Sep 28, 2018

Goran Žužić

More Concentration Inequalities & Applications

Review

Fact (Chernoff): Given independent random variables (RVs) $X_1, X_2, \dots, X_n \in [0,1]$, we have that

$$\Pr\left[\left|\sum X_i - E\left[\sum X_i\right]\right| \geq \lambda\right] \leq 2\exp\left(c \cdot \frac{-\lambda^2}{E[\sum X_i] + \lambda}\right).$$

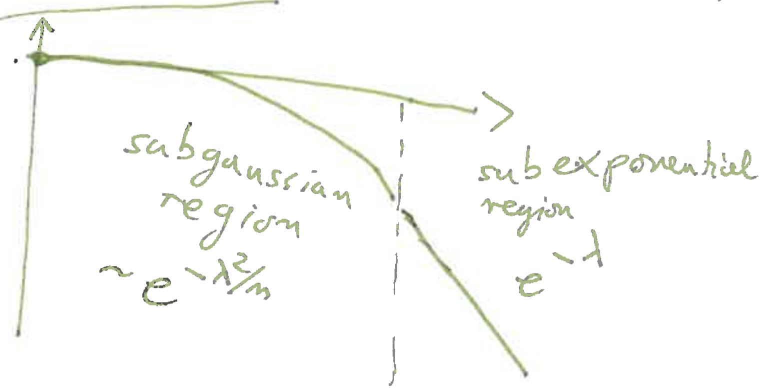
Here $c > 0$ is a sufficiently small constant.

Example: Let $X_1, X_2, \dots, X_n \in [-1, 1]$ be indep. RVs, then

$$\Pr\left[\left|\sum X_i\right| \geq \lambda\right] \leq 2\exp\left(c \cdot \frac{-\lambda^2}{n + \lambda}\right).$$ (actually not needed here)

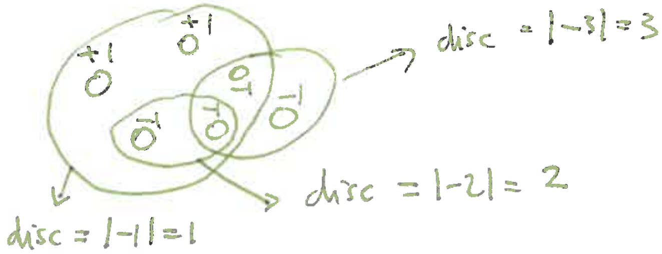
(we shift and scale s.t. X_i fit in $[0,1]$).

Behavior of tail bounds: consider $f(\lambda) = \exp\left(c \cdot \frac{-\lambda^2}{n+\lambda}\right)$.



Problem (set balancing / Discrepency): Let $\mathcal{U} = \{1, 2, \dots, n\}$ be items and $S_1, S_2, \dots, S_n \subseteq \mathcal{U}$ be n sets of items.
For some assignment $f : \mathcal{U} \to \{-1, +1\}$ we define the descrepency of a set S_i as $\text{disc}(S_i) = \left|\sum_{j \in S_i} f(j)\right|$ and the total discr. as $\max_i |\text{disc}(S_i)|$.

Prove there exist an assignment f s.t. t. disc. $\leq \sqrt{n \log n}$.



$\rightarrow$ total discrepancy of the assignment is 3.

Note: [Spencer '85] proves disc $\leq 6\sqrt{n}$; more generally if $n = \#$items, $m = \#$sets, he proves disc $\leq O(\sqrt{n \log \frac{m}{n}})$.

Algo (uniform random assignments): for each $u \in \mathcal{U}$ we assign a random Rademacher RV. Namely, $f(u) \leftarrow \begin{pmatrix} +1 & -1 \\ .5 & .5 \end{pmatrix}$.

Analysis: we prove that, under the uniform rand. assignment,

$$\text{disc} \leq O(\sqrt{n \log n}).$$

Pf. $\sum_{u \in S_i} f(u) \sim$ sum of $|S_i| \leq n$ Rad. RV's.

$$\Pr\Big[\big|\sum_{u \in S_i} f(u)\big| \geq \lambda\Big] \leq 2\exp\Big(c \cdot \frac{-\lambda^2}{n+\lambda}\Big). \qquad \text{(Chernoff)}$$

$$\Pr\Big[\max_i \big|\sum_{u \in S_i} f(u)\big| \geq \lambda\Big] \leq n \cdot 2\exp\Big(c \cdot \frac{-\lambda^2}{n+\lambda}\Big) \qquad \text{(union bound)}$$

Setting $\lambda = C \cdot \sqrt{n \log n}$ we get $\Pr[\max_i \text{disc}(S_i)] \lesssim n \cdot n^{-C} = \frac{1}{n^{C-1}} \ll 1$.

Note: if $n = \#$items and $m = \#$sets, we can get

$$\text{disc} \leq O(\sqrt{n \log m}).$$

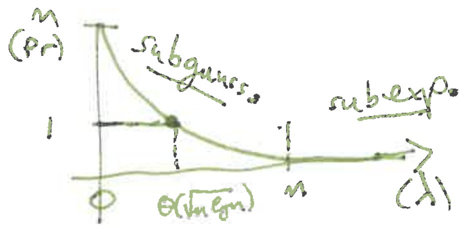
Reconstructing the expectation from tail bounds

Fact: let $X \geq 0$ be an RV. Then

$$E[X] = \int_0^\infty \Pr[X \geq \lambda]\, d\lambda.$$

Ex (from set balancing). Suppose we have

$\Pr[Y \ge \lambda] \le n \cdot \exp\left(+c \frac{-\lambda^2}{n+\lambda}\right)$. ($Y = \max_i \text{disc}(S_i)$ in set balancing)



Note that for $\lambda \ll \sqrt{n \log n}$, the bound $\Pr \le 1$ is better than the tail bound.

$$E[Y] = \int_0^\infty \Pr[Y \ge \lambda]\, d\lambda \le \int_0^{c\sqrt{n\log n}} 1 + \int_{c\sqrt{n\log n}}^{\infty} O\left(\frac{1}{\lambda^2}\right) d\lambda = c \cdot \sqrt{n \log n} + \cancel{O(1)}.$$

Unbiased Random Walk

Def: let $X_0 = 0, X_1, \ldots$ s.t. $X_{n+1} = X_n + \text{Rad} = X_n + \begin{pmatrix} +1 & -1 \\ 0.5 & 0.5 \end{pmatrix}$. independ.

then $E[X_n] = 0$, $E[|X_n|] = O(\sqrt{n})$, $\Pr[|X_n| \ge \lambda] \le 2\exp\left(c \cdot \frac{-\lambda^2}{n+\lambda}\right)$.

Question: what is $E[\max_{i=1}^{n} |X_i|]$? Chernoff + union bound will give us $O(\sqrt{n \log n})$, but this is not tight (unlike before). We can do better.

Thm: $E[\max_{i=1}^{n} |X_i|] = O(\sqrt{n})$.

Pf Let $\overset{0}{\bullet}\ \overset{1}{\bullet}\ \overset{2}{\bullet}\ \cdots\ \overset{n-1}{\bullet}$ be the individual Rad RVs.

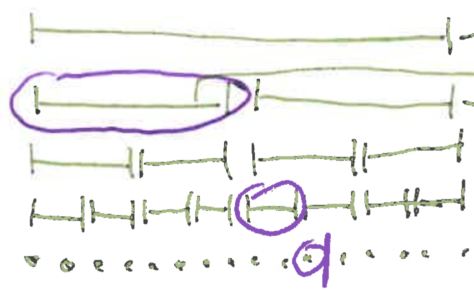
We construct a segment tree of sums.

level 0: → sum of all Rad RVs.

level 1: → sum of 1st half of RVs. → sum of 2nd ~~half~~ of RVs half

⋮

level log n:



X_K = sum of at most 1 interval on every level.

Hence $X_K = \sum_{i=0}^{\log n} \left\{ \begin{array}{l} \text{either } 0, \text{ or} \\ \text{some interval on} \\ \text{level } i \end{array} \right\}$

Let L_i = maximum absolute sum of an interval on level i.

Then $\max_K |X_K| \le L_0 + L_1 + \dots + L_{\log n}$. Also $E[\max|X_K|] \le \sum_{i=0}^{\log n} E[L_i]$.

However ~~L_i ... $\sum^m$ Rad where ...~~

$$L_i \sim \max_{j=1}^{m} \sum^{t} \text{Rad}, \quad \text{where} \quad \begin{array}{l} m = 2^i \\ t = n/2^i. \end{array}$$

so $\Pr[L_i \geq \lambda] \le \sqrt{t \cdot \log m} = O\left(\sqrt{n} \cdot \frac{\sqrt{i}}{2^{i/2}}\right)$

Hence $E \max |X_K| \le O(\sqrt{n}) \cdot \underbrace{\sum_{i=0}^{\log n} \frac{\sqrt{i}}{2^{i/2}}}_{\text{converges!}} = O(\sqrt{n})$ ✓

Martingales

<u>Def</u> A martingale is a sequence of RVs $X_0 = 0, X_1, X_2, X_3, \dots$ s.t. (i) $E[X_{K+1} \mid X_K, X_{K-1}, \dots, X_1] = X_K$ and (ii) $E[|X_K|] < \infty$.

(for simplicity)

<u>Ex</u> (unbiased random walk) $X_0 = 0, \dots, X_{K+1} = X_K + \text{Rad}$.

<u>Ex</u> $Y_n = (X_n)^2 - n$ is a martingale.

<u>Pf</u> $E[Y_{n+1} \mid Y_n] = \frac{1}{2}(X_n+1)^2 + \frac{1}{2}(X_n - 1)^2 - (n+1)$

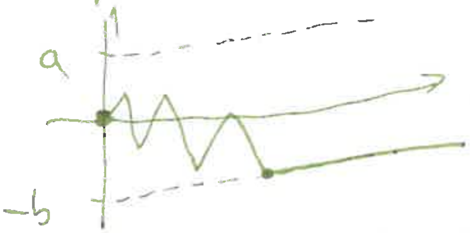
$= X_n^2 + 1 - n - 1 = X_n^2 - n = Y_n$ ✓

Fact: let $X_0=0, X_1, X_2, \ldots$ be a martingale. Then
$E[X_k] = E[X_0] = 0.$

Pf (easy)

Ex (stopped random walk). $Z_0 = 0, \ldots, Z_{k+1} = \begin{cases} Z_k & \text{if } Z_k \in \{a,-b\} \\ Z_k + \text{Rad} & \text{otherwise} \end{cases}$

This is a martingale!



Fact: a martingale Z_k (stopped at $\{a,-b\}$) will stop at time τ where $E[\tau] < \infty$.

Fact: let $p_a = \Pr[Z_k \text{ stops at } a]$, $p_b = \Pr[Z_k \text{ stops at } -b]$.
Then $p_a = \frac{b}{a+b}$ and $p_b = \frac{a}{a+b}$.

Pf: let N be such that $\Pr[Z_k \text{ stops before } N \text{ steps}] \geq 1-\varepsilon$.

$$0 = E[Z_N] = p_a^{(N)} \cdot a + p_b^{(N)}(-b) + \varepsilon \cdot \underbrace{O(1)}_{\leq \max(a,b)}.$$

Also $p_a + p_b = 1 - O(\varepsilon)$.

Plugging in & solving we get $p_a = \frac{b}{a+b} + O(\varepsilon) \longrightarrow \frac{b}{a+b}$

$p_b = \frac{a}{a+b} + O(\varepsilon) \longrightarrow \frac{a}{a+b}$.

Fact: let Z_k be the unbiased rand. walk stopped at $\{a,-b\}$.
Let τ be the stopping time. Then $E[\tau] = a \cdot b$.

idea: use $0 = E[Y_k]$ on $Y_k = (X_k)^2 - k$.

Fact (Azuma): let $0 = X_0, X_1, X_2, \ldots$ be a martingale s.t.
$|X_{k+1} - X_k| \leq 1$. Then $\Pr[|X_k| \geq \lambda] \leq 2\exp\left(\frac{-\lambda^2}{2k}\right)$.

↳ Note: we can use this on the stopped random walk to prove that $\max_{i=1}^{n} |X_i| \leq O(\sqrt{n})$ directly.