

1. Prove the following for $x \in \mathbb{C}^n$:

(a) $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$.

Proof:

For the first inequality, we have

$$\begin{aligned}\|x\|_\infty &= \max_{1 \leq i \leq n} |x_i| \\ &\leq \sum_{i=1}^n |x_i| \\ &= \|x\|_1.\end{aligned}$$

For the second half of our inequality chain, we have

$$\begin{aligned}\|x\|_1 &= \sum_{i=1}^n |x_i| \\ &\leq \sum_{i=1}^n \left(\max_{1 \leq j \leq n} |x_j| \right) \\ &= n \max_{1 \leq j \leq n} |x_j| \\ &= n\|x\|_\infty\end{aligned}$$

Thus, $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$. □

(b) $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$.

Proof:

For the first inequality, we have

$$\begin{aligned}\|x\|_\infty &= \max_{1 \leq i \leq n} |x_i| \\ &= \sqrt{\left(\max_{1 \leq i \leq n} |x_i| \right)^2} \\ &\leq \sqrt{\sum_{i=1}^n |x_i|^2} \\ &= \|x\|_2.\end{aligned}$$

For the second half of our inequality chain, we have

$$\begin{aligned}
 \|x\|_2 &= \sqrt{\sum_{i=1}^n |x_i|^2} \\
 &\leq \sqrt{\sum_{i=1}^n \left(\max_{1 \leq j \leq n} |x_j| \right)^2} \\
 &= \sqrt{n \left(\max_{1 \leq j \leq n} |x_j| \right)^2} \\
 &= \sqrt{n} \max_{1 \leq j \leq n} |x_j| \\
 &= n \|x\|_\infty.
 \end{aligned}$$

Thus, $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$. □

(c) $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2$.

Proof:

Let's work on the first half of the inequality:

$$\begin{aligned}
 \|x\|_2 &= \sqrt{\sum_{i=1}^n |x_i|^2} \\
 &\leq \sqrt{\sum_{i=1}^n |x_i|^2 + \sum_{i,j,i \neq j} |x_i| |x_j|} \\
 &= \sqrt{\left(\sum_{i=1}^n |x_i| \right) \left(\sum_{j=1}^n |x_j| \right)} \\
 &= \sqrt{\left(\sum_{i=1}^n |x_i| \right)^2} \\
 &= \sum_{i=1}^n |x_i| \\
 &= \|x\|_1.
 \end{aligned}$$

Now, let's show the second half of the inequality

$$\begin{aligned}
 \|x\|_1 &= \sum_{i=1}^n |x_i| \\
 &= \sum_{i=1}^n |1| |x_i| \\
 &\leq \|\vec{1}\|_2 \|x\|_2 \quad \text{by Cauchy-Schwartz} \\
 &= \left(\sum_{i=1}^n |1|^2 \right) \|x\|_2 \\
 &= \sqrt{n} \|x\|_2.
 \end{aligned}$$

Thus, $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$. □

2. Let $A \in \mathbb{R}^{n \times m}$ be a non-zero matrix with rank r .

- (a) Write down the singular value decomposition of A . List the properties of the matrices you use in your decomposition.

We can write the SVD of A as

$$A = U\Sigma V^T$$

Where $U \in \mathbb{R}^{n \times n}$ is made up of the left singular vectors, $\Sigma \in \mathbb{R}^{n \times m}$ is a semi-diagonal matrix with positive singular values along the main diagonal, and $V \in \mathbb{R}^{m \times m}$ is made up of the right singular vectors.

Some properties of U

- it is orthonormal
- the columns of U are made up of the normalized eigenvectors of AA^T (i.e. $U = [u_1|u_2|\dots|u_n]$).

Some properties of V

- it is orthonormal
- the columns of V are made up of the normalized eigenvectors of A^TA (i.e. $V = [v_1|v_2|\dots|v_m]$).

Finally, some properties of Σ

- the diagonal entries of Σ are given by the singular values σ_i for $i = 1, 2, \dots, r$. The rest of the entries of Σ are zero.
- by convention, we organize the singular values from the left diagonal to right diagonal as $\sigma_1 \geq \dots \geq \sigma_r \geq 0$.

- (b) Show that \mathbb{R}^m has an orthonormal basis v_1, \dots, v_m , \mathbb{R}^n has an orthonormal basis u_1, \dots, u_n and there exists $\sigma_1 \geq \dots \geq \sigma_r \geq 0$ such that

$$Av_i = \begin{cases} \sigma_i u_i, & i = 1, \dots, r \\ 0, & i = r + 1, \dots, m \end{cases},$$

$$A^T u_i = \begin{cases} \sigma_i v_i, & i = 1, \dots, r \\ 0, & i = r + 1, \dots, m \end{cases}.$$

First off, because we are guaranteed a *SVD* for all $A \in \mathbb{R}^{n \times m}$, we can get our decomposition given in part (a). Using our SVD from part (a), we know the columns of V form an orthonormal basis for \mathbb{R}^m given by v_1, \dots, v_m , and the columns of U form an orthonormal basis for \mathbb{R}^n given by u_1, \dots, u_n . So we can get the two orthonormal bases that we desired. Now, using the SVD in part (a), we have $\sigma_1 \geq \dots \geq \sigma_r \geq 0$ from the singular values matrix Σ . Putting all of this together,

$$\begin{aligned} Av_i &= U\Sigma V^T v_i = \begin{cases} \sigma_i u_i (v_i \cdot v_i), & i = 1, \dots, r \\ 0, & i = r + 1, \dots, m \end{cases} \\ &= \begin{cases} \sigma_i u_i, & i = 1, \dots, r \\ 0, & i = r + 1, \dots, m \end{cases} \end{aligned}$$

because $v_i \cdot v_j = 0$ if $i \neq j$ and $v_i \cdot v_j = 1$ if $i = j$ by the orthonormalness of V .

For the second equality, we have

$$\begin{aligned} A^T u_i &= (U \Sigma V^T)^T u_i = V \Sigma^T U^T u_i = \begin{cases} \sigma_i v_i (u_i \cdot u_i), & i = 1, \dots, r \\ 0, & i = r+1, \dots, n \end{cases} \\ &= \begin{cases} \sigma_i v_i, & i = 1, \dots, r \\ 0, & i = r+1, \dots, n \end{cases} \end{aligned}$$

because $u_i \cdot u_j = 0$ if $i \neq j$ and $u_i \cdot u_j = 1$ if $i = j$ by the orthonormalness of U .

(c) Argue that

$$\text{Range}(A) = \text{span}\{u_1, \dots, u_r\} \quad (1)$$

$$\text{Null}(A) = \text{span}\{v_{r+1}, \dots, v_m\} \quad (2)$$

$$\text{Range}(A^T) = \text{span}\{v_1, \dots, v_r\} \quad (3)$$

$$\text{Null}(A^T) = \text{span}\{u_{r+1}, \dots, u_n\}. \quad (4)$$

- To understand (1), suppose we have any $x \in \mathbb{R}^m$. Then, because we only have r singular values and the rest of Σ is either rows of zeros or columns of zeros, we have

$$Ax = U \Sigma V^T x = \sum_{i=1}^r \sigma_i u_i v_i^T x = \sum_{i=1}^r (\sigma_i v_i^T x) u_i = \sum_{i=1}^r a_i u_i$$

where $a_i = \sigma_i v_i^T x$ for $i = 1, \dots, r$ are constants. Thus, Ax can be written as a linear combination of $\{u_1, \dots, u_r\}$ for all $x \in \mathbb{R}^m$ which implies

$$\text{Range}(A) = \text{span}(u_1, \dots, u_r).$$

- To argue (2), let's find the nullspace previous part:

$$Ax = \sum_{i=1}^r a_i u_i$$

where $a_i = \sigma_i v_i^T x$ for $i = 1, \dots, r$. From this expression, we can see that nullspace is the set of all x such that $a_i = 0$ for $i = 1, \dots, r$. In other words, we need x to satisfy

$$\sigma_i v_i^T x = 0$$

for each $i = 1, \dots, r$. We know each $\{v_1, \dots, v_m\}$ is an orthogonal set and so the only x that make $a_i = 0$ are $x = b_{r+1} v_{r+1} + \dots + b_m v_m$ which implies that

$$\text{Null}(A) = \text{span}(v_{r+1}, \dots, v_m).$$

- Similar to (1), we can show (3). Suppose we have any $x \in \mathbb{R}^n$. Then,

$$A^T x = V \Sigma^T U^T x = \sum_{i=1}^r \sigma_i v_i u_i^T x = \sum_{i=1}^r (\sigma_i u_i^T x) v_i = \sum_{i=1}^r a_i v_i$$

where $a_i = \sigma_i u_i^T x$ for $i = 1, \dots, r$ constants. Thus, $A^T x$ can be written as a linear combination of $\{v_1, \dots, v_r\}$ for all $x \in \mathbb{R}^n$ which implies

$$\text{Range}(A^T) = \text{span}(v_1, \dots, v_r).$$

- We can argue (4) by applying the argument for (2) to (3) with the roles of u_i and v_i switched.

(d) Now show that $\text{Range}(A^T)$ is orthogonal to $\text{Null}(A)$.

Suppose we have any $x \in \text{Range}(A^T)$ and any $y \in \text{Null}(A)$. Then, from part (c), we know that

$$x = a_1 v_1 + \cdots + a_r v_r$$

for some $a_i \in \mathbb{R}$ and that

$$y = b_{r+1} v_{r+1} + \cdots + b_m v_m$$

for some $b_i \in \mathbb{R}$. Then,

$$\begin{aligned} x \cdot y &= \sum_{i=r+1}^m x \cdot (b_i v_i) \\ &= \sum_{i=r+1}^m \sum_{j=1}^r (a_j v_j) \cdot (b_i v_i) \\ &= \sum_{i=r+1}^m \sum_{j=1}^r a_j b_i (v_j \cdot v_i) \\ &= \sum_{i=r+1}^m \sum_{j=1}^r a_j b_i (0) \quad \text{because } i \neq j \text{ and } v_j \cdot v_i = 0 \text{ for } i \neq j \\ &= 0. \end{aligned}$$

Therefore, $\text{Range}(A^T)$ is orthogonal to $\text{Null}(A)$.

3. Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and $u, v \in \mathbb{R}^n$.

(a) Prove the following matrix identity (Sherman-Morrison)

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u}.$$

Proof:

To prove the Sherman-Morrison formula, we just need to show the RHS is the inverse of $A + uv^T$.

- First direction

$$\begin{aligned} \left(A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u} \right) (A + uv^T) &= A^{-1}A + A^{-1}uv^T - \frac{A^{-1}uv^T A^{-1}A + A^{-1}uv^T A^{-1}uv^T}{1 + v^T A^{-1}u} \\ &= I + A^{-1}uv^T - \frac{A^{-1}uv^T + A^{-1}uv^T A^{-1}uv^T}{1 + v^T A^{-1}u} \\ &= I + A^{-1}uv^T - \frac{A^{-1}u(1 + v^T A^{-1}u)v^T}{1 + v^T A^{-1}u} \\ &= I + A^{-1}uv^T - A^{-1}uv^T \\ &= I. \end{aligned}$$

- The second direction can be shown as

$$\begin{aligned}
 (A + uv^T) \left(A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u} \right) &= AA^{-1} + uv^T A^{-1} - \frac{AA^{-1}uv^T A^{-1} + uv^T A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u} \\
 &= I + uv^T A^{-1} - \frac{uv^T A^{-1} + uv^T A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u} \\
 &= I + uv^T A^{-1} - \frac{u(1 + v^T A^{-1}u)v^T A^{-1}}{1 + v^T A^{-1}u} \\
 &= I + uv^T A^{-1} - uv^T A^{-1} \\
 &= I.
 \end{aligned}$$

Therefore

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u}$$

□

- (b) Suppose that the LU factorization of A is available. Explain how the Sherman-Morrison identity can be used to solve the system $(A + uv^T)x = b$.

Firstly, let's solve for x directly to get

$$x = (A + uv^T)^{-1}b.$$

Now using the Sherman-Morrison identity, we have

$$x = A^{-1}b - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u}b.$$

With this expanded expression for x , let's break the expression up into chunks that we can use the LU factorization of A on:

$$x = y - \frac{zv^T y}{1 + v^T z} = y - \frac{z(v^T y)}{1 + v^T z}$$

where $y = A^{-1}b$ and $z = A^{-1}u$. Thus, to solve for x , I would use the LU of A to solve for y and z and then plug those into the rest of the expression to finish up the computation of x according to the parenthesis.

- (c) What is the cost of the proposed method for solving for x in part (b)?

To count the number of FLOPs to compute x in part (b), we need to first compute the cost of computing y and z using the LU factorization of A . The cost for y and z is each $2n^2 - n$. Then, we need to use this calculation to compute the rest of x . To do so, we need n subtractions for the main difference, $2(2n - 1)$ for the two dot products, n for the vector scalar multiplication, n for the division, and finally, 1 flop for the addition in the denominator. Adding all of these up yields a total cost of

$$\text{cost} = 4n^2 + 5n - 1 = O(n^2) \text{ FLOPs.}$$