Homework 11

This homework is due on Saturday, April 15, 2023 at 11:59PM. Self-grades and HW Resubmissions are due the following Saturday, April 22, 2023 at 11:59PM.

1. **Spectral Theorem for Real Symmetric Matrices**

   We want to show that every real symmetric matrix can be diagonalized by a matrix of its orthonormal eigenvectors. In other words, a symmetric matrix $S \in \mathbb{R}^{n\times n}$, i.e., a matrix S such that $S = S^\top$, can be written as $S = V\Lambda V^\top$, where $V \in \mathbb{R}^{n\times n}$ is an orthonormal matrix of eigenvectors of S and $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix of corresponding real eigenvalues of S. This is called the Spectral Theorem for real symmetric matrices.

   To do this, we will use a proof which is similar to the proof of existence of the Schur decomposition. Along the way, we will practice block matrix manipulation and the induction proof technique.

   (a) One part of the spectral theorem can be proved without any further delay. **Prove that the eigenvalues λ of a real, symmetric matrix S are real.**

   (HINT: Let λ be an eigenvalue of S with corresponding nonzero eigenvector $\vec{v}$. Evaluate $\overline{\vec{v}}^\top S\vec{v}$ in two different ways: $\overline{\vec{v}}^\top(S\vec{v})$ and $(\overline{\vec{v}}^\top S)\vec{v}$. What does this show about λ?)

   (b) For the main proof that every real symmetric matrix is diagonalized by a matrix of its orthonormal real eigenvectors, we will proceed by *induction*.

   Recall that an inductive proof trying to prove a statement that depends on n, say P_n[1], is true for all positive integers n, has two steps:

   - A base case – prove that P_1 is true.
   - An inductive step – for every $n \geq 2$, given that P_{n-1} is true, prove that P_n is true.[2]

   By doing these two steps, we show P_n is true for all n.

   In our case, the statement P_n is "every $n \times n$ symmetric matrix S can be diagonalized as $S = V\Lambda V^\top$, where V is the real orthonormal matrix of eigenvectors of S, and Λ is the real diagonal matrix of corresponding eigenvalues of S."

   Show the base case: every 1×1 symmetric matrix S can be written as $S = V\Lambda V^\top$, where V is a real and orthonormal matrix of eigenvectors of S, and Λ is a real and diagonal matrix of corresponding eigenvalues of S.

   (HINT: Every 1×1 matrix is symmetric, and also diagonal, by definition; the only real orthonormal 1×1 matrices are $\begin{bmatrix}1\end{bmatrix}$ and $\begin{bmatrix}-1\end{bmatrix}$.)

[1] Lecture used S_n, but S is already being used for symmetric matrix here.

[2] This is the so-called *weak induction* paradigm; it contrasts with *strong induction*, which you can learn in future classes like CS70.

(c) With the base case done, we are now in the inductive step. Let S be an arbitrary $n \times n$ symmetric matrix; ultimately, we want to show that $S = V\Lambda V^\top$, where V is a real and orthonormal matrix of eigenvectors of S, and Λ is a real and diagonal matrix of corresponding eigenvalues of S.

To start, let λ be an eigenvalue of S, and let $\vec{q}$ be any normalized eigenvector of S corresponding to eigenvalue λ. Let $\widetilde{Q} \in \mathbb{R}^{n\times(n-1)}$ be a set of orthonormal vectors chosen so that $Q := \begin{bmatrix} \vec{q} & \widetilde{Q} \end{bmatrix} \in \mathbb{R}^{n\times n}$ is an orthonormal matrix.[3] **Show the following equality:**

$$Q^\top S Q = \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & S_0 \end{bmatrix} \qquad \textbf{where} \qquad S_0 := \widetilde{Q}^\top S \widetilde{Q}. \tag{1}$$

(HINT: Expand Q as a block matrix $\begin{bmatrix} \vec{q} & \widetilde{Q} \end{bmatrix}$ and multiply $Q^\top S Q = \begin{bmatrix} \vec{q} & \widetilde{Q} \end{bmatrix}^\top S \begin{bmatrix} \vec{q} & \widetilde{Q} \end{bmatrix}$.)
(HINT: Since Q is orthonormal, we have $Q^\top Q = I_n$. What does this mean for the values of $\vec{q}^\top \vec{q}$ and $\widetilde{Q}^\top \vec{q}$? Use block matrix multiplication on $Q^\top Q = \begin{bmatrix} \vec{q} & \widetilde{Q} \end{bmatrix}^\top \begin{bmatrix} \vec{q} & \widetilde{Q} \end{bmatrix}$ again.)

(d) **Show that the matrix S_0 is a real symmetric matrix.**

(e) Since S_0 is a real symmetric $(n-1)\times(n-1)$ matrix, by our inductive assumption, S_0 can be orthonormally diagonalized as $S_0 = V_0 \Lambda_0 V_0^\top$, where Λ_0 is a real diagonal matrix of eigenvalues of S_0 and $V_0 \in \mathbb{R}^{(n-1)\times(n-1)}$ is a real orthonormal matrix of corresponding eigenvectors of S_0.

Define

$$V := Q \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} \qquad \textbf{and} \qquad \Lambda := V^\top S V. \tag{2}$$

i. **Show that V is orthonormal.**

ii. **Show that Λ is diagonal.**

iii. **Show that $S = V\Lambda V^\top$.**

(HINT: Use block matrix multiplication again.)

Thus, we have found a real orthonormal V and real diagonal Λ such that $S = V\Lambda V^\top = V\Lambda V^{-1}$. We have seen in a previous homework that if $A = V\Lambda V^{-1}$, then Λ are the eigenvalues of A, and V are the corresponding eigenvectors. Thus, given P_{n-1} – the fact that we can orthonormally diagonalize $(n-1)\times(n-1)$ real symmetric matrices – we have proven P_n – the fact that we can orthonormally diagonalize $n \times n$ real symmetric matrices. Thus, we've proved the Spectral Theorem for real symmetric matrices by induction!

[3] This matrix $\widetilde{Q}$ can be generated via Gram-Schmidt, for example.

2. SVD

(a) Consider the matrix

$$A = \begin{bmatrix} -1 & 1 & 5 \\ 3 & 1 & -1 \\ 2 & -1 & 4 \end{bmatrix}.$$

Observe that the columns of matrix A are mutually orthogonal with norms $\sqrt{14}, \sqrt{3}, \sqrt{42}$.

Verify numerically that columns $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ **and** $\begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix}$ **are orthogonal to each other.**

(b) **Write $A = BD$, where B is an orthonormal matrix and D is a diagonal matrix.** What is B? What is D?

(c) **Write out a singular value decomposition of $A = U\Sigma V^\top$ using the previous part.** Note the ordering of the singular values in Σ should be from the largest to smallest. *(HINT: There is no need to compute the eigenvalues of anything. Use Theorem 14, Note 14.)*

(d) Given the matrix

$$A = \frac{1}{\sqrt{50}} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} + \frac{3}{\sqrt{50}} \begin{bmatrix} -4 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}, \tag{3}$$

write out a singular value decomposition of matrix A in the form $U\Sigma V^\top$. Note the ordering of the singular values in Σ should be from the largest to smallest. *(HINT: You don't have to compute any eigenvalues for this. Some useful observations are that*

$$\begin{bmatrix} 3, 4 \end{bmatrix} \begin{bmatrix} -4 \\ 3 \end{bmatrix} = 0, \quad \begin{bmatrix} 1, -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0, \quad \left\| \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -4 \\ 3 \end{bmatrix} \right\| = 5, \quad \left\| \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\| = \sqrt{2}.$$

)

(e) Define the matrix

$$A = \begin{bmatrix} -1 & 4 \\ 1 & 4 \end{bmatrix}.$$

Find the SVD of A by following the standard algorithm introduced in Note 14, i.e. by computing the eigendecomposition of $A^\top A$. Also find the eigenvectors and eigenvalues of A. Is there a relationship between the eigenvalues or eigenvectors of A with the SVD of A?

3. Frobenius Norm

In this problem we will investigate the basic properties of the Frobenius norm.

Similar to how the norm of vector $\vec{x} \in \mathbb{R}^n$ is defined as $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$, the Frobenius norm of a matrix $A \in \mathbb{R}^{m\times n}$ is defined as

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{ij}|^2}. \tag{4}$$

A_{ij} is the entry in the i^{th} row and the j^{th} column. This is basically the norm that comes from treating a matrix like a big vector filled with numbers.

(a) With the above definitions, **show that for a 2×2 matrix A:**

$$\|A\|_F = \sqrt{\operatorname{tr}(A^\top A)}. \tag{5}$$

Note: The trace of a matrix is the sum of its diagonal entries. For example, let $A \in \mathbb{R}^{m\times n}$, then,

$$\operatorname{tr}(A) = \sum_{i=1}^{\min(n,m)} A_{ii} \tag{6}$$

Think about how/whether this expression eq. (5) generalizes to general $m \times n$ matrices.

(b) **Show for any matrix $A \in \mathbb{R}^{m\times n}$:**

$$\|A\|_F = \left\|A^\top\right\|_F \tag{7}$$

A purely written or mathematical solution will be sufficient for this problem.

(HINT: For the mathematical solution, use the trace interpretation from eq. (4).)

(c) **Show that if U and V are square orthonormal matrices, then**

$$\|UA\|_F = \|AV\|_F = \|A\|_F\,. \tag{8}$$

(HINT: Use the trace interpretation from part (a) and the equation from part (b).)

(d) **Use the SVD decomposition to show that $\|A\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$, where $\sigma_1, \ldots, \sigma_n$ are the singular values of A.**

(HINT: The previous part might be quite useful.)

Contributors:

- Siddharth Iyer.
- Yu-Yun Dai.
- Sanjit Batra.
- Anant Sahai.
- Sidney Buchbinder.
- Gaoyue Zhou.
- Druv Pai.
- Elena Jia.

- Tanmay Gautam.
- Anirudh Rengarajan.
- Aditya Arun.
- Kourosh Hakhamaneshi.
- Antroy Roy Chowdhury.