# Lecture 14

Tuesday, May 9, 2017 9:13 PM

## Markov's Inequality

$X$ be a non-negative r.v. Then for any $t > 0$,

$$\mathbb{P}[X \geq t] \leq \frac{\mathbb{E}X}{t}$$

Pf: 
$$\mathbb{E}X = \sum_x x \cdot \mathbb{P}[X = x]$$
$$= \underbrace{\sum_{x<t} x \cdot \mathbb{P}[X = x]}_{\geq 0} + \sum_{x \geq t} x\, \mathbb{P}[X = x]$$
$$\geq t \sum_{x \geq t} \mathbb{P}[X = x]$$
$$= t \cdot \mathbb{P}[X \geq t]$$

∎

## Chebyshev's Inequality

$X$ be any r.v. Then for any $t \geq 0$,

$$\mathbb{P}[|X - \mathbb{E}X| \geq t] \leq \frac{\operatorname{Var} X}{t^2}$$

The most "general" inequality with no assumptions on $X$.

Pf: 
$$\mathbb{P}[|X - \mathbb{E}X| \geq t] = \mathbb{P}[\underbrace{(X - \mathbb{E}X)^2}_{\text{a non-negative r.v.}} \geq t^2]$$
$$\leq \frac{\mathbb{E}(X - \mathbb{E}X)^2}{\ } = \frac{\operatorname{Var} X}{\ }$$

$$\underset{\text{Markov}}{\le} \frac{\mathbb{E}(X - \mathbb{E}X)^2}{t^2} = \frac{\mathrm{Var}\, X}{t^2}$$

$\blacksquare$

Often times, we will be interested in r.v.'s $Z$ which are <u>sum of independent</u> rvs.

$$Z = X_1 + X_2 + \cdots + X_n$$

and each $X_i$ is an indicator rv with

$$\mathbb{P}[X_i = 1] = p_i$$

$$\therefore \mathbb{E} Z = \sum_{i=1}^{n} p_i =: \mu$$

$$\mathrm{Var}\, Z = \sum_{i=1}^{n} \mathrm{Var}\, X_i = \sum_{i=1}^{n} p_i(1-p_i) \le \mu$$

$\because Z$ is a sum of <u>independent</u> random vars

$$\mathbb{P}_r\left[\,|Z - \mu| \ge \delta\mu\,\right] \le \, ??$$

Chebyshev gives us $\le \dfrac{\mathrm{Var}\, Z}{\delta^2 \mu^2} \le \dfrac{1}{\mu \delta^2}$

The "Chernoff Bound" gives us an exponentially better bound.

# Chernoff Bound

$Z = X_1 + \cdots + X_n$ where each $X_i$ is ind. rv with $\mathbb{E} X_i = p_i$ & $\mu = \sum p_i$

$$\mathbb{P}[Z \geq (1+\delta)\mu] \leq e^{-\mu((1+\delta)\ln(1+\delta) - \delta)}$$

...even for $-1 < \delta \leq 0$

"Interesting regions"

- when $\delta$ is "small" (think $\ll 1$)

  $\ln(1+\delta) \approx \delta$

$$\therefore \quad \mathbb{P}[Z \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/4} \quad \text{if } -1 \leq \delta \leq 1$$

- when $\delta$ is "large" ($\delta \gg 1$)

  and so only one "tail" is interesting

$$\mathbb{P}[Z \geq (1+\delta)\mu] \leq e^{-\mu\frac{\delta\ln\delta}{2}} \quad \text{if } \delta \geq 10$$

A similar stmt is true for the "other tail"

$$\mathbb{P}[Z \leq (1-\delta)\mu] \leq e^{-\mu[(1-\delta)\ln(1-\delta) + \delta]}$$

$$\leq e^{-\mu\delta^2/2} \quad \text{if } |\delta| \leq 1$$

(which'll be the case)

---

Pf:- (TRICK): Look @ $e^{tZ}$

$$\mathbb{P}[Z \geq (1+\delta)\mu] = \mathbb{P}[e^{tZ} \geq e^{(1+\delta)t\mu}]$$

for any $t > 0$.

Markov's Inequality then gives ...

$$\mathbb{P}[Z \geq (1+\delta)\mu] \leq \mathbb{E}[e^{tZ}] \cdot e^{-(1+\delta)t\mu}$$

$$\mathbb{P}[Z \geq (1+\delta)\mu] \leq \underbrace{\mathbb{E}[e^{tZ}]} \cdot e^{-(1+\delta)t\mu}$$

$$\mathbb{E}[e^{tZ}] = \mathbb{E}[e^{t\sum X_i}] = \prod_{i=1}^{n} \underbrace{\mathbb{E}[e^{tX_i}]}$$

only place where independence is used

$$\mathbb{E}[e^{tX_i}] = p_i e^t + (1-p_i)$$

$$= 1 + p_i(e^t - 1)$$

$$\leq e^{p_i(e^t-1)} \qquad \text{(only other trick)}$$

Putting things together, we get ...

$$\mathbb{P}[Z \geq (1+\delta)\mu] \leq e^{-(1+\delta)t\mu} \cdot e^{\sum_i p_i(e^t-1)}$$

$$= e^{-(1+\delta)t\mu + \mu(e^t-1)}$$

$$= e^{-\mu[t(1+\delta) - (e^t-1)]}$$

FOR ALL $t > 0$

In particular (after calculus) for $t = \ln(1+\delta)$

$$\therefore \mathbb{P}[Z \geq (1+\delta)\mu] \leq e^{-\mu[(1+\delta)\ln(1+\delta) - \delta]}$$

## Generalization (Exercise)

If $u \geq \mu$; $P[Z \geq (1+\delta)u] \leq e^{-u[(1+\delta)\ln(1+\delta) - \delta]}$

## Applications

① Toss n-coins. $X_i = \begin{cases} 1 & \text{if tails} \\ 0 & \text{" heads} \end{cases}$

- $Z = \sum X_i$ = # of tails.
- $\mathbb{E} Z = n/2$

$$\mathbb{P}\left[ \left| Z - \frac{n}{2} \right| \geq \delta \cdot \frac{n}{2} \right] \leq e^{-\frac{n\delta^2}{4}} \quad \text{for } \delta \leq 1$$

In particular if $\delta \equiv \frac{2}{\sqrt{n}}$, we get

$$\mathbb{P}\left[ \left| Z - \frac{n}{2} \right| \geq \sqrt{n} \right] \leq e^{-1}$$

& more generally

$$\mathbb{P}\left[ \left| Z - \frac{n}{2} \right| \geq c\sqrt{n} \right] \leq e^{-c^2}$$

$\delta = \frac{2c}{\sqrt{n}}$ & $c \ll \sqrt{n}$

② Discrepancy of a set system

## ② Discrepancy of a set system

- $m$ sets $S_1, \dots, S_m \subseteq \{1, \dots, n\}$
- Want to find a coloring

$$\sigma : \{1, \dots, n\} \longrightarrow \{+1, -1\}$$

s.t. $\max_{i=1}^{m} |\sigma(S_i)|$ is minimized

- Equivalently, $\sigma : \{1, \dots, n\} \longrightarrow \{0, 1\}$

s.t. $\max_{i=1}^{m} \left| \sigma(S_i) - \frac{|S_i|}{2} \right|$ is min.

- By the previous application, even a random coloring, ie, $\sigma(i) = \begin{cases} 1 & \text{wp } 1/2 \\ 0 & \text{" } 1/2 \end{cases}$

gives for every $S_i$ ...

$$\mathbb{P}\left[ \left| \sigma(S_i) - \frac{|S_i|}{2} \right| \geq t\sqrt{|S_i|} \right] \leq e^{-t^2/2}$$

$\therefore \forall i = 1 \dots m$,

$$\mathbb{P}\left[ \left| \sigma(S_i) - \frac{|S_i|}{2} \right| \geq \sqrt{4n \ln m} \right] \leq e^{-2\ln m} = \frac{1}{m^2}$$

$\therefore \Pr\left[ \exists i = 1 \dots m : \left| \sigma(S_i) - \frac{|S_i|}{2} \right| \geq \sqrt{4n\ln m} \right] \leq \frac{1}{m}$

$$\therefore \Pr\left[\exists i = 1 \ldots m : \left|\sigma(S_i) - \frac{|S_i|}{2}\right| \geq \sqrt{4n \ln m}\right] \leq \frac{1}{m}$$

UNION BND

$$\Pr[\exists i : \mathcal{E}_i] \leq \sum_i \Pr[\mathcal{E}_i]$$

$$\boxed{\text{Every set system has a discrepancy of } O\left(\sqrt{n \ln m}\right)}$$

---

## Minimizing Congestion in Directed Ntworks

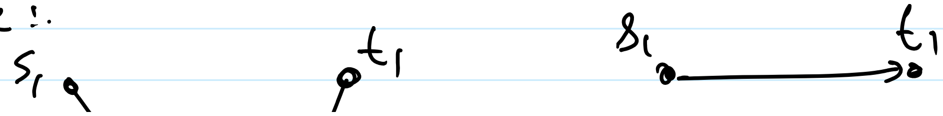
**Input:**
- Directed Graph $G = (V, A)$
- Source sink pairs $\{(s_1, t_1), (s_2, t_2) \ldots (s_k, t_k)\}$

**Output:** Paths $p_1, p_2, \ldots, p_k$ where $p_i$ goes from $s_i \rightsquigarrow t_i$

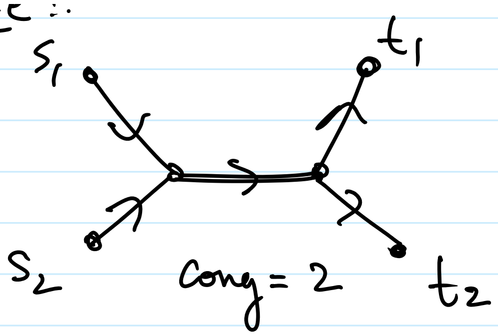
**Obj:** Minimize $\max_e \underbrace{\text{congestion}(e)}_{}$

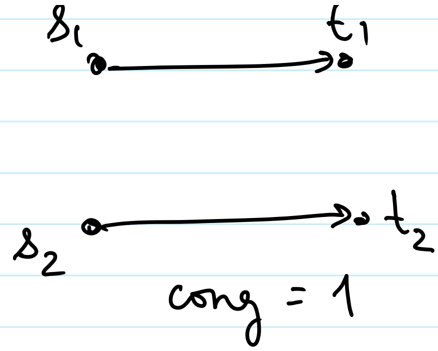
$\equiv$ # of $P_i$'s containing $e$

**Example:**



Example:




## LP relaxation

- $\mathcal{P}_i :=$ set of paths from $s_i \rightsquigarrow t_i$
- Variable : $x_p$

$$\text{Min } C$$

$$\forall e : \sum_{i=1}^{k} \sum_{p \in \mathcal{P}_i : e \in p} x_p \leq C$$

$$\forall i : \sum_{p \in \mathcal{P}_i} x_p = 1$$

- $x_p \in [0,1]$
- $C \geq 1$

As described, this LP has exponentially many variables.

For now, let's assume we can solve the LP & that $x_p > 0$ for only polynomially many $p$'s.

## Algo:

For each $i = 1 \ldots k$,

Sample one $p \in \mathcal{P}_i$ w.p. $x_p$

## Analysis

- Fix edge $e$

- Fix edge $e$
- Define $X_{e,i} = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ path contains } e \\ 0 & \text{o/w} \end{cases}$

- $\text{Cong}(e) = \sum_{i=1}^{k} X_{e,i}$

- $ALG = \max_{e} \text{Cong}(e)$

- $$\mathbb{E}[\text{Cong}(e)] = \sum_{i=1}^{k} \mathbb{E}[X_{e,i}]$$

$$= \sum_{i=1}^{k} \sum_{p \in \mathcal{P}_i : e \in p} x_p \leq LP$$

- Also <u>note</u> : $X_{e,1}, X_{e,2}, \ldots, X_{e,k}$'s are <u>independent</u>.

Just to contrast,

$X_{e,i}$ & $X_{e',i}$ for the same $i$ are <u>NOT</u> independent.

- $$\therefore \mathbb{P}\left[\text{Cong}(e) \geq (1+\delta) LP\right] \leq e^{-\frac{LP \cdot \delta \ln \delta}{2}}$$

(for $\delta \geq 10$) ... using the generalization

if $\dfrac{LP \cdot \delta \ln \delta}{\ } \geq 2 \ln m$, then if $\frac{LP \cdot \delta \ln \delta}{2} \geq 2 \ln m$, then

$$\mathbb{P}\left[\exists e : \text{Cong}(e) \geq (1+\delta) LP\right] \leq \frac{1}{m}$$

$\Rightarrow$ whp, $\text{Cong}(e) \leq (1+\delta) LP$, $\forall e$.

$\therefore$ if $\delta \geq \frac{8 \ln m}{\ln \ln m}$, this is indeed true.

$\because \ln \delta \geq \frac{\ln \ln m}{2}$
$\underline{\&\ LP \geq 1}$

$\therefore$ The Randomized Algorithm is an $O\left(\frac{\ln m}{\ln \ln m}\right)$-factor appx algo.

Note:

If LP turns out to be "large", i.e., $LP \geq \ln m$, then $\delta$ can be chosen to be $\Theta(1)$, & there is a constant factor approximation.

## Solving the LP

One way to solve the LP is to write an LP which has variables $f_{i,e}$ $\forall i = 1 .. k$, $\forall e \in A$ s.t for all $i$, the $f_{i,e}$'s constitute an <u>unit</u> flow from $s_i \rightsquigarrow t_i$.

Given $f_{i,e}$'s, we can perform FLOW DECOMPOSITION to get $x_p$'s satisfying the LP-constraint above. This would be the "easier" way to do things.

In class, we looked at a different way (and I missed a crucial step). Here is the full proof.

$$\min \quad C$$
$$\forall i = 1..k: \quad \sum_{p \in P_i} x_p = 1$$
$$\forall e: \quad C - \sum_i \sum_{p \in P_i : e \in p} x_p \geq 0$$

$$\max \quad \sum_e \beta_e$$
$$\sum_{e \in E} \beta_e \leq 1$$
$$\forall i: \forall p \in P_i \quad \sum_{e \in p} \beta_e - \alpha_i \geq 0$$
$$\beta \geq 0, \ \alpha \text{ - free.}$$

* The dual has a separation oracle. Given any $(\alpha, \beta)$, since $\beta \geq 0$, we can use Dijkstra's shortest path algorithm to separate.

$\therefore$ We can obtain $(\alpha^*, \beta^*)$ a soln to the DUAL LP.

* By complementary slackness, we know that

$\forall p \in P_i: \quad x_p^* > 0 \Rightarrow \sum \beta_e^* = \alpha_i^*$ must be true.

$\forall p \in \mathcal{P}_i : x_p^* > 0 \implies \sum_{e \in p} \beta_e^* = \alpha_i^*$ must be true.

$\therefore$ In the primal LP, we may only keep the variables $\{x_p : p \in \mathcal{P}_i\}$ only for the paths $p$ st $\beta^*(p) = \alpha_i$.

I missed this in class.

But (a) even this could be exponentially many
(b) we don't straightaway know how to enumerate all shortest paths.

Resolution of (a) is that we only care about "linearly independent" collection of paths (and # of cols is $\leq$ # edges); resolution of (b) is less obvious — how to find a linearly independent coll^n of shortest paths. Sounds like an interesting problem.

* Instead, let me give a more general resolution.

| PRIMAL | $\longrightarrow$ | DUAL |
| --- | --- | --- |
| exp many vars | | poly many vars |
| poly " constr. | | exp. many constr. |
| | | But Sep. oracle. |

What does ellipsoid do? It calls the sep. oracle atmost $T \equiv \text{poly}(n)$ times to give us the solution to the dual.

Let the sep. hyperplanes of the dual returned by the S.O. (in our case the shortest path algo) be $\{a_1^T x \geq b_1, a_2^T x \geq b_2, \ldots, a_T^T x \geq b_T\}$

Look @ the new LP

DUAL'

DUAL'

poly many vars

$a_1^T x \geq b_1$
$a_2^T x \geq b_2$
$\vdots$
$a_T^T x \geq b_T$

only the sep. hyperplanes that we encountered in ellipsoid.

DUAL' $\geq$ DUAL

$\because$ fewer constraints

For our problem, DUAL' would look like

$$\max \sum_{e \in E} \beta_e$$

$$-\sum \beta_e \leq 1$$

$$\sum_{e \in p} \beta_e \geq \alpha_1 \quad \text{for some } p\text{'s in } P_i$$

$$\text{total \# of } P_i\text{'s} \leq T = \text{poly}(m)$$

MAIN OBS :- If we run ellipsoid on DUAL', then since Ellipsoid is deterministic we will have the SAME BEHAVIOUR on DUAL' as in DUAL.

Whenever we asked for a sep. oracle in DUAL and returned $a_i^T x \geq b_i$, say, the same sep. hyp is present in DUAL' as well!

$\therefore \quad \text{DUAL} = \text{DUAL}'$

$\Rightarrow \quad \text{PRIMAL} = \underbrace{\text{PRIMAL}'}$

The only variables that survive are the $T \equiv \text{poly}(m)$ variables $x_p$.