Advanced Graph Algorithms and Optimization	Spring 2020

Spectral Graph Theory II

Rasmus Kyng — *Problem Set 6 — Wednesday, April 22*

The exercises for this week will not count toward your grade, but you are highly encouraged to solve them all. The solution is due on **Sunday, April 26 by 12:00 noon**.

Throughout these exercises, we will use the following notation:

- S^n is the set of symmetric real matrices $n \times n$ matrices.
- S^n_+ is the set of positive semi-definite $n \times n$ matrices.
- S^n_{++} is the set of positive definite $n \times n$ matrices.

Whenever we say a matrix is positive semi-definite or positive definite, we require it to be real and symmetric.

Exercise 1.

In this exercise, you will prove the following lemma from Lecture 9. 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 $O((1+K)\log(K/\epsilon)(T_{\text{matvec}} + T_{\text{sol}} + n))$ 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, which you proved in Problem 2 of Graded Homework 1.

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 $O(\sqrt{\kappa}\log(\kappa/\epsilon)(T_{\text{matvec}}+n))$ 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 2.

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 $\boldsymbol{X} \in S^n_{++}$.

Exercise 3.

1. Let $f(\boldsymbol{X}) = \boldsymbol{X}^2$ for $\boldsymbol{X} \in S^n$. Prove that f is midpoint operator convex, that is, for any $\boldsymbol{X}, \boldsymbol{Y} \in S^n$ we have
$$f\left(\frac{1}{2}\boldsymbol{X} + \frac{1}{2}\boldsymbol{Y}\right) \preceq \frac{1}{2} f\left(\boldsymbol{X}\right) + \frac{1}{2} f\left(\boldsymbol{Y}\right).$$
Remark. You're not asked to prove anything beyond midpoint operator convexity, but under very mild conditions midpoint convexity implies general convexity, and midpoint operator convexity implies operator convexity. This is also true in this case: f is in fact operator convex.