HOMEWORK 0
Due: Friday, Sep 3, 11:59pm EST on Gradescope

Note: This homework covers some prerequisite material that you should be comfortable with, they will help you follow the lectures, and appreciate them. The relevant background is covered in basic classes on algorithms (e.g. 451), complexity, linear algebra, and probability and statistics. You may not have taken all four, or you may be rusty on some of the material. That's fine (and even expected). In that case, please brush up on the fundamentals as you solve this homework.

Exercises

These cover material that is important for the course. However, exercises are not graded, you need not submit solutions.

1. Matrix Norms, Singular Values, SVD. Let $A \in \mathbb{R}^{n\times n}$ be a not-necessarily-symmetric matrix. The *spectral norm* $\|A\|_2$ of A is defined as $\max_{v:\|v\|_2=1}\|Av\|_2$. The *Frobenius norm* of A is defined as $\sqrt{\sum_{i,j} A_{i,j}^2}$. The *trace* of A is defined as $\sum_{i\le n} A_{i,i}$. Let $\sigma_1 = \max_{v:\|v\|_2=1}\|Av\|_2$ and let v_1 be a unit vector such that $\|Av_1\|_2 = \sigma_1$. For each $i > 1$, inductively define $\sigma_i = \max_{v:\|v\|_2=1,\langle v,v_j\rangle=0 \text{ for every } j<i}\|Av\|_2$ and v_i to be a vector achieving this bound. Prove the following basic facts:

(a) **Spectral Norm vs Singular Value:** Prove that $\sigma_1^2 \ge \sigma_2^2 \ge \ldots \sigma_n^2$ are the n eigenvalues of the matrix $A^\top A$.

(b) **Spectral Norms vs Row/Column Norms:** $\|A\|_2 \le \max_{r,c}\{\sum_j |A(r,j)|, \sum_i |A(i,c)|\}$.

(c) **Comparison of Norms:** $\|A\|_2 \le \|A\|_F = \mathrm{tr}(A^\intercal A)^{1/2} = \sqrt{\sum_{i\le n}\sigma_i^2}$.

(d) **Orthogonal Invariance:** A matrix U is said to be *orthogonal* if $UU^\intercal = U^\intercal U = I$. Prove that $\|AU\|_2 = \|A\|_2$, $\|AU\|_F = \|A\|_F$ and $\mathrm{tr}(U^\intercal AU) = \mathrm{tr}(A)$.

(e) Define the *Frobenius inner product* $\langle A, B\rangle_F := \sum_{i,j} A_{ij}B_{ij}$. Show that $\langle A, B\rangle_F = \mathrm{tr}(AB)$.

(f) **Positive Semidefiniteness:** A symmetric matrix A is called *positive semidefinite (PSD)* if $v^\intercal Av \ge 0$ for every vector v. Prove that the following are equivalent:

   i. A is PSD,

   ii. all eigenvalues of A are non-negative,

   iii. $A = VV^\intercal$ for some matrix V, and

   iv. for every positive semidefinite matrix B, $\langle A, B\rangle_F \ge 0$.

Problems

1. The Spectral Hand. Let $G = (V, E)$ be a d-regular, undirected graph on vertices V and edge set E. Let $|V| = n$. For each edge $\{i, j\}$ of G, define matrix $L_{i,j} \in \mathbb{R}^{n \times n}$ such that $L_{i,j}(i, j) = L_{i,j}(j, i) = -1$ and $L_{i,j}(i, i) = L_{i,j}(j, j) = 1$. All other entries of $L_{i,j}$ are 0. The matrix $L = L_G = \sum_{\{i,j\} \in E} L_{i,j}$ is called the *Laplacian* matrix of G.

(a) (1 points) For any x, prove that $x^\intercal Lx = \sum_{\{i,j\}:\{i,j\} \in E}(x_i - x_j)^2$. Infer that L is PSD.

Observe that when $x \in \{0, 1\}^n$, $x^\intercal Ax$ equals ~~twice~~ the number of edges in the cut defined by the vertices $S = \{i \mid x_i = 1\}$.

(b) (3 points) Prove that max cut in G is at most $n\|L\|_2$ (Hint: Use part (a).) You can even prove $\frac{n}{4}\|L\|_2$, but we're happy with the weaker bound.

(c) (1 points) Prove that L has an eigenvalue equal to 0.

(d) (5 points) Prove that G is disconnected then 0 is an eigenvalue of L with multiplicity ≥ 2. (Hint: prove that there is a 2 dimensional subspace—related to the connected components—such that for every vector v in the subspace, $Lv = 0$.)

2. Let's Take a Moment. Let $X_1, X_2 \ldots, X_n$ be independent random variables such that $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^2 = 1$ for every i. Let $\overline{X} = \frac{1}{n}\sum_i X_i$. In this exercise, we will study the polynomial and the exponential moment methods to bound the probability that $\overline{X}$ deviates from 0.

(a) (2 points) Prove that $\Pr[\overline{X} \geq M/\sqrt{n}] \leq \frac{1}{M^2}$. (Hint: Apply Markov's inequality to $\overline{X}^2$.) When is this inequality tight?

(b) (5 points) Suppose that $\mathbb{E}[X_i^{2t-1}] = 0$ for every positive integer t and $\mathbb{E}[X_i^{2t}] \leq C^{2t}$ for some $C > 0$ and every $t \leq k$ for some positive integer k. ~~Compute $\mathbb{E}\overline{X}^{2t}$.~~ Give the best bound you can on $\mathbb{E}\overline{X}^{2k}$.

(c) (3 points) Under the hypothesis of part (b), prove that for some absolute constant $\cancel{C}C'$ and every $M > 0$, $\Pr[\overline{X} \geq \sqrt{C'k}\frac{M}{\sqrt{n}}] \leq \frac{1}{M^{2k}}$. (Hint: Apply Markov's inequality to an appropriate power of $\overline{X}$.) When is this inequality tight?

(d) (0 points) Convince yourself that the conclusions of parts (a) and (c) continue to hold even if X_is are pairwise and $2k$-wise independent, respectively. Recall that a set of random variables is *t-wise independent* if every subset of size at most t are mutually independent.

(e) (Bonus, Exponential Moment Method.) Prove that $\Pr[\overline{X} \geq M/\sqrt{n}] \leq e^{-CM^2}$ for some some absolute constant C. (Hint: Apply Markov's inequality to $Y = \exp(s\overline{X})$ for an appropriately chosen parameter s.)

3. LP Duality. Consider the problem of maximizing (or minimizing) $\sum_i c_i x_i$ over all $x \in \mathbb{R}^n$, $x_i \geq 0$ for $1 \leq i \leq n$ such that $\sum_j A_{ij}x_j \leq b_i$ for $1 \leq i \leq m$. This optimization problem is called a *linear program (LP)* in n variables with m linear inequality constraints, and the linear function $\sum_i c_i x_i$ is called the *objective* function. Points $x \in \mathbb{R}^n$ satisfying all the constraints are called *feasible*, and the maximum (or minimum) value of the objective function over all feasible x is called the *value* of the LP. Let's restrict to the maximization version in the following. Prove the following:

(a) (3 points) We want to derive an upper bound on the value of the LP. Let $\ell_1, \ell_2, \ldots, \ell_m$ be non-negative reals such that $\sum_i \ell_i A_{ik} \geq c_k$ for every $k \leq n$. Prove that the value of the LP is at most $\sum_{i \leq m} \ell_i b_i$.

(b) (5 points) Consider the problem of finding the smallest upper bound of the above form on the value of the LP. Prove that this is an LP with m variables and $n + m$ linear constraints. This LP is called the *dual* to the original LP.

(c) (2 points) What is the dual LP of the dual LP derived in part (b)?