CS : Deep Learning Fall 2020

Lecture 2 — September 4, 2020

Prof. Zhao Song *Scribe: Zhao Song*

1 Concentration vs Discrepancy

1.1 Chernoff bound vs Spencer Theorem

Using the well known Chernoff bound and a union bound, it is easy to derive the following result:

Theorem 1. *Given n vectors $a_1, a_2, \cdots, a_n \in \{\pm\}^n$. Let $\sigma_1, \sigma_2, \cdots, \sigma_n$ denote n i.i.d. random sign variables. Then we have*

$$\Pr_{\sigma \sim \{\pm 1\}} \left[\forall j \in [n], |\langle a_j, \sigma \rangle| \lesssim \sqrt{n \log n} \right] \geq 1 - 1/\operatorname{poly}(n).$$

The above result can be thought of as a concentration inequality. An interesting question to ask is if we don't need to high probability, then is $\sqrt{n \log n}$ is the best bound we could hope. This is a discrepancy type question.

Spencer is able to show the following exciting result

Theorem 2. *Given n vectors $a_1, a_2, \cdots, a_n \in \{\pm\}^n$. Let $\sigma_1, \sigma_2, \cdots, \sigma_n$ denote n i.i.d. random sign variables. Then we have*

$$\Pr_{\sigma \sim \{\pm 1\}} \left[\forall j \in [n], |\langle a_j, \sigma \rangle| \lesssim \sqrt{n} \right] \geq 1/2^n.$$