

## Lecture 03: Chernoff Bounds and Intro to Spectral Graph Theory

Lecturer: Yuan Zhou

Scribe: Zhenhua Chen

## 1 Chernoff Bounds

In the last lecture, we proved an upper bound for a random variable in a particular case, namely:

$$\Pr[e^{\lambda S} \geq e^{10\lambda\sqrt{n \ln n}}] = e^{\frac{n\lambda^2}{2} - 10\lambda\sqrt{n \ln n}}.$$

When picking  $\lambda = 10\sqrt{\frac{\ln n}{n}}$ , the upper bound becomes:

$$\Pr[e^{\lambda S} \geq e^{10\lambda\sqrt{n \ln n}}] = e^{50 \ln n - 100 \ln n} = \frac{1}{n^{50}}.$$

For this lecture, we will use a similar method to give a upper bound for a more general case.

**Theorem 1** (Chernoff Bounds). *Suppose  $0 \leq X_i \leq 1$  be independent Bernoulli variables.  $\mathbb{E}(X_i) = P_i$ . Let  $X = X_1 + X_2 + \dots + X_n$ ,  $\mu = \mathbb{E}[X]$ , then we have:*

$$\Pr[X \geq (1 + \delta)\mu] \leq e^{\frac{-\delta^2\mu}{2+\delta}} (\delta > 0), \quad \text{and} \quad \Pr[X \leq (1 - \delta)\mu] \leq e^{\frac{-\delta^2\mu}{2}} (0 < \delta < 1).$$

*Proof of the first inequality [2]:* Fix  $\lambda > 0$ , then  $e^{\lambda x}$  is strictly increasing. Based on Markov Inequality we have:

$$\Pr[X \geq (1 + \delta)\mu] = \Pr[e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}] \leq \frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda(1+\delta)\mu}}.$$

Where

$$\mathbb{E}[e^{\lambda X}] = \prod_{i=1}^n \mathbb{E}[e^{\lambda X_i}] = \prod_{i=1}^n (P_i e^{\lambda} + (1 - P_i)) = \prod_{i=1}^n (1 + P_i(e^{\lambda} - 1)).$$

Note that  $1 + x \leq e^x, x > 0$ , so

$$\prod_{i=1}^n (1 + P_i(e^{\lambda} - 1)) \leq \prod_{i=1}^n e^{P_i(e^{\lambda} - 1)} = e^{(\sum_{i=1}^n P_i)(e^{\lambda} - 1)} \leq e^{\mu(e^{\lambda} - 1)}.$$

Thus,

$$\Pr[X \geq (1 + \delta)\mu] \leq \frac{e^{\mu(e^\lambda - 1)}}{e^{\lambda(1 + \delta)\mu}} = e^{\mu(e^\lambda - 1 - \lambda(1 + \delta))}.$$

Pick  $\lambda = \ln(1 + \delta)$ , then,

$$\Pr[X \geq (1 + \delta)\mu] \leq e^{\mu[(1 + \delta) - 1 - (1 + \delta)\ln(1 + \delta)]} = \left[ \frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right]^\mu.$$

Taking natural logarithm on the right side, we can get:

$$\mu(\delta - (1 + \delta)\ln(1 + \delta)).$$

Note that  $\ln(1 + x) \geq \frac{x}{1 + x/2}$  when  $x > 0$ , we obtain:

$$\mu(\delta - (1 + \delta)\ln(1 + \delta)) \leq -\frac{\delta^2}{2 + \delta}\mu.$$

Therefore we can get the upper bound:

$$\Pr[X \geq (1 + \delta)\mu] \leq \left[ \frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right]^\mu \leq e^{-\frac{\delta^2}{2 + \delta}\mu}.$$

Similarly we can prove the other inequality as follows.

*Proof of the second inequality:* Fix  $0 < \lambda < 1$ , then  $e^{-\lambda x}$  is strictly decreasing. Based on Markov Inequality we have:

$$\Pr[X \leq (1 - \delta)\mu] = \Pr[e^{-\lambda X} \geq e^{-\lambda(1 - \delta)\mu}] \leq \frac{\mathbb{E}[e^{-\lambda X}]}{e^{-\lambda(1 - \delta)\mu}}.$$

Where

$$\mathbb{E}[e^{-\lambda X}] = \prod_{i=1}^n \mathbb{E}[e^{-\lambda X_i}] = \prod_{i=1}^n (P_i e^{-\lambda} + (1 - P_i)) = \prod_{i=1}^n (1 - P_i(1 - e^{-\lambda})).$$

Note that  $1 - x \leq e^{-x}$ , so

$$\prod_{i=1}^n (1 - P_i(1 - e^{-\lambda})) \leq \prod_{i=1}^n e^{-P_i(1 - e^{-\lambda})} = e^{-(\sum_{i=1}^n P_i)(1 - e^{-\lambda})} \leq e^{\mu(e^{-\lambda} - 1)}.$$

Thus,

$$\Pr[X \leq (1 - \delta)\mu] \leq \frac{e^{\mu(e^{-\lambda} - 1)}}{e^{-\lambda(1 - \delta)\mu}} = e^{\mu(e^{-\lambda} - 1 + \lambda(1 - \delta))}.$$

Pick  $\lambda = -\ln(1 - \delta)$ , then,

$$\Pr[X \leq (1 - \delta)\mu] \leq e^{\mu[(1 - \delta) - 1 - (1 - \delta)\ln(1 - \delta)]} = \left[ \frac{e^{-\delta}}{(1 - \delta)^{1 - \delta}} \right]^\mu.$$

Note that  $(1 - x)\ln(1 - x) > -x + \frac{x^2}{2}$  when  $0 < x < 1$ , therefore:

$$\Pr[X \leq (1 - \delta)\mu] \leq \left[ \frac{e^{-\delta}}{(1 - \delta)^{1 - \delta}} \right]^\mu \leq \left[ \frac{e^{-\delta}}{e^{-\delta + \frac{\delta^2}{2}}} \right]^\mu = e^{-\mu\delta^2/2}.$$

## 1.1 Hoeffding's Inequality

**Theorem 2.** *Hoeffding bound.* Let  $X_1, X_2, \dots, X_N$  be independent variables. Assume  $X_i$  is strictly bounded by  $[a_i, b_i]$ . Let  $X = \sum X_i$  and  $\mu = \mathbb{E}(X)$ , for all  $t > 0$ ,  $\Pr[x \geq \mu + t] \leq e^{-\frac{2t^2}{\sum(b_i - a_i)^2}}$ , at the same time,  $\Pr[x \leq \mu - t] \leq e^{-\frac{2t^2}{\sum(b_i - a_i)^2}}$ .

**Remark 1.** *Hoeffding's inequality applies more general cases when the independent random variables  $X_i$  have different ranges. However Chernoff bounds are tighter in some special situation (e.g. when  $\mu$  is very small).*

## 2 Spectral Graph Theory

In mathematics, spectral graph theory is the study of properties of a graph in relationship to the characteristic polynomial, eigenvalues, and eigenvectors of matrices associated to the graph, such as its adjacency matrix or Laplacian matrix [1]. Let us begin with some basics of spectral graph theory.

### 2.1 Graph

The graph we talk here can be defined as the set of vertexes and edges, namely,  $G = (V, E)$ . We assume the graph has some particular properties:

- Undirected. Being undirected is crucial. Most of the properties we will prove do not extend to directed graphs.
- Finite.
- Okay with multiple edges, weighted edges, or self loops.
- Regular. It is not necessary but will make our life easier.

In case you are familiar with the concept of regular graph, we give the definition here:

**Definition 3.** *If every vertex of a graph has the same number of degree, then this graph is a regular graph. If the degree number is  $d$ , then we call it  $d$ -regular graph.*

For example, Figure 1 is a 5-degree regular graph, we can model this graph by adjacent matrix  $A_G \in \mathbb{R}^{n \times n}$ ,  $n = |V|$ . We claim  $A_G(u, v) = \text{edges between } (u, v)$ , then the  $A_G$  for

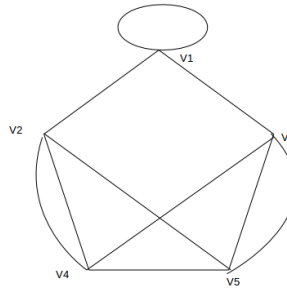


Figure 1: 5-degree regular graph

Figure 1 is:

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 \\ 0 & 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \end{bmatrix}$$

Note that we can conclude from  $A_G$ : 1).  $A_G$  is symmetric since  $G$  is undirected; 2). Sum of rows and columns equal to the graph degree.

## 2.2 Random walk matrix

Given a graph, imagine we need to model the probabilities of stepping from one vertex to another vertex, assume that each node needs to be reached with equal probability, then how should we do this? One straightforward way is to put the probabilities of each option into an adjacent-matrix-style matrix  $W_G$ , each element in this matrix represent a probability of an edge being chosen, namely:

**Definition 4.**  $W_G = \frac{1}{d} \times A_G$ ,  $d$  is the degree of the associated graph.

The  $W_G$  here is called **random walk matrix**. You can find the definition of  $A_G$  in section 2.1.

**Example:** Consider the graph of Figure 1 in section 2.1, we can get the random walk matrix from it by computing its weights. The weight of each edge  $w_d = \frac{1}{d}A_G$ , so the whole matrix is:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \end{bmatrix}$$

Imagine at time  $t_0$ , there exists 1 unit gas at node  $v_3$ . At  $t_1$ ,  $\frac{1}{4}$  gas at node  $v_3$  flows to  $v_1$  (since the probability of this path being chosen equals to  $\frac{1}{4}$ ). Similarly, another  $\frac{1}{4}$  flows to  $v_4$  and the remaining  $\frac{1}{2}$  flows to node  $v_5$ .

Assume  $\vec{p} \in \mathbb{R}^v$ ,  $\vec{p}^{(t)}$  represents the gas distribution at time  $t$ . Now how to compute the gas distribution at time  $t + 1$ ? Assume the initial gas distribution starts with:

$$\vec{p}^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

then,

$$\vec{p}^{(1)} = W_G \cdot \vec{p}^{(0)} = \begin{bmatrix} \frac{1}{4} \\ 0 \\ 0 \\ \frac{1}{4} \\ \frac{1}{2} \end{bmatrix}.$$

Similarly,  $\vec{p}^{(2)} = W_G \cdot \vec{p}^{(1)}$ .

## 2.3 The Laplacian Matrix

Given an undirected graph matrix  $W_G$ , the normalized Laplacian matrix  $L_G$  can be defined as:

**Definition 5.**  $L_G = I - W_G$ , where  $I$  is the identity matrix.

Fix  $\vec{x} \in \mathbb{R}^v$ , the quadratic form associated with  $L_G$ , then:

$$\mathcal{E}(\vec{X}) = \vec{X}^T L_G \vec{X} = \frac{1}{d} \sum_{u,v} (\vec{X}(u) - \vec{X}(v))^2.$$

Based on above deduction, we can conclude that  $L_G$  is positive semi-definite (PSD).

## 2.4 Review of eigenvalues and eigenvectors

**Definition 6.**  $\vec{X}$  is an eigenvector of  $M$  with eigenvalue  $\lambda$  if  $M\vec{X} = \lambda\vec{X}$ .

**Fact:**  $\lambda$  is an eigenvalue of  $M$  if  $\det(M - \lambda I) = 0$ ,  $I$  is the identity matrix.

Let us recall the spectral theorem from linear algebra:

**Theorem 7.** *The spectral theorem. If  $M \in \mathbb{R}^{n \times n}$  is symmetric, then there  $\exists \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$   $n$  mutually orthogonal with vectors  $\vec{\psi}_1, \vec{\psi}_2, \dots, \vec{\psi}_n$  where  $\forall i$ ,  $\vec{\psi}_i$  is an eigenvector of  $M$  with eigenvalue  $\lambda_i$ .*

Moreover,

$$M = [\vec{\psi}_1, \vec{\psi}_2, \dots, \vec{\psi}_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{\psi}_1^T \\ \vec{\psi}_2^T \\ \dots \\ \vec{\psi}_n^T \end{bmatrix} = \sum_{i=1}^n \lambda_i \vec{\psi}_i \vec{\psi}_i^T.$$

**Observation:** Since  $W_G$  is real symmetric, by the spectral theorem, we can write it as:

$$W_G = \lambda_1 \vec{\psi}_1 \vec{\psi}_1^T + \dots + \lambda_n \vec{\psi}_n \vec{\psi}_n^T.$$

Assume:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n,$$

By  $L_G = I - W_G$ , we have:

$$L_G = (1 - \lambda_1) \vec{\psi}_1 \vec{\psi}_1^T + \dots + (1 - \lambda_n) \vec{\psi}_n \vec{\psi}_n^T.$$

Note that we already proved that  $L_G$  is PSD, therefore,

$$1 - \lambda_1 \geq 0.$$

Thus we have

$$1 \geq \lambda_1 \geq \dots \geq \lambda_n.$$

## References

- [1] Spielman, Daniel. *Spectral Graph Theory*. 2011.
- [2] Michel Goemans. *Chernoff bounds, and some applications*. 2015. Available online at <http://math.mit.edu/~goemans/18310S15/chernoff-notes.pdf>