Lecture 5

Continue Chernoff-Hoeffding Bounds

An Application to Routing for Congestion Minimization

A classical problem from both theory and practice.

Given $G = (V, E)$ a directed graph. Let $(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)$ be source-sink pairs.

The Edge-Disjoint-Paths (EDP) problem

asks the following: Can we find paths $P_1, P_2, \ldots, P_k$ such that

(i) P_i is an $s_i - t_i$ path

(ii) $P_1, \ldots, P_k$ are edge-disjoint

A fundamental but difficult problem. We will consider a relaxation.

Given G and the pairs, find the paths such that no edge is used in too many paths.

We write an LP relaxation

Let $\mathcal{P}_i$ be the set of all $s_i \to t_i$ paths - an exponential set.

$$\begin{aligned}
\min \quad & \lambda & \\
\text{s.t.} \quad & \sum_{p \in \mathcal{P}_i} x_p = 1 & \forall i \in [k] \\
& \sum_{p \in \mathcal{P}_i, p \ni e} x_p \leq \lambda & \forall e \in E \\
& x_p \geq 0 & p \in \bigcup_i \mathcal{P}_i
\end{aligned}$$

We will not discuss how to solve the above LP but it can be done via the Ellipsoid method or by writing a different edge-flow based formulation that I write below where $x(e, i)$ is the flow on edge e for pair i.

$$\begin{aligned}
\min \quad & \lambda & \\
\text{s.t.} \quad & \sum_{e \in \delta^+(s_i)} x(e, i) - \sum_{e \in \delta^-(s_i)} x(e, i) = 1 & \forall i \in [k] \\
& \sum_{e \in \delta^+(v)} x(e, i) - \sum_{e \in \delta^-(v)} x(e, i) = 0 & \forall v \notin \{s_i, t_i\}, \forall i \in [k] \\
& \sum_{i=1}^{k} x(e, i) \leq \lambda & \forall e \in E \\
& x(e, i) \geq 0 & e \in E, i \in [k]
\end{aligned}$$

Let λ^* be the optimum value of the LP relaxation. Note that λ^* can be smaller than 1. The true lower bound is $\max\{\lambda^*, 1\}$.

Can we convert the fractional solution to an integral solution with small "congestion"?

Rounding

1. Solve LP relaxation.

2. For each pair (s_i, t_i) independently pick a path P_i where p_i is chosen with probability $\bar{x}_{p_i}$.

Raghavan-Thompson were the first to do this rounding and analyze via Chernoff bounds in 1987.

Theorem. *The algorithm outputs a set of random paths $P_1, \ldots, P_k$ such that $\forall i \in [k]$, P_i is an $s_i - t_i$ path. And with probability $1 - \frac{1}{poly(m)}$ the load on any edge is at most $O(\frac{\log m}{\log \log m}) \max\{\lambda^*, 1\}$.*

Proof Sketch: Consider any edge e. Let Y_i be the indicator for P_i, the path chosen for (s_i, t_i), using e. Let $Y = \sum_{i=1}^{k} Y_i$.

$$P_r[Y_i = 1] = \sum_{p \in \mathcal{P}_i, p \ni e} x_p = x(e,i) \text{ the total flow on e for commodity i.}$$

Why? $Y_1, \ldots, Y_k$ are independent.

$$E[Y] = \sum E[Y_i] = \sum_{i=1}^{k} x(e,i) \leq \lambda^*$$

Hence by Chernoff bound, then

$$P_r[Y \geq c\frac{\log(m)}{\log\log(m)}] \leq \frac{1}{m^{c'}}$$

for sufficiently large constants c and c'. We apply union bound over all the m edges and we have the load on any edge

$$\cdots \leq m \cdot \frac{1}{m^{c'}} \leq \frac{1}{m^{c'-1}} \cdot m$$

Exercise: You can use Chernoff bounds to prove two related bounds

(i) Show that for $\epsilon \in (0,1)$, $P_r[Y \geq (1+\epsilon)\lambda^* + \frac{c \log m}{\epsilon^2}] \leq \frac{1}{m^{c'}}$. Thus when $\lambda^* = \Omega(\log m)$ we get a very good approximation.

(ii) when $\lambda^* \geq c \log m$, $P_r[Y > \lambda^* \pm \sqrt{c \log m \lambda^*}] \leq \frac{1}{m^{c'}}$. Note the preceding bound has a multiplicative factor on λ^*.

Additive Chernoff Bound

To motivate this bound consider the random walk on the line. We had $Y = \sum_{i=1}^{n} X_i$ where $X_i \in \{-1, 1\}$ and $E[X_i] = 0$ and hence $E[Y] = 0$. In this setting we cannot expect a multiplicative bound. We will state a general bound that handles this kind of setting.

Hoeffding Bound

Let $X = \sum_{i=1}^{n} X_i$ where

(i) X_i are independent

(ii) $X_i \in [a_i, b_i] \quad \forall i \in [n]$

(iii) $E[X_i] = 0 \quad \forall i \in [n]$

Then,

$$P_r[X \geq \eta] \leq e^{-\frac{\eta^2}{2\sum_i (b_i - a_i)^2}}$$

and

$$P_r[X \leq -\eta] \leq e^{-\frac{\eta^2}{2\sum_i (b_i - a_i)^2}}$$

Comments:

(i) Suppose $X_i \in [-1, 1]$. Then $\sum_{i=1}^{n} (b_i - a_i)^2 = 4n$.

(ii) Why assume $E[X_i] = 0$? We can replace X_i by $Y_i = X_i - E[X_i]$ and $E[Y_i] = 0$ and if $X_i \in [a_i, b_i]$ then $Y_i \in [a_i - E[X_i], b_i - E[X_i]]$ and hence the term $(b_i - a_i)$ does not change.

Without assuming $E[\cdot] = 0$ we have

$$P_r[X - E[X] \geq t] \leq e^{-\frac{t^2}{2\sum_{i=1}^{n}(b_i - a_i)^2}}$$

which is the standard form.

Proof Sketch: As before we consider e^{tX} for a parameter $t > 0$.

$$P_r[X \geq \eta] = P_r[e^{tX} \geq e^{t\eta}] \leq \frac{E[e^{tX}]}{e^{t\eta}} \quad \text{by Markov}$$

So it boils down to estimating/upper bounding $E[e^{tX}] = \prod_{i=1}^{n} E[e^{tX_i}]$ and choosing the best t.

How do we bound $E[e^{tX_i}]$? We will only sketch the argument. e^{ty} is a convex function in the interval $[a_i, b_i]$ (on the entire real line). We know $E[X_i] = 0$ and $X_i \in [a_i, b_i]$. What is the distribution that maximizes $E[e^{tX_i}]$ since that is what gives us the weakest bound?

Due to convexity it turns out that we should put all the probability mass on the extremes of the interval $[a_i, b_i]$ (i.e. the worst case is when $X_i \in \{a_i, b_i\}$ subject to $E[X_i] = 0$; note $a_i \leq 0$ and $b_i \geq 0$). Say X_i is b_i with $p = \frac{-a_i}{b_i - a_i}$ and a_i with prob $1 - p = \frac{b_i}{b_i - a_i}$. One can prove above by convexity. Assuming above,

$$E[e^{tX_i}] \leq pe^{tb_i} + (1-p)e^{ta_i}$$

(completely obvious). By calculus we can show that this is $\leq e^{t^2 \frac{(b_i - a_i)^2}{8}}$.

Assuming above we have

$$E[e^{tX}] \leq e^{\frac{t^2}{8}\sum_{i=1}^{n}(b_i - a_i)^2}$$

Thus $P_r[X \geq \eta] \leq \frac{E[e^{tX}]}{e^{t\eta}} \leq e^{\frac{t^2 \sum(b_i - a_i)^2}{8} - t\eta}$.

Minimizing over t, we have

$$\frac{t}{4}\sum_{i=1}^{n}(b_i - a_i)^2 = \eta$$

$$t^* = \frac{4\eta}{\sum_{i=1}^{n}(b_i - a_i)^2}$$

plugging in, the exponent is

$$\begin{aligned} &\frac{t^{*2}}{8}\sum(b_i - a_i)^2 - t^*\eta \\ =&\frac{16\eta^2}{8(\sum(b_i - a_i)^2)} - \frac{4\eta^2}{\sum(b_i - a_i)^2} \\ =&-\frac{2\eta^2}{\sum_{i=1}^{n}(b_i - a_i)^2} \end{aligned}$$

Thus

$$P_r[X \geq \eta] \leq e^{-\frac{2\eta^2}{\sum_{i=1}^{n}(b_i - a_i)^2}}$$

Lower Tail is similar.

Application

Random Walk on the Line. $X = \sum_{i=1}^{n} X_i$, $X_i \in \{-1, 1\}$, $E[X_i] = 0$. $b_i = 1, a_i = -1$.

$$P_r[X > t\sqrt{n}] \leq e^{-\frac{t^2 n}{8n}} \leq e^{-t^2/8}$$

$$P_r[X < t\sqrt{n}] \leq e^{-t^2/8}$$