CS365 Spring '23
Foundations of Data Science
Prof. C.E. Tsourakakis Assignment 8

Instructions

- The homework is due on **Friday 4/14 at 5pm ET**.
- No extension will be provided, unless for serious documented reasons.
- Start early!
- Study the material taught in class, and feel free to do so in small groups, but the solutions should be a product of your own work.
- This is not a multiple choice homework; reasoning, and mathematical proofs are required before giving your final answer.

1 [15 points]

Let $A \in \mathbf{R}^{m\times n}$ be a real $m \times n$ matrix.

$m\times m$ $n\times n$

1. (5pts) Prove that the eigenvalues of AA^T and A^TA are real and non-negative.
2. (5pts) Prove that the two matrices have the same set of non-negative eigenvalues.
3. (5pts) How does this set of eigenvalues relates to the set of singular values? What about the left, right singular vectors with respect to the eigenvectors of the matrices AA^T and A^TA?

2 [20 points]

1. (5 pts) Let $A^{n\times n}$ be a real square matrix. Suppose the rows of A are orthonormal. Prove that the columns have to be orthonormal. Is this statement true when the matrix is not square?
2. (5 pts) Prove that a linear system $Ax = b$ is consistent if and only if $rank(A) = rank([A|b])$. Comment on the geometric interpretation of equation $rank(A) = rank([A|b])$.
3. (5 pts + 5 pts) What is the SVD of the matrix $M = [0, 1, 2]^{1\times 3}$? Compute it in two ways:
   (a) Using exercise 1.3.
   (b) By "eyeballing" M.
   Hint: Understand the subspaces spanned by the columns and rows in order to decide the left and singular vectors.

3 SVD for least squares [20 points]

Suppose you are given a system of linear equations $A^{m\times n}x^{n\times 1} = b^{m\times 1}$ where the number of rows m is greater than the number of columns n (overdetermined system of linear equations). Given that the number of equations m is greater than the number of unknowns maybe there is no x that satisfies the linear system. Thus it is natural to try to find an x that minimizes the error $||Ax - b||_2$.

1. (10 pts) Assume that A is full rank, i.e., $rank(A) = n < m$. Prove that the unique minimizer $x^\star = (A^TA)^{-1}A^Tb$. Be explicit about where you use the assumption that A is full rank and the objective value $||Ax^\star - b||_2$.

2. (10 pts) Solve the same optimization problem when A is rank deficient.

   Hint: Use the SVD decomposition

4 Coding [30 points]

Check the Jupyter notebook on our Git repo.

1.

1) $\because$ AA^T, A^TA are symmetric

$\therefore$ λ is real.

For AA^T: $\lambda = ||v||^2\lambda = v^T\lambda v = v^TAA^Tv = (A^Tv)^TA^Tv = ||A^Tv||^2 \geq 0$

A^TA: $\lambda = ||v||^2\lambda = v^T\lambda v = v^TA^TAv = (Av)^TAv = ||Av||^2 \geq 0$

$\therefore$ $\boxed{\lambda \geq 0}$

2)
Proof:

suppose $x \in \mathbb{R}$, $\lambda = A^TA$ such that

$$(A^TA)x = \lambda x$$
$$A(A^TA)x = A\lambda x$$
$$AA^TAx = A\lambda x = \lambda Ax$$
$$(AA^T)Ax = \lambda\cdot(Ax), \quad \lambda, x \neq 0$$
$$\Rightarrow \quad \lambda = AA^T$$

Thus, they have the same non-negative eigenvalues.

3)

$\sigma = \sqrt{\lambda}$, where λ is eigenvalues of A^TA

$A = U\Sigma V^T$

U: left singular vectors. $\Rightarrow$ Eigenvectors of AA^T make up columns of U.

V: right singular vectors. $\Rightarrow$ Eigenvectors of A^TA make up columns of V.

2.

1) Since A is a square matrix, then $AA^T = A^TA$.

We've known that A's rows are orthonormal, so

According to the definition, $U^TU = I$ iff U has orthonormal column.

$\Rightarrow (A^T)^T A^T = AA^T = I$

$\because AA^T = A^TA$

$\therefore A^TA = I \quad \therefore A$ has orthonormal columns.

when the matrix is not square:

$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} - b_1 - \\ - b_2 - \end{bmatrix}$ $\quad \|b_1\| = 1,\ \|b_2\| = 1, \quad b_1 b_2^T = 0$

$\therefore$ rows are orthonormal

$B = \begin{bmatrix} | & | & | \\ b_1' & b_2' & b_3' \\ | & | & | \end{bmatrix}$ $\quad \|b_1'\| = 1,\ \|b_2'\| = 1 \quad \|b_3'\| \neq 1$

$\therefore$ columns are NOT orthonormal.

Thus, the statement is false when the matrix isn't square.

2) $\text{Rank}(A) = p$, it means vectors in A form a p-dimension space. If $\text{Rank}([A|b]) = p$, it means all vectors in A and the vector b all lie in the span of column of A (which is the p-dimensional space).

According to the definition, $Ax = b$ is consistent if and only if b lies in the span of col A. Therefore, we can say that if $\text{Rank}(A) = \text{Rank}([A|b])$, then $Ax = b$ is consistent.

If $Ax = b$ is consistent, it means there exists x such that $x_1 \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + \cdots + x_n \begin{bmatrix} | \\ a_n \\ | \end{bmatrix} = b$. Suppose $\text{Rank}(A) = p$, then

$A = \begin{bmatrix} - A_1 - \\ \vdots \\ - A_p - \\ \cdots 0 \cdots \\ \cdots 0 \cdots \end{bmatrix} \Rightarrow b = \begin{bmatrix} b_1 \\ \vdots \\ b_p \\ 0 \\ \vdots \end{bmatrix}$ according to the linear combination.

Therefore, $\text{Rank}([A|b]) = p = \text{Rank}(A)$.

Thus, if $Ax = b$ is consistent, then $\text{Rank}(A) = \text{Rank}([A|b])$

$\Rightarrow Ax = b$ is consistent $\Longleftrightarrow \text{Rank}(A) = \text{Rank}([A|b])$ ∎

3) $M = [0 \ 1 \ 2]$

a) $M^TM = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} [0 \ 1 \ 2] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix}$

$\det(M^TM - \lambda I) = 0$

$\det \begin{bmatrix} -\lambda & 0 & 0 \\ 0 & 1-\lambda & 2 \\ 0 & 2 & 4-\lambda \end{bmatrix} = 0$

$-\lambda \cdot \begin{vmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 2 \\ 0 & 4-\lambda \end{vmatrix} + 0 \cdot \begin{vmatrix} 0 & 1-\lambda \\ 0 & 2 \end{vmatrix} = 0$

$(-\lambda)[(1-\lambda)(4-\lambda) - 4] = 0$

$(-\lambda)(4 - 5\lambda + \lambda^2 - 4) = 0$

$(-\lambda)(\lambda^2 - 5\lambda) = 0$

$\lambda_1 = 5, \ \lambda_2 = 0, \ \lambda_3 = 0$

$\lambda = 5$

$M^TM - I = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -4 & 2 \\ 0 & 2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

$(M^TM - I)x = 0$

$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{l} x_1 = 0 \\ 2x_2 - x_3 = 0 \end{array}$

$\Rightarrow$ general solution: $\begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \end{bmatrix} \cdot x_3$

$V_1 = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}$, $\|V_1\| = \sqrt{\frac{1}{4} + 1} = \frac{\sqrt{5}}{2}$, $V_1 = \begin{bmatrix} 0 \\ \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \end{bmatrix}$

$\lambda = 0$

$M^TM - I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

$(M^TM - I)x = 0$

$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{l} x_1 \text{ free} \\ x_2 + 2x_3 = 0 \\ x_2 = -2x_3 \end{array}$

$\Rightarrow$ general solution:

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} x_3$

$\therefore V_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $V_3 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$

$\|V_2\| = 1$ $\quad \|V_3\| = \sqrt{5}$

$V_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\quad V_3 = \begin{bmatrix} 0 \\ \frac{-2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \end{bmatrix}$

$\Sigma = [\sqrt{5} \ 0 \ 0]$

$V = \begin{bmatrix} 0 & 1 & 0 \\ \frac{\sqrt{5}}{5} & 0 & \frac{-2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & 0 & \frac{\sqrt{5}}{5} \end{bmatrix}$

$U_i = \frac{1}{\sigma} M V_i = \frac{1}{\sqrt{5}} [0, 1, 2] \begin{bmatrix} 0 \\ \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \end{bmatrix} = 0 + \frac{1}{\sqrt{5}} \cdot \frac{\sqrt{5}}{5} + \frac{2}{\sqrt{5}} \cdot \frac{2\sqrt{5}}{5}$

$= 1 = U$

$\therefore$ SVD: $M = [1] \cdot [\sqrt{5} \ 0 \ 0] \begin{bmatrix} 0 & \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ 1 & 0 & 0 \\ 0 & \frac{-2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}$

b) $M = [0\ 1\ 2]$ $\quad M^T = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

$\Sigma \Rightarrow \det\begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{vmatrix} \Rightarrow \begin{matrix} \lambda_1 = 5 \\ \sigma_1 = \sqrt{5} \end{matrix} \Rightarrow \Sigma = [\sqrt{5}, 0, 0]$

$U \begin{cases} [u_1 \cdots u_r] = \text{Rank}(M) \\ [u_{r+1} \cdots u_m] = \text{Null}(M^T) \end{cases} \quad V \begin{cases} [V_1 \cdots V_r] = \text{Rank}(M^T) \\ [V_{r+1} \cdots V_n] = \text{Null}(M) \end{cases}$

$M \overset{RREF}{\sim} [0\ 1\ 2] \Rightarrow \text{Rank}(M) = \text{span}\{[1]\}$

$[0\ 1\ 2] \rightarrow x_1$ free, $x_2 + 2x_3 = 0 \quad x_2 = -2x_3$

$x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} x_2 \Rightarrow \text{Nul}(M) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}\right\}$

$\xrightarrow{\text{Normalized}} \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{\sqrt{5}}{5} \\ \frac{-2\sqrt{5}}{5} \end{bmatrix}\right\}$

$M^T \sim \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \end{bmatrix} \Rightarrow \text{Rank}(M^T) = \text{Span}\left\{\begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}\right\}$ $\quad$ Normalize $\rightarrow \text{Span}\left\{\begin{bmatrix} 0 \\ \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} \end{bmatrix}\right\}$

$\because$ there's only 1 column and it has pivot

$\therefore \text{Nul}(M^T) = \{0\}$

In conclusion:

$U = [1]$, $\Sigma = [\sqrt{5}\ 0\ 0]$, $V = \begin{bmatrix} 0 & 1 & 0 \\ \frac{\sqrt{5}}{5} & 0 & \frac{\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & 0 & \frac{-2\sqrt{5}}{5} \end{bmatrix}$

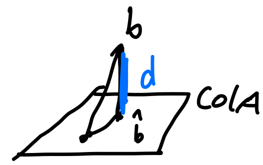
$\therefore$ SVD: $M = [1] \cdot [\sqrt{5}\ 0\ 0] \begin{bmatrix} 0 & \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ 1 & 0 & 0 \\ 0 & \frac{-2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix}$

3

1) $\because \hat{b} = \text{proj}_{\text{Col}A}\, b = A\hat{x}$

$\therefore \|A\hat{x} - b\| = \|\hat{b} - b\| = d \leq \|Ax - b\|$

$\boxed{\text{Thus, } \hat{x} \text{ is a minimizer}}$



$A\hat{x} = b$

$A^TA\hat{x} = A^Tb$

If $\hat{x}$ is unique, A^TA should be invertible.

$\because$ A is full rank

$\therefore Nul(A) = \{0\} = Nul(A^TA)$

- $A \cdot x = 0$ where $x \in Nul(A)$

$\because A^TA \cdot x = 0$

$\therefore x \in Nul(A^TA)$

- $A^TAx' = 0$ where $x' \in Nul(A^TA)$

$x'^T A^T A x' = 0$

$(Ax')^T Ax' = 0$

$Ax' = 0$

$\therefore x' \in Nul(A)$

$\therefore \boxed{Nul(A) = Nul(A^TA)}$

According to invertible matrix theorem, if $Nul(A^TA) = \{0\}$ is true, then equivalently, A^TA is invertible.

$\Rightarrow \hat{x} = (A^TA)^{-1}A^Tb$ is the unique minimizer.

2) $\because$ A is rank deficient, $Nul(A) \neq \{0\}$

$\therefore Nul(A^TA) \neq \{0\}$

according IMT. A^TA is not invertible.

$\because A = U\Sigma V^T$, and in reduced $U\Sigma V^T$, Σ is invertible.

$\therefore$ We have pseudoinverse $A^+ = V\Sigma^{-1}U^T \approx A^T$

$A^TA\hat{x} = A^Tb$

$A^+A\hat{x} = A^+b$

$V\Sigma^{-1}U^TU\Sigma V^T\hat{x} = V\Sigma^{-1}U^Tb$

$\hat{x} = V\Sigma^{-1}U^Tb$ is the optimal solution.