LUMS
A Not-for-Profit University

Department of Electrical Engineering
School of Science and Engineering

EE212 Mathematical Foundations for Machine Learning and Data Science

ASSIGNMENT 2 - Solution

Due Date: 23:55, Monday, October 18, 2021. (Submit online on LMS)
Format: 6 problems, for a total of 100 marks
Instructions:

- You are allowed to collaborate with your peers but copying your colleague's solution is strictly prohibited. This is not a a group assignment. Each student must submit his/her own assignment, scanned in **a single PDF document.**
- Solve the assignment on blank A4 sheets and either scan the document using a scanner or use CamScanner proficiently.
- Upload the solved assignment on LMS in the "Assignments" tab under Assignment 01.
- Naming convention should be as follows: "Name_RollNumber_Assignment_1.pdf"
- Feel free to contact the instructor or the teaching assistants if you have any concerns.

- You represent the most competent individuals in the country, do not let plagiarism come in between your learning. In case any instance of plagiarism is detected, the disciplinary case will be dealt with according to the university's rules and regulations.

Problem 1 (10 marks)

Solve the following systems of linear equations. If a solution does not exist, specify.

(a) [**5 marks**] System A:
$2r + 2s + 9t = 7$
$r - 3t = 8$
$s + 5t = -2$

(b) [**5 marks**] System B:
$3j - 7k + 7l = -8$
$-4j + 6k - l = 7$
$j - 3k + 4l = -4$

Solution.

(a) System A:

$$\begin{bmatrix}1 & 0 & -3 & 8\\ 2 & 2 & 9 & 7\\ 0 & 1 & 5 & -2\end{bmatrix} \to \begin{bmatrix}1 & 0 & -3 & 8\\ 2 & 2 & 9 & 7\\ 0 & 1 & 5 & -2\end{bmatrix} \to \begin{bmatrix}1 & 0 & -3 & 8\\ 0 & 2 & 15 & -9\\ 0 & 1 & 5 & -2\end{bmatrix} \to$$
$$\begin{bmatrix}1 & 0 & -3 & 8\\ 0 & 1 & 5 & -2\\ 0 & 0 & 5 & -5\end{bmatrix} \to \begin{bmatrix}1 & 0 & -3 & 8\\ 0 & 1 & 5 & -2\\ 0 & 0 & 1 & -1\end{bmatrix} \to \begin{bmatrix}1 & 0 & 0 & 5\\ 0 & 1 & 0 & 3\\ 0 & 0 & 1 & -1\end{bmatrix}.$$

The solution is (5,3,-1).

(b) System B:
Replace $R2$ by $R2 + (-3)R1$.
Replace $R3$ by $R3 + 4(R1)$.
Replace $R3$ by $R3 + (3)R2$.

$$\begin{bmatrix}1 & -3 & 4 & -4\\ 3 & -7 & 7 & -8\\ -4 & 6 & -1 & 7\end{bmatrix} \to \begin{bmatrix}1 & -3 & 4 & -4\\ 0 & 2 & -5 & 4\\ 0 & -6 & 15 & -9\end{bmatrix} \to \begin{bmatrix}1 & -3 & 4 & -4\\ 0 & 2 & -5 & 4\\ 0 & 0 & 0 & 3\end{bmatrix}$$

The last row requires that 0=3. Thus, the system is inconsistent and its solution set is empty.

Problem 2 (10 marks)

Solve the following system of linear equations.

$$x + 2y + 3z = 1$$
$$x + 2y + 4z = 2$$

Solution.

$$AX = b$$

Where:

$$A = \begin{bmatrix}1 & 2 & 3\\ 1 & 2 & 4\end{bmatrix}, \quad X = \begin{bmatrix}x\\ y\\ z\end{bmatrix} \quad and \quad b = \begin{bmatrix}1\\ 2\end{bmatrix}$$

A is a fat matrix (has more columns than rows) hence it has a right pseudo inverse, we will write it as $A^\dagger$ where:

$$A^\dagger = A^T(AA^T)^{-1}$$

$$A^T = \begin{bmatrix}1 & 1\\ 2 & 2\\ 3 & 4\end{bmatrix}$$

$$AA^T = \begin{bmatrix}1 & 2 & 3\\ 1 & 2 & 4\end{bmatrix}\begin{bmatrix}1 & 1\\ 2 & 2\\ 3 & 4\end{bmatrix} = \begin{bmatrix}14 & 17\\ 17 & 21\end{bmatrix}$$

$$(AA^T)^{-1} = \begin{bmatrix} 4.2 & -3.4 \\ -3.4 & 2.8 \end{bmatrix}$$

$$A^T(AA^T)^{-1} = \begin{bmatrix} 0.8 & -0.6 \\ 1.6 & -1.2 \\ -1 & 1 \end{bmatrix} = A^\dagger$$

$$X = A^\dagger b$$

$$X = \begin{bmatrix} 0.8 & -0.6 \\ 1.6 & -1.2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$X = \begin{bmatrix} -0.4 \\ -0.8 \\ 1 \end{bmatrix}$$

Therefore:

$$x = -0.4, \quad y = -0.8 \quad and \quad z = 1$$

Problem 3 (10 marks)

In the following matrices, the black squares signify any non-zero value and the asterisks signify any value including zero. For the following augmented matrices, state if each is consistent (that there is at least one set of values for each of the unknowns that satisfies the system) and whether its solution is unique.

(a) [**5 marks**] Matrix A:

$$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & 0 \end{bmatrix}$$

(b) [**5 marks**] Matrix B:

$$\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$$

Solution.

(a) The system is consistent, with a unique solution.

(b) The system is consistent. There are many solutions because $C2$ is a free variable. A variable is a basic variable if it corresponds to a pivot column. Otherwise, the variable is known as a free variable. In order to determine which variables are basic and which are free, it is necessary to row reduce the augmented matrix to echelon form.

Problem 4 (20 marks)

You are given a matrix M of size $m \times n$ which is observably skinny *(has more rows than* *columns)* and is full-rank *(has n linearly independent columns)*. x_{ls} is the least square solution of $Mx = y$ and $y_{ls} = Mx_{ls}$.

Given that the relationship $r = y - y_{ls}$ describes the residual vector of this problem, show that it satisfies the following expression:

$$||r||^2 = ||y||^2 - ||y_{ls}||^2.$$

Solution. Let us first show that $r \perp y_{ls}$. Since $y_{ls} = Ax_{ls} = AA^{\dagger}y = A(A^TA)^{-1}A^Ty$

$$\begin{aligned} {y_{ls}}^T r &= y_{ls}^T(y - y_{ls}) = y_{ls}^T y - y_{ls}^T y_{ls} \\ &= y^TA(A^TA)^{-1}A^Ty - y^TA(A^TA)^{-1}A^TA(A^TA)^{-1}A^Ty \\ &= y^TA(A^TA)^{-1}A^Ty - y^TA(A^TA)^{-1}(A^TA)(A^TA)^{-1}A^Ty \\ &= y^TA(A^TA)^{-1}A^Ty - y^TA(A^TA)^{-1}A^Ty \\ &= 0. \end{aligned}$$

Thus,

$$\begin{aligned} ||y||^2 &= ||y_{ls} + r||^2 \\ &= (y_{ls} + r)^T(y_{ls} + r) \\ &= ||y_{ls}||^2 + 2y_{ls}^T r + ||r||^2 \\ &= ||y_{ls}||^2 + ||r||^2. \end{aligned}$$

Therefore,

$$||r||^2 = ||y||^2 - ||y_{ls}||^2.$$

Problem 5 (30 marks)

You are taking part in a Machine Learning-themed scavenger hunt. After solving a handful of clues and navigating the campus grounds to the locations they reveal, you now find yourself with a piece of paper that reads the following:

You are given an $m \times n$ matrix M along with five facts about it.
1. M^TM has two non-zero eigenvalues: 9 and 4.
2. One of the two normalized eigenvectors for M^TM is $e_1 = \frac{1}{\sqrt{5}}\begin{bmatrix}1 & 2\end{bmatrix}^T$.
3. M has SVD= $U\Sigma V^T$ where the third column of U is $u_3 = \begin{bmatrix}2/3 & -1/3 & -2/3\end{bmatrix}^T$.
4. $Mv_1 = \frac{1}{\sqrt{5}}\begin{bmatrix}5 & 2 & 4\end{bmatrix}^T$ where v_1 is the first column of the matrix V.
5. $Mv_2 = \frac{2}{\sqrt{5}}\begin{bmatrix}0 & 2 & -1\end{bmatrix}^T$ where v_2 is the second column of the matrix V.
The square of the Σ_{22} entry will determine the floor of SSE where you will find your next clue.

To solve the clue, you take a five-step approach:

(a) **[5 marks]** Determine the size of the matrix M.

(b) **[7 marks]** Write out the U matrix.

(c) [**7 marks**] Write out the V matrix.

(d) [**7 marks**] Write out the Σ matrix.

(e) [**4 marks**] Which floor of SSE must you head to next?

Solution.

(a) M is a mxn matrix so M_{mxn}.
Dimensions of u_3 should be mx1, it is 3x1, hence $m = 3$.
For the matrix product Mv_1 to have dimensions 3x1, the dimensions should be $M_{3xn}(v_1)_{nx1}$. Hence $n = 2$.

(b)

$$U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}$$

$$u_1 = \frac{1}{\sqrt{\lambda_1}} A v_1; \quad \lambda_1 = 9 \quad and \quad Av_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 2 & 4 \end{bmatrix}^T \quad \text{(given)}$$

$$u_2 = \frac{1}{\sqrt{\lambda_2}} A v_2; \quad \lambda_2 = 4 \quad and \quad Av_2 = \frac{2}{\sqrt{5}} \begin{bmatrix} 0 & 2 & -1 \end{bmatrix}^T \quad \text{(given)}$$

$$u_3 = \begin{bmatrix} \frac{2}{3} & \frac{-1}{3} & \frac{-2}{3} \end{bmatrix}^T \text{(given)}$$

Hence:

$$U = \begin{bmatrix} \frac{5}{3\sqrt{5}} & 0 & \frac{2}{3} \\ \frac{2}{3\sqrt{5}} & \frac{2}{\sqrt{5}} & -\frac{1}{3} \\ \frac{4}{3\sqrt{5}} & -\frac{1}{\sqrt{5}} & -\frac{2}{3} \end{bmatrix}$$

(c)

$$V = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$$

$$v_1 = e_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

We know that $v_1^T v_2 = 0$ as they are both orthogonal. So

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} v_{2a} \\ v_{2b} \end{bmatrix} = 0$$

$$v_{2a} = -2v_{2b}$$

Both vectors should be normalized so

$$v_{2a}^2 + v_{2b}^2 = 1$$

Solve these two equations simultaneously to get:

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

(d)

$$\Sigma = \begin{bmatrix} \sqrt{9} & 0 \\ 0 & \sqrt{4} \\ 0 & 0 \end{bmatrix}$$

(e) 4^{th} floor

Problem 6 (20 marks)

After a tiring day of sorting through signal data, you decide to wrap up by organizing it into a clean matrix. Upon formulating the data matrix D, you observe the following relationship:

$$rank(D) = rank(D^T D) = rank(DD^T)$$

Initially you're puzzled by this relationship, but then you recall the concept of SVD and decide to use it to explain why this may be.
Using the SVD of an $m \times n$ matrix D, prove that the observed relationship holds.

Solution.

$$D_{mxn} = (U_1)_{mxr}(\Sigma_1)_{rxr}(V_1^T)_{rxn} \quad (1)$$

$$D_T D = (U_1\Sigma_1 V_1^T)^T(U_1\Sigma_1 V_1^T)$$

$$= V_1\Sigma_1^T U_1^T U_1\Sigma_1 V_1^T$$

$$= V_1\Sigma_1^T\Sigma_1 V_1^T; \quad \text{Remember } U_1^T U_1 = I$$

$$D^T D = V_1(\Sigma_2)_{rxr}V_1^T \quad (2)$$

Where Σ_1, Σ_1^T are diagonal matrices.

$$DD^T = (U_1\Sigma_1 V_1^T)(U_1\Sigma_1 V_1^T)^T$$

$$= U_1\Sigma_1 V_1^T V_1\Sigma_1^T U_1^T$$

$$= U_1\Sigma_1\Sigma_1^T U_1^T; \quad \text{Remember } V_1^T V_1 = I$$

$$DD^T = U_1(\Sigma_3)_{rxr}U_1^T \quad (3)$$

Using Eq. (1), (2) and (3) we can say that:

rank(D)=rank(Σ_1)=r
rank($D^T D$)=rank(Σ_2)=r
rank(DD^T)=rank(Σ_3)=r

Therefore:

$$rank(D) = rank(D^T D) = rank(DD^T)$$

— End of Assignment —