Please provide complete and well-written solutions to the following exercises.

Due October 4, in the discussion section.

Homework 1

Exercise 1. Let A be an $n \times n$ real symmetric matrix. Show that the following three conditions are equivalent:

- A is positive semidefinite
- All eigenvalues of A are nonnegative.
- There exists a real $n \times n$ matrix B such that $A = BB^T$.

(Hint: you should probably use the Spectral Theorem for Symmetric Matrices.)

Exercise 2. Show that the two different versions of the Spectral Theorem for Symmetric Matrices from the notes are equivalent. That is, show that the following statements are equivalent:

- Let A be an $n \times n$ real symmetric matrix. Then there exists an orthogonal $n \times n$ matrix Q whose columns are each eigenvectors of A, and there exists a real diagonal matrix D whose diagonal entries are the eigenvalues of A such that $Q^{-1}AQ = D$. That is, $A = QDQ^{-1}$.
- Equivalently, if $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ are the eigenvalues of A (where some eigenvalues are allowed to be the same), then $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$, and there exist vectors $v_1, \ldots, v_n \in \mathbf{R}^n$ which are an orthonormal basis of $\mathbf{R}^n$ such that $A = \sum_{i=1}^n \lambda_i v_i v_i^T$.

Exercise 3. Let A be an $n \times n$ real symmetric matrix. Let $\lambda_1 \geq \cdots \geq \lambda_n \in \mathbf{R}$ be the eigenvalues of A, ordered according to their size. Let $x \in \mathbf{R}^n$. Show that

$$\lambda_1 \left|\left|x\right|\right|^2 \geq x^T A x \geq \lambda_n \left|\left|x\right|\right|^2.$$

Exercise 4. Prove the **Cauchy-Schwarz inequality**: For any $x, y \in \mathbf{R}^n$, we have

$$\left|\langle x, y\rangle\right| \leq \left|\left|x\right|\right| \left|\left|y\right|\right|.$$

(Hint: subtract the projection of y onto x. That is, if $x \neq 0$, let $v := \frac{\langle x,y\rangle}{\langle x,x\rangle} x$, and expand out the inequality $\left|\left|y - v\right|\right|^2 \geq 0$.)

Exercise 5. Prove the **triangle inequality**: For any $x, y, z \in \mathbf{R}^n$,

$$\left|\left|x - y\right|\right| \leq \left|\left|x - z\right|\right| + \left|\left|z - y\right|\right|.$$

(Hint: it may be conceptually easier to show $\left|\left|x + y\right|\right| \leq \left|\left|x\right|\right| + \left|\left|y\right|\right|$. To show this inequality, square both sides, and use the Cauchy-Schwarz inequality.) Then, deduce the **reverse triangle inequality**:

$$\left|\left|x - y\right|\right| \geq \left|\,\left|\left|x\right|\right| - \left|\left|y\right|\right|\,\right|.$$

Exercise 6 (The Power Method)**.** This exercise gives an algorithm for finding the eigenvectors and eigenvalues of a symmetric matrix. The Power Method described below is not the best algorithm for this task, but it is perhaps the easiest to describe and analyze.

Let A be an $n \times n$ real symmetric matrix. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the (unknown) eigenvalues of A, and let $v_1, \ldots, v_n \in \mathbf{R}^n$ be the corresponding (unknown) eigenvectors of A such that $||v_i|| = 1$ and such that $Av_i = \lambda_i v_i$ for all $1 \leq i \leq n$.

Given A, our first goal is to find v_1 and λ_1. For simplicity, assume that $1/2 < \lambda_1 < 1$, and $0 \leq \lambda_n \leq \cdots \leq \lambda_2 < 1/4$. Suppose we have found a vector $v \in \mathbf{R}^n$ such that $||v|| = 1$ and $|\langle v, v_1 \rangle| > 1/n$. (A randomly chosen v will satisfy $|\langle v, v_1 \rangle| > 1/(10\sqrt{n})$, [which is a nice optional exercise for those who have taken 170A], so this assumption is valid in practice.) Let k be a positive integer. Show that

$$A^k v$$

approximates v_1 well as k becomes large. More specifically, show that for all $k \geq 1$,

$$\left|\left|A^k v - \langle v, v_1 \rangle \lambda_1^k v_1\right|\right|^2 \leq \frac{n-1}{16^k}.$$

Since $|\langle v, v_1 \rangle| \, \lambda_1^k > 2^{-k}/n$, this inequality implies that $A^k v$ is approximately an eigenvector of A with eigenvalue λ_1. That is, by the triangle inequality,

$$\left|\left|A(A^k v) - \lambda_1 (A^k v)\right|\right| \leq \left|\left|A^{k+1} v - \langle v, v_1 \rangle \lambda_1^{k+1} v_1\right|\right| + \lambda_1 \left|\left|\langle v, v_1 \rangle \lambda_1^k v_1 - A^k v\right|\right| \leq 2\frac{\sqrt{n-1}}{4^k}.$$

Moreover, by the reverse triangle inequality,

$$\left|\left|A^k v\right|\right| = \left|\left|A^k v - \langle v, v_1 \rangle \lambda_1^k v_1 + \langle v, v_1 \rangle \lambda_1^k v_1\right|\right| \geq \frac{1}{n} 2^{-k} - \frac{\sqrt{n-1}}{4^k}.$$

In conclusion, if we take k to be large (say $k > 10 \log n$), and if we define $z := A^k v$, then z is approximately an eigenvector of A, that is

$$\left|\left|A \frac{A^k v}{||A^k v||} - \lambda_1 \frac{A^k v}{||A^k v||}\right|\right| \leq 4n^{3/2} 2^{-k} \leq 4n^{-4}.$$

And to approximately find the first eigenvalue λ_1, we simply compute

$$\frac{z^T A z}{z^T z}.$$

That is, we have approximately found the first eigenvector and eigenvalue of A.

Remarks. To find the second eigenvector and eigenvalue, we can repeat the above procedure, where we start by choosing v such that $\langle v, v_1 \rangle = 0$, $||v|| = 1$ and $|\langle v, v_2 \rangle| > 1/(10\sqrt{n})$. To find the third eigenvector and eigenvalue, we can repeat the above procedure, where we start by choosing v such that $\langle v, v_1 \rangle = \langle v, v_2 \rangle = 0$, $||v|| = 1$ and $|\langle v, v_3 \rangle| > 1/(10\sqrt{n})$. And so on.

Google's PageRank algorithm uses the power method to rank websites very rapidly. In particular, they let n be the number of websites on the internet (so that n is roughly 10^9). They then define an $n \times n$ matrix C where $C_{ij} = 1$ if there is a hyperlink between websites i and j, and $C_{ij} = 0$ otherwise. Then, they let B be an $n \times n$ matrix such that B_{ij} is 1 divided by the number of 1's in the i^{th} row of C, if $C_{ij} = 1$, and $B_{ij} = 0$ otherwise. Finally, they define

$$A = (.85)B + (.15)D/n$$

where D is an $n \times n$ matrix all of whose entries are 1.

The power method finds the eigenvector v_1 of A, and the size of the i^{th} entry of v_1 is proportional to the "rank" of website i.

Exercise 7 (This exercise is optional; any exercise in this course involving programming is optional)**.** Write a program in Matlab that computes the first, second, and third eigenvectors and eigenvalues of a symmetric matrix A of arbitrary size. Compare your results with the Matlab programs `eigs` and `eig`.

Then, under the assumptions of the previous exercise ($1/2 < \lambda_1 < 1$, and $0 \leq \lambda_n \leq \cdots \leq \lambda_2 < 1/4$), provide an upper bound on the number of arithmetic operations that are required to compute the first three decimal places of the first eigenvalue λ_1. Your upper bound could involve either the size n of the matrix A, or the number m of nonzero entries of A.

Note that the power method iteratively applies the matrix to a vector, instead of multiplying matrices together. The latter operation can require many more arithmetic operations than the former.

Exercise 8. Show that the intersection of two convex sets is convex.

Exercise 9. Let $f\colon \mathbf{R} \to \mathbf{R}$ so that $f(x) = x^2$. Show that f is convex.

Exercise 10. Let $f\colon \mathbf{R} \to \mathbf{R}$ be a function with three continuous derivatives. Show that f is convex if and only if $f''(x) \geq 0$ for all $x \in \mathbf{R}$. (Hint: for the reverse implication, you may need to use Taylor's Theorem with integral remainder.)

Exercise 11. Prove Taylor's Theorem with Integral Remainder when $k = 2$:

Let $f\colon \mathbf{R} \to \mathbf{R}$ be a function with three continuous derivatives. Then, for any $x, y \in \mathbf{R}$,

$$f(x) = f(y) + (x - y)f'(y) + \int_y^x f''(t)(x - t)dt.$$

(Hint: integrate by parts.)