{i=1}^{n} X_i\right)\right]$

We have

Note that $\mathbf{E}[\exp(\delta(S_n - S_0))] = \mathbf{E}\left[\exp\left(\delta \cdot \sum_{i=1}^{n} X_i\right)\right]$

We have

$$\begin{aligned}
&\mathbf{E}\left[\exp\left(\delta \cdot \sum_{i=1}^{n} X_i\right)\right] = \mathbf{E}\left[\prod_{i=1}^{n} \exp(\delta X_i)\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[\prod_{i=1}^{n} \exp(\delta X_i) \middle| X_0, \ldots, X_{n-1}\right]\right] \\
&= \mathbf{E}\left[\prod_{i=1}^{n-1} \exp(\delta X_i) \cdot \mathbf{E}\left[\exp(\delta X_n) \middle| X_0, \ldots, X_{n-1}\right]\right]
\end{aligned}$$

We can prove

$$\mathbf{E}[\exp(\delta X_n) \mid X_0, \ldots, X_{n-1}] \leq \exp\left(\frac{\delta^2(b_n - a_n)^2}{8}\right)$$

similar to the case of the Hoeffding inequality

We can prove

$$\mathbf{E}[\exp(\delta X_n) \mid X_0, \ldots, X_{n-1}] \leq \exp\left(\frac{\delta^2(b_n - a_n)^2}{8}\right)$$

similar to the case of the Hoeffding inequality

Then an induction on n finishes the proof

Methods of Bounded Differences

Methods of Bounded Differences

A function $f(X_1, \ldots, X_n)$ satisfies *bounded differences condition with bounds* $\{c_i\}_{1 \leq i \leq n}$ if

Methods of Bounded Differences

A function $f(X_1, \ldots, X_n)$ satisfies *bounded differences condition with bounds* $\{c_i\}_{1 \le i \le n}$ if

$$\forall i, \forall x_1, \ldots, x_n, \forall x_i' : |f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)| \le c_i$$

Methods of Bounded Differences

A function $f(X_1, \ldots, X_n)$ satisfies *bounded differences condition with bounds* $\{c_i\}_{1 \le i \le n}$ if

$$\forall i, \forall x_1, \ldots, x_n, \forall x_i' : |f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)| \le c_i$$

We use f and $\{X_i\}_{i \ge 1}$ to define a Doob martingale $\{Z_i\}$

Methods of Bounded Differences

A function $f(X_1, \ldots, X_n)$ satisfies *bounded differences condition with bounds* $\{c_i\}_{1 \le i \le n}$ if

$$\forall i, \forall x_1, \ldots, x_n, \forall x_i' : |f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)| \le c_i$$

We use f and $\{X_i\}_{i \ge 1}$ to define a Doob martingale $\{Z_i\}$

When $\{X_i\}$ are independent, the bounded differences condition implies $B_i \le Z_i - Z_{i-1} \le B_i + c_i$ for some B_i

McDiarmid’s Inequality

McDiarmid's Inequality

Let f be a function on n variables satisfying bounded differences condition with $\{c_i\}_{1 \le i \le n}$

McDiarmid's Inequality

Let f be a function on n variables satisfying bounded differences condition with $\{c_i\}_{1 \le i \le n}$

Let $X_1, \ldots, X_n$ be independent variables

McDiarmid's Inequality

Let f be a function on n variables satisfying bounded differences condition with $\{c_i\}_{1 \leq i \leq n}$

Let $X_1, \ldots, X_n$ be independent variables

$$\Pr[\underbrace{f(X_1, \ldots, X_n)}_{= Z_n} - \underbrace{\mathbf{E}[f(X_1, \ldots, X_n)]}_{= Z_0} \geq t] \leq 2e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}$$

McDiarmid's Inequality

Let f be a function on n variables satisfying bounded differences condition with $\{c_i\}_{1 \le i \le n}$

Let $X_1, \ldots, X_n$ be independent variables

$$\Pr[\underbrace{f(X_1, \ldots, X_n)}_{= Z_n} - \underbrace{\mathbf{E}[f(X_1, \ldots, X_n)]}_{= Z_0} \ge t] \le 2e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}$$

This is a consequence of Azuma-Hoeffding and $Z_i - Z_{i-1} \in [B_i, B_i + c_i]$ for all i

Proof of

$$B_i \leq Z_i - Z_{i-1} \leq B_i + c_i$$

Proof of

$$B_i \leq Z_i - Z_{i-1} \leq B_i + c_i$$

$$Z_i - Z_{i-1} = \mathbf{E}[f(\bar{X}) \mid \bar{X}_i] - \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}]$$

Proof of $B_i \le Z_i - Z_{i-1} \le B_i + c_i$

$$Z_i - Z_{i-1} = \mathbf{E}[f(\bar{X}) \mid \bar{X}_i] - \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}]$$

Therefore,

$$Z_i - Z_{i-1} \le \sup_x \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}, X_i = x] - \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}]$$

$$Z_i - Z_{i-1} \ge \inf_y \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}, X_i = y] - \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}]$$

$$:= B_i$$

It suffices to bound

$$\sup_{x,y} \left(\mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}, X_i = x] - \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}, X_i = y] \right)$$

$$= \sup_{x,y} \left(\mathbf{E}[f_i(\bar{X}, x) \mid \bar{X}_{i-1}] - \mathbf{E}[f_i(\bar{X}, y) \mid \bar{X}_{i-1}] \right)$$

$$= \sup_{x,y} \mathbf{E}[f_i(\bar{X}, x) - f_i(\bar{X}, y \mid \bar{X}_{i-1}]$$

It suffices to bound

$$\sup_{x,y} \left(\mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}, X_i = x] - \mathbf{E}[f(\bar{X}) \mid \bar{X}_{i-1}, X_i = y] \right)$$

$$= \sup_{x,y} \left(\mathbf{E}[f_i(\bar{X}, x) \mid \bar{X}_{i-1}] - \mathbf{E}[f_i(\bar{X}, y) \mid \bar{X}_{i-1}] \right)$$

$$= \sup_{x,y} \mathbf{E}[f_i(\bar{X}, x) - f_i(\bar{X}, y \mid \bar{X}_{i-1}]$$

This quantity is upper bounded by c_i by the independence of $\{X_i\}_{i \geq 0}$

Application: Pattern Matching

Application: Pattern Matching

- $X = (X_1, \ldots, X_n) \in \{0,1\}^n$ - a random string
- $B = (B_1, \ldots, B_k) \in \{0,1\}^k$ - a fixed string

Application: Pattern Matching

- $X = (X_1, \ldots, X_n) \in \{0,1\}^n$ - a random string
- $B = (B_1, \ldots, B_k) \in \{0,1\}^k$ - a fixed string

What is the expected number of occurrences of B in X?

Application: Pattern Matching

- $X = (X_1, \ldots, X_n) \in \{0,1\}^n$ - a random string
- $B = (B_1, \ldots, B_k) \in \{0,1\}^k$ - a fixed string

What is the expected number of occurrences of B in X?

By linearity of expectation,

$$\mathbf{E}[F] = (n - k + 1)2^{-k}$$

F satisfies the bounded differences property with k

So $\Pr[\,|F - \mathbf{E}[F]\,| \geq \delta k\sqrt{n}] \leq 2e^{-2\delta^2}$

$F = F(X_1, \ldots, X_n)$ is a function of X

F satisfies the bounded differences property with k

So $\Pr[\,|F - \mathbf{E}[F]|\, \geq \delta k \sqrt{n}] \leq 2e^{-2\delta^2}$

$F = F(X_1, \ldots, X_n)$ is a function of X

So we can construct a Doob martingale where

F satisfies the bounded differences property with k

So $\Pr[\,|F - \mathbf{E}[F]|\, \geq \delta k\sqrt{n}] \leq 2e^{-2\delta^2}$

$F = F(X_1, \ldots, X_n)$ is a function of X

So we can construct a Doob martingale where

$$Z_0 = \mathbf{E}[F], \quad Z_i = \mathbf{E}[F \mid \bar{X}_i]$$

F satisfies the bounded differences property with k

So $\Pr[\,|F - \mathbf{E}[F]| \geq \delta k \sqrt{n}] \leq 2e^{-2\delta^2}$

$F = F(X_1, \ldots, X_n)$ is a function of X

So we can construct a Doob martingale where

$$Z_0 = \mathbf{E}[F], \quad Z_i = \mathbf{E}[F \mid \bar{X}_i]$$

If we change one bit of X, how much can F changes?

F satisfies the bounded differences property with k

So $\Pr[\,|F - \mathbf{E}[F]\,| \geq \delta k \sqrt{n}] \leq 2e^{-2\delta^2}$

Chromatic Number

Chromatic Number

Recall the vertex exposure martingale

$$Z_0 = \mathbf{E}[\chi(G)], \quad Z_i = \mathbf{E}[\chi(G) \mid X_1, \ldots, X_i]$$

Chromatic Number

Recall the vertex exposure martingale

$$Z_0 = \mathbf{E}[\chi(G)], \quad Z_i = \mathbf{E}[\chi(G) \mid X_1, \ldots, X_i]$$

Revealing the edges of a new vertex changes the chromatic number by at most 1

Chromatic Number

Recall the vertex exposure martingale

$$Z_0 = \mathbf{E}[\chi(G)], \quad Z_i = \mathbf{E}[\chi(G) \mid X_1, \ldots, X_i]$$

Revealing the edges of a new vertex changes the chromatic number by at most 1

So McDiarmid's Inequality implies
$\Pr[\,|\chi(G) - \mathbf{E}[\chi(G)]\,| \geq \delta\sqrt{n}] \leq 2e^{-2\delta^2}$

Chromatic Number

Recall the vertex exposure martingale

$$Z_0 = \mathbf{E}[\chi(G)], \quad Z_i = \mathbf{E}[\chi(G) \mid X_1, \ldots, X_i]$$

Revealing the edges of a new vertex changes the chromatic number by at most 1

So McDiarmid's Inequality implies
$\Pr[\,|\chi(G) - \mathbf{E}[\chi(G)]\,| \geq \delta\sqrt{n}] \leq 2e^{-2\delta^2}$

We obtain concentration without even knowing the expectation!