Advanced Graph Algorithms and Optimization Spring 2021

Spectral Graph Theory

R. Kyng & M. Probst *Problem Set 6 — Tuesday, April 20th*

The exercises for this week will not count toward your grade, but you are highly encouraged to solve them all.

Exercise 1.

In this exercise, we want you to complete the proof of Theorem 9.3.3 in Chapter 9. Refer to the lectures notes for definitions of the terms used here.

1. Prove that Equation (9.4) is satisfied, i.e. that for all edges $e \in E$ we have $\|\boldsymbol{X}_e\| \leq \frac{1}{\alpha}$.
2. Prove that Equation (9.5) is satisfied, i.e. that $\left\|\sum_e \mathbb{E}\left[\boldsymbol{X}_e^2\right]\right\| \leq \frac{1}{\alpha}$.
3. Explain how we can use a scalar Chernoff bound to prove that $\left|\tilde{E}\right| \leq O(\epsilon^{-2}\log(n/\delta)n)$ with probability at least $1-\delta/2$. You may pick any constant that suits you to establish the $O(\cdot)$ bound.

Exercise 2.

In this exercise, you will prove the Lemma 10.2.2 from Chapter 10. We restate it here:

Lemma. *Given a matrix $\boldsymbol{M} \in S_{++}^n$, a vector $\boldsymbol{d} \in \mathbb{R}^n$ and a decomposition $\boldsymbol{M} \approx_\kappa \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{L}}^\top$, we can find $\tilde{\boldsymbol{x}}$ that ϵ-approximately solves $\boldsymbol{M}\boldsymbol{x} = \boldsymbol{d}$, using $O((1+K)\log(K/\epsilon)(T_{matvec} + T_{sol} + n))$ time, where*

- *T_{matvec} denotes the time required to compute $\boldsymbol{M}\boldsymbol{z}$ given a vector $\boldsymbol{z}$, i.e. a "matrix-vector multiplication".*
- *T_{sol} denotes the time required to compute $\boldsymbol{\mathcal{L}}^{-1}\boldsymbol{z}$ or $(\boldsymbol{\mathcal{L}}^\top)^{-1}\boldsymbol{z}$ given a vector $\boldsymbol{z}$.*

The lemma uses our definition of ϵ-approximate solution, we will also restate:

Definition (ϵ-approximate solution to $\boldsymbol{M}\boldsymbol{x} = \boldsymbol{d}$.)**.** *Given PSD matrix $\boldsymbol{M}$ and $\boldsymbol{d} \in \ker(\boldsymbol{M})^\perp$, let $\boldsymbol{M}\boldsymbol{x}^* = \boldsymbol{d}$. We say that $\tilde{\boldsymbol{x}}$ is an ϵ-approximate solution to the linear equation $\boldsymbol{M}\boldsymbol{x} = \boldsymbol{d}$ if*

$$\|\tilde{\boldsymbol{x}} - \boldsymbol{x}^*\|_{\boldsymbol{M}}^2 \leq \epsilon \|\boldsymbol{x}^*\|_{\boldsymbol{M}}^2 .$$

You may assume the following theorem.

Theorem (Accelerated Gradient Descent for Solving PD Linear Equations)**.** *Suppose we are given matrix a $\boldsymbol{A} \in S_{++}^n$ and a vector $\boldsymbol{b} \in \mathbb{R}^n$, and l and u s.t.*

$$l \leq \lambda_{\min}(\boldsymbol{A}) \text{ and } \lambda_{\max}(\boldsymbol{A}) \leq u.$$

Let $\kappa = \frac{u}{l}$. We can find $\tilde{\boldsymbol{x}}$ that ϵ-approximately solves $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, in time $O(\sqrt{\kappa}\log(\kappa/\epsilon)(T_{matvec}+n))$ where T_{matvec} denotes the time required to compute $\boldsymbol{A}\boldsymbol{z}$ given a vector $\boldsymbol{z}$, i.e. a "matrix-vector multiplication".

Here are some intermediate steps that might be helpful for proving the lemma:

1. Show that for all $\boldsymbol{x}$,
$$\frac{1}{1+K} \leq \frac{\boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{L}}^\top \boldsymbol{x}} \leq 1+K$$
2. Show that for all $\boldsymbol{y}$,
$$\frac{1}{1+K} \leq \frac{\boldsymbol{y}^\top \boldsymbol{\mathcal{L}}^{-1}\boldsymbol{M}(\boldsymbol{\mathcal{L}}^\top)^{-1}\boldsymbol{y}}{\boldsymbol{y}^\top \boldsymbol{y}} \leq 1+K$$
3. It might be a good idea to approximately solve a linear equation in $\boldsymbol{\mathcal{L}}^{-1}\boldsymbol{M}(\boldsymbol{\mathcal{L}}^\top)^{-1}$? You'll have to figure out the right way to convert both a linear equation and a solution.

Exercise 3.

Consider $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{R}^n$.

1. Assume that $\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top$ is invertible. Determine c such that
$$\left(\boldsymbol{I} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{I} - \frac{\boldsymbol{u}\boldsymbol{v}^\top}{c}.$$
2. Assume that both $\boldsymbol{A}$ and $\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top$ are invertible. Prove that
$$\left(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top \boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^\top \boldsymbol{A}^{-1}\boldsymbol{u}}.$$

   Hint: You might use that $(\boldsymbol{B}\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}$ for two invertible matrices $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{n\times n}$.

Exercise 4.

Consider a matrix function $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$. For $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{n\times n}$, we define
$$Df(\boldsymbol{X})[\boldsymbol{Y}] = \frac{\partial}{\partial t}\bigg|_{t=0} f(\boldsymbol{X} + t\boldsymbol{Y}).$$

Remark. *Note that if we think of $\boldsymbol{X}$ and $\boldsymbol{Y}$ each as a vector of numbers, then this is the (matrix-valued) directional derivative of f at $\boldsymbol{X}$ in the direction of $\boldsymbol{Y}$.*

Consider $f(\boldsymbol{X}) = \boldsymbol{X}^{-1}$ for an invertible matrix $\boldsymbol{X} \in \mathbb{R}^{n\times n}$. Prove that
$$Df(\boldsymbol{X})[\boldsymbol{Y}] = -\boldsymbol{X}^{-1}\boldsymbol{Y}\boldsymbol{X}^{-1}.$$

Hint: You might need to use Exercise 3.

Exercise 5.

1. Consider $\boldsymbol{A} \in S_{++}^n$ and matrix $\boldsymbol{\Delta} \in S_{+}^n$. Prove that $(\boldsymbol{A} + \boldsymbol{\Delta})^{-1} \preceq \boldsymbol{A}^{-1}$.

2. Let T be a convex set. We say that a function $f : T \to \mathbb{R}^{n \times n}$, is operator convex if for any two matrices $\boldsymbol{A}, \boldsymbol{B} \in T$ and any $\theta \in [0, 1]$

$$f\left(\theta \boldsymbol{X} + (1 - \theta)\, \boldsymbol{Y}\right) \preceq \theta f\left(\boldsymbol{X}\right) + (1 - \theta)\, f\left(\boldsymbol{Y}\right).$$

   Prove that $f(\boldsymbol{X}) = \boldsymbol{X}^{-1}$ is operator convex over the set $T = S_{++}^n$.

   Hint: You could first show that operator convexity is implied by the second directional derivative $D^2 f(\boldsymbol{X})[\boldsymbol{Y}, \boldsymbol{Y}]$ being positive semi-definite for all $\boldsymbol{Y} \in S^n$ and $\boldsymbol{X} \in S_{++}^n$.