Homework 10

This homework is due on Friday, April 8, 2022, at 11:59PM. Self-grades and HW Resubmissions are due on the following Friday, April 15, 2022, at 11:59PM.

1. **Reading Lecture Notes**

   Staying up to date with lectures is an important part of the learning process in this course. Here are links to the notes that you need to read for this week: Note 15 Note 16

   (a) Consider two vectors $\vec{x} \in \mathbb{R}^m$ and $\vec{y} \in \mathbb{R}^n$. **What are the dimensions of the matrix $\vec{x}\vec{y}^\top$ and what is the rank of $\vec{x}\vec{y}^\top$?**

   Solution: Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, then

$$\vec{x}\vec{y}^\top = \vec{x}\begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ y_1\vec{x} & y_2\vec{x} & \dots & y_n\vec{x} \\ | & | & & | \end{bmatrix} \tag{1}$$

$$\vec{x}\vec{y}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\vec{y}^\top = \begin{bmatrix} - & x_1\vec{y}^\top & - \\ - & x_2\vec{y}^\top & - \\ & \vdots & \\ - & x_m\vec{y}^\top & - \end{bmatrix} \tag{2}$$

   We will have n multiples of $\vec{x}$, and since $\vec{x}$ has m entries, the matrix has dimensions $m \times n$. We can see that each column of $\vec{x}\vec{y}^\top$ is a multiple of $\vec{x}$ and each row is a multiple of $\vec{y}^\top$, thus $\vec{x}\vec{y}^\top$ has rank 1, unless either of the vectors $\vec{x}$ or $\vec{y}$ was the (appropriate shape) all zero vector $\vec{0}$. If the result is the zero matrix, the rank would be 0.

   (b) Consider a matrix $A \in \mathbb{R}^{m\times n}$ and the rank of A is r. Suppose its SVD is $A = U\Sigma V^\top$ where $U \in \mathbb{R}^{m\times m}$, $\Sigma \in \mathbb{R}^{m\times n}$, and $V \in \mathbb{R}^{n\times n}$. **Write A in terms of the singular values of A and outer products of the columns of U and V.**

   Solution: We have $A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top$ where $\vec{u}_i$ and $\vec{v}_i$ are the i-th columns of U and V. Note that we only sum to r since A has rank r and hence it has r non-zero singular values $\sigma_1, ..., \sigma_r$. This is the outer product form of the SVD.

2. 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?)

Solution: Using the fact that S is real and symmetric so $\overline{S} = S = S^\top$, we get

$$\overline{\vec{v}}^\top(S\vec{v}) = \overline{\vec{v}}^\top(\lambda\vec{v}) = \lambda\overline{\vec{v}}^\top\vec{v} = \lambda\|\vec{v}\|^2 \tag{3}$$

$$(\overline{\vec{v}}^\top S)\vec{v} = (S\overline{\vec{v}})^\top\vec{v} = (\overline{S\vec{v}})^\top\vec{v} = (\overline{\lambda\vec{v}})^\top\vec{v} = \overline{\lambda}(\overline{\vec{v}}^\top\vec{v}) = \overline{\lambda}\|\vec{v}\|^2. \tag{4}$$

where $\|\vec{v}\|^2 = \sum_{i=1}^n |v_i|^2$. Since $\vec{v} \neq \vec{0}_n$, we know that $\|\vec{v}\|^2 > 0$, and so $\lambda = \overline{\lambda}$. Thus λ is real.

(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 $\begin{bmatrix}-1\end{bmatrix}$.)

Solution: Let $S = \begin{bmatrix}s\end{bmatrix}$. Since $\begin{bmatrix}1\end{bmatrix}$ is a real and orthonormal matrix, and $\begin{bmatrix}s\end{bmatrix}$ is diagonal, $S = \begin{bmatrix}1\end{bmatrix}\begin{bmatrix}s\end{bmatrix}\begin{bmatrix}1\end{bmatrix}^\top$ is an orthonormal diagonalization of S. Since $S\vec{x} = s\vec{x}$ for all $\vec{x} \in \mathbb{R}^1$, we see that $\begin{bmatrix}s\end{bmatrix}$ is a matrix of eigenvalues of S, and also that any vector is an eigenvector so an orthonormal matrix of eigenvectors of S is $\begin{bmatrix}1\end{bmatrix}$.

It is also possible to answer with $S = \begin{bmatrix}-1\end{bmatrix}\begin{bmatrix}s\end{bmatrix}\begin{bmatrix}-1\end{bmatrix}^\top$.

(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.

[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.

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:**

$$Q^\top S Q = \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & S_0 \end{bmatrix} \qquad \textbf{where} \qquad S_0 := \widetilde{Q}^\top S \widetilde{Q}. \tag{5}$$

(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.)

Solution: We use block-matrix multiplication:

$$\begin{aligned}
Q^\top S Q &= \begin{bmatrix} \vec{q}^\top \\ \widetilde{Q}^\top \end{bmatrix} S \begin{bmatrix}\vec{q} & \widetilde{Q}\end{bmatrix} && (6)\\
&= \begin{bmatrix} \vec{q}^\top \\ \widetilde{Q}^\top \end{bmatrix} \begin{bmatrix}S\vec{q} & S\widetilde{Q}\end{bmatrix} && (7)\\
&= \begin{bmatrix} \vec{q}^\top \\ \widetilde{Q}^\top \end{bmatrix} \begin{bmatrix}\lambda\vec{q} & S\widetilde{Q}\end{bmatrix} && (8)\\
&= \begin{bmatrix} \lambda\vec{q}^\top\vec{q} & \lambda\vec{q}^\top\widetilde{Q} \\ \lambda\widetilde{Q}^\top\vec{q} & \widetilde{Q}^\top S\widetilde{Q} \end{bmatrix}. && (9)
\end{aligned}$$

To simplify, we follow the hint, and expand $Q^\top Q = I_n$.

$$\begin{aligned}
Q^\top Q &= I_n && (10)\\
\begin{bmatrix} \vec{q}^\top \\ \widetilde{Q}^\top \end{bmatrix} \begin{bmatrix}\vec{q} & \widetilde{Q}\end{bmatrix} &= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & I_{n-1} \end{bmatrix} && (11)\\
\begin{bmatrix} \vec{q}^\top\vec{q} & \vec{q}^\top\widetilde{Q} \\ \widetilde{Q}^\top\vec{q} & \widetilde{Q}^\top\widetilde{Q} \end{bmatrix} &= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & I_{n-1} \end{bmatrix}. && (12)
\end{aligned}$$

Thus we get $\vec{q}^\top\vec{q} = 1$, $\widetilde{Q}^\top\vec{q} = \vec{0}_{n-1}$, $\vec{q}^\top\widetilde{Q} = \vec{0}_{n-1}^\top$, and so we have

$$Q^\top S Q = \begin{bmatrix} \lambda\vec{q}^\top\vec{q} & \vec{q}^\top\widetilde{Q} \\ \lambda\widetilde{Q}^\top\vec{q} & \widetilde{Q}^\top S\widetilde{Q} \end{bmatrix} = \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & \widetilde{Q}^\top S\widetilde{Q} \end{bmatrix} \tag{13}$$

(d) **Show that the matrix S_0 is a real symmetric matrix.**
Solution: We show that $S_0^\top = S_0$.

$$\begin{aligned}
S_0^\top &= (\widetilde{Q}^\top S\widetilde{Q})^\top && (14)\\
&= (\widetilde{Q})^\top (S)^\top (\widetilde{Q}^\top)^\top && (15)\\
&= \widetilde{Q}^\top S^\top \widetilde{Q} && (16)\\
&= \widetilde{Q}^\top S \widetilde{Q} && (17)\\
&= S_0. && (18)
\end{aligned}$$

where the second-to-last equality is because S is symmetric so $S^\top = S$.
It is not necessary to write in the solution, but to show that S_0 is real, note that $\widetilde{Q}$ is real and S is real, so $S_0 = \widetilde{Q}^\top S\widetilde{Q}$ is real as a matrix product of real matrices.

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

(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

$$V := Q\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} \qquad \textbf{and} \qquad \Lambda := V^\top S V. \tag{19}$$

i. **Show that V is orthonormal.**
Solution: We compute $V^\top V$.

$$\begin{aligned}
V^\top V &= \left(Q\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}\right)^\top \left(Q\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}\right) && (20)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}^\top Q^\top Q \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} && (21)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}^\top \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} && (22)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0^\top \end{bmatrix} \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} && (23)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0^\top V_0 \end{bmatrix} && (24)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & I_{n-1} \end{bmatrix} && (25)\\
&= I_n. && (26)
\end{aligned}$$

It is not necessary to write in the solution, but to show that V is real, note that Q is real and $\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}$ is real, so $V = Q\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}$ is real as a matrix product of real matrices.

ii. **Show that Λ is diagonal.**
Solution: We compute $\Lambda = V^\top S V$.

$$\begin{aligned}
\Lambda &= V^\top S V && (27)\\
&= \left(Q\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}\right)^\top S \left(Q\begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}\right) && (28)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix}^\top Q^\top S Q \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} && (29)\\
&= \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0^\top \end{bmatrix} \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & S_0 \end{bmatrix} \begin{bmatrix} 1 & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0 \end{bmatrix} && (30)\\
&= \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & V_0^\top S_0 V_0 \end{bmatrix} && (31)\\
&= \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & \Lambda_0 \end{bmatrix}. && (32)
\end{aligned}$$

We already know Λ_0 is diagonal so Λ is diagonal.
It is not necessary to write in the solution, but to show that Λ is real, note that λ is real (shown in part (a)) and Λ_0 is real by the induction, so $\Lambda = \begin{bmatrix} \lambda & \vec{0}_{n-1}^\top \\ \vec{0}_{n-1} & \Lambda_0 \end{bmatrix}$ is real.

iii. **Show that $S = V\Lambda V^\top$.**
Solution: We have

$$\Lambda = V^\top S V \tag{33}$$
$$\implies V\Lambda = SV \tag{34}$$
$$\implies V\Lambda V^\top = S. \tag{35}$$

(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. 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.**

Solution: Taking the inner product of the two vectors, we have

$$\left\langle \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix} \right\rangle = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}^\top \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix} = 5 - 1 - 4 = 0. \tag{36}$$

So the two columns 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?

Solution: We compute the norm for each column and divide each column by its norm to obtain matrix B. Matrix D is formed by placing the norms on the diagonal.

$$B = \begin{bmatrix} -\frac{1}{\sqrt{14}} & \frac{1}{\sqrt{3}} & \frac{5}{\sqrt{42}} \\ \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{42}} \\ \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{3}} & \frac{4}{\sqrt{42}} \end{bmatrix} \tag{37}$$

$$D = \begin{bmatrix} \sqrt{14} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{42} \end{bmatrix} \tag{38}$$

(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 16.)*

Solution: Using part b, we can write

$$A = BD = BDI = \begin{bmatrix} -\frac{1}{\sqrt{14}} & \frac{1}{\sqrt{3}} & \frac{5}{\sqrt{42}} \\ \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{42}} \\ \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{3}} & \frac{4}{\sqrt{42}} \end{bmatrix} \begin{bmatrix} \sqrt{14} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{42} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{39}$$

Reordering the columns and rows of B and I so that the diagonal entries of D are in non-decreasing order, we have

$$\begin{bmatrix} \frac{5}{\sqrt{42}} & -\frac{1}{\sqrt{14}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{42}} & \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{3}} \\ \frac{4}{\sqrt{42}} & \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{42} & 0 & 0 \\ 0 & \sqrt{14} & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{40}$$

Then by Note 16, Theorem 14, this is an SVD of A.

(d) Given the matrix

$$A = \frac{1}{\sqrt{50}} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} + \frac{3}{\sqrt{50}} \begin{bmatrix} -4 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}, \tag{41}$$

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*

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

)

Solution:

The singular value decomposition can be written in the form

$$A=\sum_{i=1}^{2}\sigma_i\vec{u}_i\vec{v}_i^\top, \tag{42}$$

with unit orthonormal vectors $\{\vec{u}_i\}$ and $\{\vec{v}_i\}$. From the given observations, we can see that the vectors we were provided are orthogonal, so we can just normalize them to get the desired answer. Taking it step by step:

$$\begin{aligned}
A&=\frac{1}{\sqrt{50}}\begin{bmatrix}3\\4\end{bmatrix}\begin{bmatrix}1&-1\end{bmatrix}+\frac{3}{\sqrt{50}}\begin{bmatrix}-4\\3\end{bmatrix}\begin{bmatrix}1&1\end{bmatrix} &(43)\\
&=\frac{5}{\sqrt{50}}\begin{bmatrix}\frac{3}{5}\\\frac{4}{5}\end{bmatrix}\begin{bmatrix}1&-1\end{bmatrix}+\frac{3\cdot5}{\sqrt{50}}\begin{bmatrix}-\frac{4}{5}\\\frac{3}{5}\end{bmatrix}\begin{bmatrix}1&1\end{bmatrix} &(44)\\
&=\frac{5\sqrt{2}}{\sqrt{50}}\begin{bmatrix}\frac{3}{5}\\\frac{4}{5}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{bmatrix}+\frac{3\cdot5\sqrt{2}}{\sqrt{50}}\begin{bmatrix}-\frac{4}{5}\\\frac{3}{5}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{bmatrix} &(45)\\
&=1\begin{bmatrix}\frac{3}{5}\\\frac{4}{5}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{bmatrix}+3\begin{bmatrix}-\frac{4}{5}\\\frac{3}{5}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{bmatrix}. &(46)
\end{aligned}$$

From this, we can derive

$$\vec{u}_1=\begin{bmatrix}-\frac{4}{5}\\\frac{3}{5}\end{bmatrix},\vec{u}_2=\begin{bmatrix}\frac{3}{5}\\\frac{4}{5}\end{bmatrix},\vec{v}_1=\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix},\vec{v}_2=\begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix} \tag{47}$$

and corresponding singular values $\sigma_1=3,\sigma_2=1$ because we need to order them by size in decreasing order. This gives the singular value decomposition

$$A=\begin{bmatrix}-\frac{4}{5}&\frac{3}{5}\\\frac{3}{5}&\frac{4}{5}\end{bmatrix}\begin{bmatrix}3&0\\0&1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{bmatrix}^\top \tag{48}$$

(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 16, 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?

Solution: Since we have a square matrix, we will arbitrarily use $A^\top A$ for our SVD:

$$\begin{aligned}
A^\top A&=\begin{bmatrix}-1&1\\4&4\end{bmatrix}\begin{bmatrix}-1&4\\1&4\end{bmatrix} &(49)\\
&=\begin{bmatrix}2&0\\0&32\end{bmatrix} &(50)
\end{aligned}$$

Next, we find the eigenvalues of the above matrix.

$$\det(A^\top A - \lambda I) = (2-\lambda)(32-\lambda) = 0 \tag{51}$$

Hence, the eigenvalues are $\lambda_1 = 32$ and $\lambda_2 = 2$, and the singular values are $\sigma_1 = \sqrt{32} = 4\sqrt{2}$ and $\sigma_2 = \sqrt{2}$.
Next, we find the right singular vectors (i.e. the columns of V). Finding null$(A^\top A - \lambda_1 I)$ and null$(A^\top A - \lambda_2 I)$ will give us $\vec{v}_1$ and $\vec{v}_2$ respectively.
Hence, $\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (the eigenvectors are already normalized here).
Lastly, we find the right singular vectors (the columns of U)

$$\begin{aligned} \vec{u}_1 &= \frac{1}{\sigma_1} A v_1 && (52) \\ &= \frac{1}{4\sqrt{2}} \begin{bmatrix} -1 & 4 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} && (53) \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} && (54) \end{aligned}$$

Similarly, we get $\vec{u}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$.
So the full SVD representation of A is given below

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{55}$$

Now that we have found the SVD of A, we will find the eigenvalues and eigenvectors of A. Let us start with the eigenvalues:

$$\det(A - \lambda I) = (-1-\lambda)(4-\lambda) - 4 = \lambda^2 - 3\lambda - 8 = 0 \tag{56}$$

Using the quadratic formula, the eigenvalues are $\lambda_1 = \frac{3+\sqrt{41}}{2} \approx 4.7$ and $\lambda_2 = \frac{3-\sqrt{41}}{2} \approx -1.7$.

Since we already used $\vec{v}_1, \vec{v}_2$ for the SVD, let us denote the eigenvectors of A as $\vec{r}_1, \vec{r}_2$.
Finding null$(A - \lambda_1 I)$ and null$(A - \lambda_2 I)$ will give us $\vec{r}_1$ and $\vec{r}_2$ respectively.
Hence, the normalized eigenvectors of A are $\vec{r}_1 \approx \begin{bmatrix} -0.98 \\ 0.17 \end{bmatrix}$ and $\vec{r}_2 \approx \begin{bmatrix} -0.57 \\ -0.82 \end{bmatrix}$.

We notice that there is no relationship between the eigenvalues or eigenvectors of A with the SVD of A.

4. The Moore-Penrose pseudoinverse

Say we have a set of linear equations given by $A\vec{x} = \vec{y}$. If A is invertible, then the unique solution for $\vec{x}$ is $\vec{x} = A^{-1}\vec{y}$. However, what if A is not a square matrix, and we still wanted to find an $\vec{x}$ that satisfied $A\vec{x} = \vec{y}$? We know that we could use a linear least-squares approach for "tall" matrices A where it isn't possible to find a solution that exactly matches all the measurements. The linear least-squares solution gives us a reasonable answer that asks for the "best" match in terms of reducing the norm of the error vector.

How about when the matrix A is "wide", i.e. A has more columns than rows? In this case, there are generally going to be lots of possible solutions — so which should we choose? To address this, we introduce the *Moore-Penrose pseudoinverse*, which generalizes the idea of the matrix inverse and can be calculated using the singular value decomposition.

Since the SVD of a matrix always exists, the Moore-Penrose pseudoinverse does as well. Another useful property of the Moore-Penrose pseudoinverse $A^{\dagger}$ is that the solution it gives, $\hat{\vec{x}} = A^{\dagger}\vec{y}$, satisfies a minimality property: $\|\hat{\vec{x}}\| \leq \|\vec{z}\|$ for all $\vec{z}$ such that $A\vec{z} = \vec{y}$.

(a) Say we have the matrix

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

To find the Moore-Penrose pseudoinverse we start by calculating the SVD of A. That is to say, we find orthonormal matrices U and V, and diagonal matrix Σ, such that $A = U\Sigma V^{\top}$.
Here we give you the decomposition of A as:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{57}$$

where:

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{58}$$

$$\Sigma = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \tag{59}$$

$$V^{\top} = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{60}$$

It is a good idea to be able to calculate the SVD yourself as you may be asked to solve similar questions on your own in the exam. **You do not have to do any work for this part.**

Solution: Though you did not have to do any work for deriving the SVD the following solutions will walk you through how to solve for the SVD:

$$A = U\Sigma V^{\top} \tag{61}$$

$$AA^{\top} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}. \tag{62}$$

Which has characteristic polynomial $\lambda^2 - 6\lambda + 8 = 0$, producing eigenvalues 4 and 2. Solving $Av = \lambda_i v$ produces eigenvectors $[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}]^{\top}$ and $[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^{\top}$ associated with eigenvalues 4 and 2 respectively. The singular values are the square roots of the eigenvalues of $AA^{\top}$, so

$$\Sigma = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \tag{63}$$

and

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{64}$$

We can then solve for the $\vec{v}$ vectors using $A^\top \vec{u}_i = \sigma_i \vec{v}_i$, producing $\vec{v}_1 = [0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^\top$ and $\vec{v}_2 = [1, 0, 0]^\top$. The last $\vec{v}$ must be orthonormal to the other two, so we can pick $[0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^\top$. The SVD is:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{65}$$

(b) Suppose we have non-zero singular values $\sigma_1, \ldots, \sigma_r$, and that we have written the SVD matrices so that Σ is in the form

$$\Sigma = \underbrace{\begin{bmatrix} \sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix}}_{\text{Dimension: } m \times n}. \tag{66}$$

Consider the action of Σ on $\vec{v} \in \mathbb{R}^n$, i.e. $\Sigma\vec{v}$. **What is the effect of Σ on each element of $\vec{v}$?** Let us define the following matrix:

$$\widetilde{\Sigma} = \underbrace{\begin{bmatrix} \frac{1}{\sigma_1} & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \frac{1}{\sigma_2} & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{\sigma_r} & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix}}_{\text{Dimension: } n \times m}. \tag{67}$$

What is $\widetilde{\Sigma}\Sigma$? What is the effect of $\widetilde{\Sigma}\Sigma$ on $\vec{v} \in \mathbb{R}^n$?

Solution: As described in the problem, we can represent $\vec{v}$ as $\vec{v} = [v_1, v_2, ..., v_n]^T$. Then, we have

$$\Sigma\vec{v} = [\sigma_1 v_1, \sigma_2 v_2, ..., \sigma_r v_r, 0, 0, ..., 0]. \tag{68}$$

Therefore, the effect of Σ is to scale the i-th element of $\vec{v}$: v_i by σ_i, when $i \leq r$. When $i > r$, Σ wipes out the original value with a 0.

Following the matrix multiplication, we have

$$\widetilde{\Sigma}\Sigma = \underbrace{\begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix}}_{\text{Dimension: } n \times n}, \tag{69}$$

where $\widetilde{\Sigma}\Sigma$ is a diagonal matrix with the first r diagonal values equal to **1**, and **0** for the rest of the elements. Therefore, $\widetilde{\Sigma}\Sigma\vec{v} = [v_1, v_2, ..., v_r, 0, ..., 0]^T$. This operation keeps the first r values of $\vec{v}$, and turns the rest to **0**.

(c) Consider when $A = U\Sigma V^\top$ acts on $\vec{x}$ to give the result $\vec{y}$, i.e.

$$A\vec{x} = U\Sigma V^\top \vec{x} = \vec{y}. \tag{70}$$

Observe that $V^\top \vec{x}$ rotates $\vec{x}$ without changing its length, and U rotates $\Sigma V^\top \vec{x}$ again. The *Moore-Penrose pseudoinverse* $A^\dagger$ is given as

$$A^\dagger = V\widetilde{\Sigma}U^\top, \tag{71}$$

where $\widetilde{\Sigma}$ is given in (c). Consider if we apply the Moore-Penrose pseudoinverse to find a candidate solution $A^\dagger y$:

$$\vec{y} = U\Sigma V^\top \vec{x} \tag{72}$$

$$A^\dagger y = (V\widetilde{\Sigma}U^\top)(U\Sigma V^\top)\vec{x}. \tag{73}$$

Qualitatively, what are the effects of the matrices V, $\widetilde{\Sigma}$, and $U^\top$ in the Moore-Penrose pseudoinverse when finding a solution?

Solution: Given

$$A^\dagger y = V\widetilde{\Sigma}U^\top(U\Sigma V^\top)\vec{x}, \tag{74}$$

and $U^\top U = I$, we know that $U^\top$ "undoes" the rotation effect of U and brings the vector $U\Sigma V^\top \vec{x}$ back to $\Sigma V^\top \vec{x}$.

Now we have

$$A^\dagger y = V\widetilde{\Sigma}U^\top(U\Sigma V^\top)\vec{x} \tag{75}$$

$$= V\widetilde{\Sigma}\Sigma V^\top \vec{x}. \tag{76}$$

$\widetilde{\Sigma}$ scales the first r values of $\Sigma V^\top \vec{x}$ by $\frac{1}{\sigma_i}$, for $i = 1, 2, 3, ..., r$, and returns **0** for the rest. Meanwhile, as shown in part (b), we know that $\widetilde{\Sigma}\Sigma$ keeps the first r values of $V^\top \vec{x}$.

Finally, since $VV^\top = I$, matrix V again "undoes" the rotation effect of $V^\top$ without changing its length.

(d) **What does the Moore-Penrose pseudoinverse give as a solution $\vec{x}$ in the following system of equations?**

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Confirm that your solution indeed satisfies the system of equations.

Solution: From the above, we have the solution given by:

$$\vec{x} = A^\dagger \vec{y} = V\widetilde{\Sigma}U^\top \vec{y} \tag{77}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \tag{78}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \tag{79}$$

Therefore, a reasonable solution to the system of equations is:

$$\vec{x} = \begin{bmatrix} 3 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \tag{80}$$

Confirming that the solution works, we have

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}. \tag{81}$$

5. **(OPTIONAL) Make Your Own Problem.**

   Write your own problem about content covered in the course thus far, and provide a thorough solution to it.

   NOTE: This can be a totally new problem, a modification on an existing problem, or a Jupyter part for a problem that previously didn't have one. Please cite all sources for anything (including course material) that you used as inspiration.

   NOTE: High-quality problems may be used as inspiration for the problems we choose to put on future homeworks or exams.

6. **Homework Process and Study Group**

   Citing sources and collaborators are an important part of life, including being a student!

   We also want to understand what resources you find helpful and how much time homework is taking, so we can change things in the future if possible.

   (a) **What sources (if any) did you use as you worked through the homework?**

   (b) **If you worked with someone on this homework, who did you work with?**
   List names and student ID's. (In case of homework party, you can also just describe the group.)

   (c) **Roughly how many total hours did you work on this homework? Write it down here where you'll need to remember it for the self-grade form.**

Contributors:

- Siddharth Iyer.
- Yu-Yun Dai.
- Sanjit Batra.
- Anant Sahai.
- Sidney Buchbinder.
- Gaoyue Zhou.
- Druv Pai.
- Elena Jia.
- Tanmay Gautam.
- Justin Yim.
- Sally Hui.

