

## Question 1

Let  $A$  be a nonsingular  $n \times n$  matrix with real entries and  $\vec{\mathbf{b}} \in \mathbb{R}^n$ .

Let  $A = P\Sigma Q^\top$  be the singular value decomposition (SVD) of  $A$ . Then, the system  $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$  can be written as:

$$\begin{aligned} A\vec{\mathbf{x}} = \vec{\mathbf{b}} &\iff (P\Sigma Q^\top)\vec{\mathbf{x}} = \vec{\mathbf{b}} \\ &\iff \vec{\mathbf{x}} = (P\Sigma Q^\top)^{-1}\vec{\mathbf{b}} \\ &\iff \vec{\mathbf{x}} = (Q^\top)^{-1}\Sigma^{-1}P^{-1}\vec{\mathbf{b}} \end{aligned}$$

Note that since  $A$  is a real nonsingular matrix, the matrices  $P = \begin{pmatrix} | & & | \\ \vec{\mathbf{p}}_1 & \cdots & \vec{\mathbf{p}}_n \\ | & & | \end{pmatrix}$  and  $Q = \begin{pmatrix} | & & | \\ \vec{\mathbf{q}}_1 & \cdots & \vec{\mathbf{q}}_n \\ | & & | \end{pmatrix}$  are orthogonal. Thus,  $Q^\top Q = QQ^\top = I \iff Q^\top = Q^{-1}$  and  $P^\top P = PP^\top = I \iff P^\top = P^{-1}$ . As such, the solution can be written as:

$$\begin{aligned} \vec{\mathbf{x}} &= Q\Sigma^{-1}P^\top\vec{\mathbf{b}} \\ &= \begin{pmatrix} | & & | \\ \vec{\mathbf{q}}_1 & \cdots & \vec{\mathbf{q}}_n \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}^{-1} \begin{pmatrix} -\vec{\mathbf{p}}_1^\top & - \\ \vdots & \\ -\vec{\mathbf{p}}_n^\top & - \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \end{aligned}$$

Here, we're basically solving for  $\vec{\mathbf{x}}$  by computing the inverse of  $A$  with its SVD. Thus, it is crucial that  $A$  is nonsingular. Otherwise,  $A$  would be invertible. Additionally, note that the solution requires us to compute the inverse of  $\Sigma$ , which we note  $\Sigma^{-1} = \text{diag}(1/\sigma_1, \dots, 1/\sigma_n)$ . This means that all singular values of  $A$ ,  $\sigma_1, \dots, \sigma_n$  must be all positive. If  $A$  is singular, then its determinant is zero, which means zero is a singular value of  $A$ , and thus  $\Sigma$  would also be invertible.

## Question 2

From Question 1, we solved the system  $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$  by substituting the SVD of  $A$  and computing its inverse. Where  $A = P\Sigma Q^\top$  is the SVD of  $A$ , the resulting expression is  $\vec{\mathbf{x}} = Q\Sigma^{-1}P^\top\vec{\mathbf{b}}$ . Thus,  $A^{-1} = Q\Sigma^{-1}P^\top$  and as demonstrated in the previous question, the singular value of  $A^{-1}$  is simply the reciprocal of the singular values of  $A$ .

## Question 3

Let  $A$  be a real-valued  $m \times n$  matrix and let  $\|A\| = \sqrt{\sum_{j=1}^n \sum_{i=1}^m a_{ij}^2}$ .

**Part A**

Let  $B$  be a  $p \times r$  matrix and  $C$  be a  $r \times p$  matrix. Then,  $BC$  is a  $p \times p$  matrix and  $CB$  is an  $r \times r$  matrix.

The trace of an  $n \times n$  square matrix  $M$  is the sum of the entries on the main diagonal, that is  $\text{trace}(M) = \sum_{i=1}^n m_{ii}$ . As such,  $\text{trace}(BC) = \sum_{i=1}^p \sum_{j=1}^r b_{ij}c_{ij}$  and  $\text{trace}(CB) = \sum_{i=1}^r \sum_{j=1}^p c_{ij}b_{ij}$ . By commutativity, we can clearly see that  $\text{trace}(BC) = \text{trace}(CB)$ .

**Part B**

Similar to 3(a), we note that  $AA^\top$  is an  $m \times m$  matrix and  $A^\top A$  is an  $n \times n$  matrix. As such,  $\text{trace}(AA^\top) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}a_{ij} = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$  and  $\text{trace}(A^\top A) = \sum_{i=1}^n \sum_{j=1}^m a_{ij}a_{ij} = \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2$ . Again by commutativity, we see that  $\text{trace}(AA^\top) = \text{trace}(A^\top A) = \|A\|^2$ .

**Part C**

Let  $U$  be an  $m \times m$  orthogonal matrix. Then,  $U^\top U = I_m$ . From 3(b), we note that  $\|A\|^2 = \text{trace}(A^\top A) = \text{trace}(AA^\top)$ . Then,

$$\begin{aligned} \|UA\|^2 &= \text{trace}((UA)^\top(UA)) \\ &= \text{trace}(A^\top U^\top U A) \\ &= \text{trace}(A^\top A) \\ &= \|A\|^2. \end{aligned}$$

As such, it follows that  $\|UA\| = \|A\|$ .

**Part D**

Let  $A = P\Sigma Q^\top$  be the singular value decomposition SVD of  $A$ , where  $P = \begin{pmatrix} | & & | \\ \vec{\mathbf{p}}_1 & \cdots & \vec{\mathbf{p}}_m \\ | & & | \end{pmatrix}$  and  $Q = \begin{pmatrix} | & & | \\ \vec{\mathbf{q}}_1 & \cdots & \vec{\mathbf{q}}_n \\ | & & | \end{pmatrix}$  are orthogonal matrices. And  $\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}$  is the diago-

nal matrix where  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$  are the singular values of  $A$ . Then,

$$\begin{aligned}
 \|A\|^2 &= \text{trace}(A^\top A) \\
 &= \text{trace}((P\Sigma Q^\top)^\top (P\Sigma Q^\top)) \\
 &= \text{trace}((Q^\top)^\top \Sigma^\top P^\top P \Sigma Q^\top) \\
 &= \text{trace}(Q \Sigma^\top \Sigma Q^\top) \\
 &= \text{trace}(Q \Sigma^2 Q^\top) && \because \Sigma^\top = \Sigma \\
 &= \text{trace}(\Sigma^2) && \because Q Q^\top = I \\
 &= \sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2.
 \end{aligned}$$

As such, it follows that  $\|A\| = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2}$ .

## Question 4

### Part A

Let  $A$  be an  $m \times n$  real-valued matrix with rank  $r$  and let  $A = P\Sigma Q^\top$  be its singular value decomposition. Again, let  $P = \begin{pmatrix} | & & | \\ \vec{p}_1 & \dots & \vec{p}_r \\ | & & | \end{pmatrix}$ ,  $\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}$ , and

$Q^\top = \begin{pmatrix} - & \vec{q}_1^\top & - \\ & \vdots & \\ - & \vec{q}_r^\top & - \end{pmatrix}$ . As demonstrated in class, the SVD of  $A$  can be decomposed into the following:

$$\begin{aligned}
 A &= P\Sigma Q^\top \\
 &= \sigma_1 \vec{p}_1 \vec{q}_1^\top + \sigma_2 \vec{p}_2 \vec{q}_2^\top + \dots + \sigma_r \vec{p}_r \vec{q}_r^\top,
 \end{aligned}$$

where each term is an  $m \times n$  matrix of rank 1 and  $\text{rank}(A)$  is the minimum number of rank-1 matrices  $A_1, A_2, \dots, A_r$  such that  $A = A_1 + A_2 + \dots + A_r$ .

### Part B

Let  $A_k = P_k \Sigma_k Q_k^\top = \sigma_1 \vec{p}_1 \vec{q}_1^\top + \sigma_2 \vec{p}_2 \vec{q}_2^\top + \dots + \sigma_r \vec{p}_r \vec{q}_r^\top$ . Then,

$$A - A_k = \sigma_{k+1} \vec{p}_{k+1} \vec{q}_{k+1}^\top + \sigma_{k+2} \vec{p}_{k+2} \vec{q}_{k+2}^\top + \dots + \sigma_r \vec{p}_r \vec{q}_r^\top.$$

As such, the  $\Sigma_k = \text{diag}(\sigma_{k+1}, \sigma_{k+2}, \dots, \sigma_r)$ . Thus, from 3(d) we have that

$$\|A_k\| = \sqrt{\text{trace}(\Sigma_k^2)} = \sqrt{\sigma_{k+1}^2 + \sigma_{k+2}^2 + \dots + \sigma_r^2}.$$

## Question 5

Refer to Section %% Question 5 in the math425hw7.m file for the relevant code.