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

Due Date: 23:55, Saturday. July 25, 2020 (Submit online on LMS)
Format: 12 problems, for a total of 100 marks
Instructions:

- You are not allowed to submit a group assignment. Each student must submit his/her own hand-written assignment, scanned in a single PDF document.
- You are allowed to collaborate with your peers but copying your colleague's solution is strictly prohibited. Anybody found guilty would be subjected to disciplinary action in accordance with the university rules and regulations.

Problem 1 (10 marks)

Potential customers are divided into m market segments, which are groups of customers with similar demographics, e.g., college educated women aged 25–29. A company markets its products by purchasing advertising in a set of n channels, i.e., specific TV or radio shows, magazines, web sites, blogs, direct mail, and so on. The ability of each channel to deliver impressions or views by potential customers is characterized by the $reachmatrix$, the $m \times n$ matrix R, where R_{ij} is the number of views of customers in segment i for each dollar spent on channel j. (We assume that the total number of views in each market segment is the sum of the views from each channel, and that the views from each channel scale linearly with spending.) The n-vector c will denote the company's purchases of advertising, in dollars, in the n channels. The m-vector v gives the total number of impressions in the m market segments due to the advertising in all channels. Finally, we introduce the m-vector a, where a_i gives the profit in dollars per impression in market segment i. The entries of R, c, v, and a are all nonnegative.

(a) [**2 marks**] Express the total amount of money the company spends on advertising using vector/matrix notation.

(b) [**2 marks**] Express v using vector/matrix notation, in terms of the other vectors and matrices.

(c) [**2 marks**] Express the total profit from all market segments using vector/matrix notation.

(d) [**2 marks**] How would you find the single channel most effective at reaching market segment 3, in terms of impressions per dollar spent?

(e) [**2 marks**] What does it mean if R_{35} is very small (compared to other entries of R)?

Problem 2 (15 marks)

An $n \times n$ matrix A is called skew-symmetric if $A^T = -A$, i.e., its transpose is its negative. (A symmetric matrix satisfies $A^T = A$.)

(a) [**3 marks**] Find all 2×2 skew-symmetric matrices.

(b) [**3 marks**] Explain why the diagonal entries of a skew-symmetric matrix must be zero.

(c) [**5 marks**] Show that for a skew-symmetric matrix A, and any n-vector x, $(Ax) \perp x$. This means that Ax and x are orthogonal.*Hint.* First show that for any n x n matrix A and n-vector x, $x^T(Ax) = \sum_{i,j=1}^{n} A_{ij}x_ix_j$.

(d) [**4 marks**] Now suppose A is any matrix for which $(Ax) \perp x$ for any n-vector x. Show that A must be skew-symmetric. *Hint.* You might find the formula

$$(e_i + e_j)^T(A(e_i + e_j)) = A_{ii} + A_{jj} + A_{ij} + A_{ji},$$

valid for any $n \times n$ matrix A, useful. For $i = j$, this reduces to $e_i^T(Ae_i) = A_{ii}$.

Problem 3 (5 marks)

Suppose A is an $m \times n$ matrix and x is an n-vector. A famous inequality relates $\|x\|$, $\|A\|$, and $\|Ax\|$:

$$\|Ax\| \leq \|A\|\|x\|$$

The left-hand side is the (vector) norm of the matrix-vector product; the right-hand side is the (scalar) product of the matrix and vector norms. Show this inequality. *Hints.* Let a_i^T be the ith row of A. Use the Cauchy–Schwarz inequality to get $(a_i^T x)^2 \leq \|a_i\|^2\|x\|^2$. Then add the resulting m inequalities.

Problem 4 (5 marks)

We consider a set of n basic foods (such as rice, beans, apples) and a set of m nutrients or components (such as protein, fat, sugar, vitamin C). Food j has a cost given by c_j (say, in dollars per gram), and contains an amount N_{ij} of nutrient i (per gram). (The nutrients are given in some appropriate units, which can depend on the particular nutrient.) A daily diet is represented by an n-vector d, with d_i the daily intake (in grams) of food i. Express the condition that a diet d contains the total nutrient amounts given by the m-vector n^{des}, and has a total cost B (the budget) as a set of linear equations in the variables $d_1, \ldots, d_n$. (The entries of d must be non-negative, but we ignore this issue here.)

Problem 5 (4 marks)

Suppose that A is an $m \times n$ matrix, D is a diagonal matrix, and $B = DA$. Describe B in terms of A and the entries of D. You can refer to the rows or columns or entries of A.

Problem 6 (5 marks)

The sum of the diagonal entries of a square matrix is called the trace of the matrix, denoted $\mathbf{tr}(A)$.Suppose A and B are $m \times n$ matrices. Show that

$$\mathbf{tr}(A^T B) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}B_{ij}$$

Problem 7 (6 marks)

We consider m students, n classes, and p majors. Each student can be in any number of the classes (although we'd expect the number to range from 3 to 6), and can have any number of the majors (although the common values would be 0, 1, or 2). The data about the students' classes and majors are given by an m x n matrix C and an $m \times p$ matrix M, where

$$C_{ij} = \begin{cases} 1 & \text{student } i \text{ is in class } j \\ 0 & \text{student } i \text{ is not in class } j, \end{cases}$$

and

$$M_{ij} = \begin{cases} 1 & \text{student } i \text{ is in major } j \\ 0 & \text{student } i \text{ is not in major } j, \end{cases}$$

(a) [**3 marks**] Let E be the n-vector with E_i being the enrollment in class i. Express E using matrix notation, in terms of the matrices C and M.

(b) [**3 marks**] Define the $n \times p$ matrix S where S_{ij} is the total number of students in class i with major j. Express S using matrix notation, in terms of the matrices C and M.

Problem 8 (10 marks)

Let $G \in \mathbf{R}^{m \times n}$ represent a contingency matrix of m students who are members of n groups:

$$G_{ij} = \begin{cases} 1 & \text{student } i \text{ is in group j} \\ 0 & \text{student } i \text{ is not in group j} \end{cases}$$

(A student can be in any number of the groups.)

(a) [**2 marks**] What is the meaning of the 3rd column of G?

(b) [**2 marks**] What is the meaning of the 15th row of G?

(c) [**2 marks**] Give a simple formula (using matrices, vectors, etc.) for the n-vector M, where M_i is the total membership (i.e., number of students) in group i.

(d) [**2 marks**] Interpret $(GG^T)_{ij}$ in simple English.

(e) [**2 marks**] Interpret $(G^TG)_{ij}$ in simple English.

Problem 9 (5 marks)

Let Z be a tall $m \times n$ matrix with linearly independent columns, and let X and Y be left inverses of Z. Show that for any scalars α and β satisfying $\alpha + \beta = 1$, $\alpha X + \beta Y$ is also a left inverse of Z. It follows that if a matrix has two different left inverses, it has an infinite number of different left inverses.

Problem 10 (10 marks)

Find the inverse of the $n \times n$ running sum matrix,

$$S = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix}$$

Does your answer make sense?

Problem 11 (5 marks)

Suppose A is an $n \times p$ matrix and B is a $p \times n$ matrix, so $C = AB$ makes sense. Explain why C cannot be invertible if A is tall and B is wide, i.e., if $p < n$. *Hint.* First argue that the columns of B must be linearly dependent.

Problem 12 (20 marks)

Let a be $n \times n$ symmetric matrix which can be written as $A = S\Lambda S^T$ using EVD or as $A = U\Sigma V^T$ using SVD.

(a) [**4 marks**] Assuming all eigenvalues of A are positive i.e. $\lambda_i \geq 0$, show that $S = V = U$.

(b) [**16 marks**] For an $m \times n$ matrix A, following information is known:

- $A^T A$ has 9 and 4 as its two non-zero eigenvalues and one of the two normalised eigenvectors is $\underline{e}_i = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \end{bmatrix}^T$
- It has SVD as $U\Sigma V^T$ with matrix U having its third column equal to $\underline{u}_3 = \begin{bmatrix} 2/3 & -1/3 & -2/3 \end{bmatrix}^T$
- It is also given $A\underline{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 2 & 4 \end{bmatrix}^T$ and $A\underline{v}_2 = \frac{2}{\sqrt{5}} \begin{bmatrix} 0 & 2 & -1 \end{bmatrix}^T$ where $\underline{v}_1$ and $\underline{v}_2$ are first two columns of matrix V

i. What is the size of A?

ii. Write down U matrix.

iii. Write down V matrix.

iv. Write down Σ matrix.

— End of Assignment —