AI4OPT Tutorial Series: Practice problems

November 15, 2022

1. Let $\mathbf{W} \in \mathbb{R}^{n \times n} = (W_{ij})$, $\mathbf{W} = \mathbf{W}^\top$, $\{W_{ij} : i < j\} \sim \mathcal{N}(0, 1/n)$, $\{W_{ii} : 1 \leq i \leq n\} \sim \mathcal{N}(0, 2/n)$ be independent. The behavior of the largest eigenvalue of $\mathbf{W}$ is crucial in many applications. A fundamental result from Random Matrix Theory establishes that $\lambda_{\max}(\mathbf{W}) \to 2$ a.s. Here we will prove a one-sided upper bound.

   To this end, we will use the Sudakov-Fernique inequality for gaussians.

   Lemma 1 (Sudakov-Fernique inequality)**.** *Let X and Y be n-dimensional Gaussian vectors with $\mathbb{E}[X_i] = \mathbb{E}[Y_i]$ for all $1 \leq i \leq n$ and $\mathbb{E}(X_i - X_j)^2 \leq \mathbb{E}(Y_i - Y_j)^2$ for all $i \neq j$. Then $\mathbb{E}[\max_{1 \leq i \leq n} X_i] \leq \mathbb{E}[\max_{1 \leq i \leq n} Y_i]$.*

   (i) Use the Sudakov-Fernique inequality to prove that $\limsup_{n \to \infty} \mathbb{E}[\lambda_{\max}(\mathbf{W})] \leq 2$. **Hint:** Use $\lambda_{\max}(\mathbf{W}) = \max_{\|x\|_2 = 1} x^\top \mathbf{W} x$. Compare with the gaussian process $H(x) = 2g^\top x$, where $g \sim \mathcal{N}(0, I_n/n)$.

   (ii) Repeat the argument for a spiked matrix $\mathbf{M} = \frac{\lambda}{n} \mathbf{v}\mathbf{v}^\top + \mathbf{W}$ ($\mathbf{v}$ is uniform on the sphere with radius $\sqrt{n}$) to prove the upper bound

   $$\mathbb{E}[\lambda_{\max}(\mathbf{W})] \leq \begin{cases} 2 & \text{if } \lambda \leq 1 \\ \lambda + \frac{1}{\lambda} & \text{o.w.} \end{cases}$$

2. In this exercise, we will sketch some arguments to establish *exact recovery* for community detection using semidefinite programming. Let $G = ([n], E)$ be sampled from a Stochastic Block Model (SBM) with two communities as follows: let $\sigma \in \{\pm 1\}^n$ denote the true community assignment, and for simplicity assume that $\sum_i \sigma_i = 0$. The edges are sampled independently with probability $p_n = \frac{a \log n}{n}$ if $\sigma_i = \sigma_j$ and $q_n = \frac{b \log n}{n}$ o.w. Here a, b are constants independent of n. In this regime, if $a + b > 2$ and $\sqrt{a} - \sqrt{b} > \sqrt{2}$, then the underlying community assignment σ can be *exactly* recovered (upto a global flip). We discuss a semidefinite program which works in this regime.

   Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ denote the adjacency matrix of the graph. Define $\mathbf{B} = (B_{ij}) \in \mathbb{R}^{n \times n}$ as follows:

   $$B_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } A_{ij} = 1 \\ -1 & o.w. \end{cases}$$

Consider the semidefinite program $\max \operatorname{Tr}(BX)$ subject to $X \succeq 0$, $X_{ii} = 1$ for all $1 \leq i \leq n$. We claim that $X_* = \sigma\sigma^\top$ is the unique solution to the SDP with high probability.

(i) To this end, we will use convex duality. First prove that

$$\begin{aligned} &\max \operatorname{Tr}(BX) &&\leq \min \operatorname{Tr}(Z) \\ &X \succeq 0 && Z \text{ is diagonal} \\ &X_{ii} = 1 && Z - B \succeq 0. \end{aligned}$$

Hint: Z is diagonal and $X_{ii} = 1$ implies $\operatorname{Tr}(ZX) = \operatorname{Tr}(X)$. Further, $X \succeq 0$ and $Z - B \succeq 0$ implies $\operatorname{Tr}[(Z - B)X] \geq 0$.

(ii) We wish to prove that X_* is the unique solution to the original SDP. Prove that if there exists a feasible Z for the dual problem such that $\operatorname{Tr}[(Z - B)X_*] = 0$, then X_* is a solution to the primal SDP. In addition, if $\lambda_2(Z - B) > 0$, X_* is the unique solution to the primal SDP (here $\lambda_2(Z - B)$ is the second smallest eigenvalue of $Z - B$).

(iii) Note that $\operatorname{Tr}[(Z - B)X_*] = 0$ if $(Z - B)\sigma_* = 0$. Use this equation to construct a guess for Z. This is called the *dual witness*.

(iv) By construction, Z is diagonal and $(Z - B)\sigma_* = 0$. One has to check that $Z - B \succeq 0$ and $\lambda_2(Z - B) > 0$. These can be done using appropriate concentration inequalities. See [1] and references therein for additional details.

(v) In fact if $\sqrt{a} - \sqrt{b} < \sqrt{2}$, exact recovery is information theoretically impossible. For more discussion on this, see [1] and references therein.

References

[1] Bandeira, A.S., 2015. Ten lectures and forty-two open problems in the mathematics of data science. Lecture Notes.