HOMEWORK 0

Instructions : This homework familiarizes you with basic ideas from linear algebra and spectral graph theory that will be helpful (and will be assumed to be prior knowledge) in this class. We recommend that you do it yourself. Solutions will be posted in one week for your reference. This problem set is not graded.

The following exercises recall some basic facts. We will use them without proof in this course so it will be useful to solve all of them.

Preliminaries All matrices and vectors will be over the reals. In all the exercises below you can use the fact that any $n \times n$ matrix A has a singular value decomposition (SVD)

$$A = \sum_{i=1}^{r} \sigma_i u_i \otimes v_i \tag{1}$$

with $\sigma_i \in \mathbb{R}$ and $u_i, v_i \in \mathbb{R}^n$, and for every i, j $\|u_i\| = 1$, $\|v_j\| = 1$ (where $\|v\| = \sqrt{\sum v_i^2}$), and for all $i \neq j$, $\langle u_i, u_j \rangle = 0$ and $\langle v_i, v_j \rangle = 0$. The SVD of a matrix can be computed in polynomial time.

Equivalently $A = U \Sigma V^\top$ where Σ is a diagonal matrix and U and V are orthogonal matrices (satisfying $U^\top U = V^\top V = I$). If A is symmetric then there is such a decomposition with $u_i = v_i$ for all i (i.e., $U = V$). In this case the values $\sigma_1, \ldots, \sigma_r$ are known as *eigenvalues* of A and the vectors $v_1, \ldots, v_r$ are known as *eigenvectors*. You can use that there's a polynomial time algorithm A can be found in polynomial time. (You can ignore issues of numerical accuracy in all exercises.)

Exercises

Exercise 1 (Matrix Norms)**.** *Let A be an $n \times n$ matrix with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_n$. The* spectral *norm $\|A\|_2$ of A is the maximum of $\|Av\|_2$ over all vectors $v \in \mathbb{R}^n$ with $\|v\|_2 = 1$. The* Frobenius *norm $\|A\|_F$ of A is $\sqrt{\sum A_{i,j}^2}$.*

1. *Prove that $\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2$. When are these two inequalities tight (give examples)?*
2. *Prove that for symmetric A (i.e., $A = A^\top$), $\|A\|_2 \leq \max_i \sum_j |A_{i,j}|$*[1].
3. *Prove that for the adjacency matrix A of a d-regular graph $\|A\|_2 = d$.*

[1] **Hint:** You can do this via the following stronger inequality: for any (not necessarily symmetric) matrix A, $\|A\| \leq \sqrt{\alpha\beta}$ where $\alpha = \max_i \sum_j |A_{i,j}|$ and $\beta = \max_j \sum_i |A_{i,j}|$.

Exercise 2 (Trace Moment Method)**.** *Let A be a symmetric $n \times n$ matrix.*[2]

[2] Tao's blogpost explains this and other methods to prove sharp upper-bounds on the norms of random matrices

1. *Let* $\mathrm{Tr}(A) = \sum A_{i,i}$. *Prove that for every even* k, $\|A\|_2 \le \mathrm{Tr}(A^k)^{1/k} \le n^{1/k}\|A\|_2$.

2. *(Bonus) Let* A *be a symmetric matrix such that* $A_{i,i} = 0$ *for all* i *and* $A_{i,j}$ *is chosen to be a random value in* $\{\pm 1\}$ *independently of all others.* ***(a)*** *Prove that (for* n *sufficiently large) with probability at least* 0.99, $\|A\|_2 \le n^{0.9}$. ***(b)*** *Prove that with probability at least* 0.99, $\|A\|_2 \le n^{0.51}$.

Note: While $\|A\|_2$ can be computed in polynomial time, both $\max_i \sum_j |A_{i,j}|$ and $\|A\|_F$ give even simpler to compute upper bounds for $\|A\|_2$. However the examples in Exercise 1 show that they are not always tight. It is often easier to compute $\mathrm{Tr}(A^k)^{1/k}$ than trying to compute $\|A\|_2$ directly, and as k grows this yields a better and better estimate.

Exercise 3 (Positive Semidefinite Matrices)**.** *Let* A *be an* $n \times n$ *symmetric matrix. Prove that the following are equivalent:*

1. *For every vector* $v \in R^n$, $v^\top A v \ge 0$ *(note:* $v^\top A v = \sum_{i,j} A_{i,j} v_i v_j$*).*

2. *All eigenvalues of* A *are non-negative.*

3. *There are linear functions* $L_1, \ldots, L_m$ *such that the quadratic polynomial* P_A *defined as* $P_A(x) = \sum A_{i,j} x_i x_j$ *can be written as* $P_A = \sum_i L_i^2$. *In particular, the polynomial* P_A *is a* sum of squares*.*

4. $A = B^\top B$ *for some* $r \times n$ *matrix* B

Any of the above statements can be taken as a definition of positive semidefinite matrices (denoted by $A \succeq 0$*.)*

Exercise 4 (Gaussian Distributions)**.** *The standard Gaussian random variable* g *on* $\mathbb{R}^n$ *has the following probability density function:* $\gamma(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp(-\frac{\|x\|_2^2}{2})$.

1. *(Rotation Invariance) Let* g *be a a standard Gaussian random variable. Prove that for any matrix* U *such that* $U^\top U = I = UU^\top$, Ug *is also a standard Gaussian random variable*[3]*.*

2. *(Orthogonality implies Independence) Prove*[4] *that for any orthogonal pair of vectors* v, w*, and standard Gaussian* g*, the random variables* $\langle g, v\rangle$ *and* $\langle g, w\rangle$ *are independent.*

[3] **Hint:** For any v, how is $\|Uv\|_2$ related to $\|v\|_2$?

[4] First deal with the special case when $v = e_1$ and $w = e_2$, the standard basis vectors.

Exercise 5 (Covariance Matrices)**.** *Let* D *be a probability distribution on* $\mathbb{R}^n$*. The* covariance matrix *of* D *is defined (assuming it exists) as* $\mathbb{E}_D(x - \mathbb{E}_D\, x)(x - \mathbb{E}_D\, x)^\top$.

1. *Prove that for any vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$, the matrix $M = \sum_{i \leq n} v_i v_i^\top$ is positive semidefinite.*

2. *Prove that the covariance matrix of any distribution D is positive semidefinite.*

3. *Prove that for any positive semidefinite matrix $M \in \mathbb{R}^{n \times n}$, there exists a probability distribution D on $\mathbb{R}^n$ with covariance M (and mean $\mathbf{0}$).*

Exercise 6 (Basic Convexity and Hyperplane Separation)**.** *Let $K \subseteq \mathbb{R}^n$. K is said to be* convex *if for every $x, y \in K$ and any $\alpha \in [0,1]$, $\alpha x + (1-\alpha) y \in K$. Let $B = \{x \mid \sum_i x_i^2 \leq 1\}$ be the unit ball in ℓ_2 norm. K is said to be* bounded *if $K \subseteq cB$ for some finite c. A point x is said to be a* limit point *of K if x can be approximated arbitrarily well by points in K. That is, for every $\epsilon > 0$, there exists a $y \in K$ such that $\|x - y\|_2 \leq \epsilon$. K is said to be* closed *if every limit point of K is contained in K. The* halfspace *defined by a vector h and a threshold τ is the set: $H_{h,\tau} = \{x \in \mathbb{R}^n \mid \langle h, x \rangle \geq \tau\}$.*

In the following, we will prove the following fundamental theorem of convex analysis: For any convex set K that is closed, bounded[5] *and convex and $v \notin K$, there exists a halfspaces $H_{h,\tau}$ such that $K \subseteq H_{h,\tau}$ while $v \notin H_{h,\tau}$. The hyperplane $\{x \mid \langle h, x \rangle = \tau\}$ is said to be a* separating *hyperplane for K and v.*

[5] Both these assumptions are not necessary. But the proof is slightly more technical.

1. *(Projection) Let K be a closed, bounded, convex set. Let $v \notin K$. Let $\mathrm{dist} : K \to \mathbb{R}$ be the function defined by $\mathrm{dist}(x) = \|v - x\|^2$. Verify that dist is continuous over K. Apply the extreme value theorem*[6] *to conclude that there exists a points $m \in K$ such that $0 < \mathrm{dist}(m) \leq \mathrm{dist}(x)$ for every $x \in K$.*

[6] The extreme value theorem says that a continuous function on a closed, bounded subset of $\mathbb{R}^n$ attains its supremum and infimum.

2. *Let $f_v : K \to \mathbb{R}$ be defined as $f_v(x) = \langle m - v, x \rangle - \frac{\|m\|_2^2 - \|v\|_2^2}{2}$. Our goal is to prove that $\{x \mid f_v(x) = 0\}$ is a good separating hyperplane. Towards this, verify that $f_v((m+v)/2) = 0$. Prove that $f_v(v) < 0$. Have you used convexity of K yet?*

3. *We will now prove that $f_v(x) > 0$ for every $x \in K$. Suppose towards a contradiction that there exists a $y \in K$ such that $f_v(y) \leq 0$. Prove that $\nabla f_v(m)^\top (y - m) < 0$. That is, "moving" in the direction of y decreases f_v.*

4. *Conclude that there is a positive $1 > \alpha > 0$ such that $f_v(m + \alpha(y - m)) < f_v(m)$. Derive a contradiction to finish the proof.*

Exercise 7 (Separation Oracle for PSD Matrices)**.**

1. *Prove that* $\mathbb{S}_+ = \{M \mid M \succeq 0\} \subseteq \mathbb{R}^{n\times n}$ *is a closed, convex set. Use the definition of a closed set given in the previous exercise, where the distance between two matrices* M_1 *and* M_2 *is* $\|M_1 - M_2\|_F$.

2. *Suppose* $N \in \mathbb{R}^{n\times n}$ *is not positive semidefinite. Show that there's a polynomial time algorithm to compute a matrix* B *such that* $\sum_{i,j} B_{i,j}M_{i,j} \geq 0$ *for every* $M \in \mathbb{S}_+$ *while* $\sum_{i,j} B_{i,j}N_{i,j} < 0$.