Theoretical Machine Learning - Assignment 1

Gaurav Kumar Rampuria

May 16, 2024

1. Problem 1

1.1.

Consider a general matrix $\mathtt{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & ... & a_{1n} \\ a_{21} & a_{22} & a_{23} & ... & a_{2n} \\ a_{31} & a_{32} & a_{33} & ... & a_{3n} \\ . & . & . & ... & . \\ . & . & . & ... & . \\ a_{m1} & a_{m2} & a_{m3} & ... & a_{mn} \end{vmatrix}$ of order m x n, and $\mathtt{x} = \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ x_n \end{vmatrix}$ of order n x 1.

Their matrix multiplication gives a vector $= \begin{vmatrix} a_{11}x_1 + a_{12}x_2 + ...a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + ...a_{2n}x_n \\ a_{31}x_1 + a_{32}x_2 + ...a_{3n}x_n \\ . \\ . \\ a_{m1}x_1 + a_{m2}x_2 + ...a_{mn}x_n \end{vmatrix}$

Now we use differentiation of one vector, A$\mathtt{x}$ by another vector, $\mathtt{x}$. This will give us a matrix of dimension m x n.

$$\begin{vmatrix} \frac{\partial a_{11}x_1}{\partial x_1} & \frac{\partial a_{12}x_2}{\partial x_2} & \frac{\partial a_{13}x_3}{\partial x_3} & ... & \frac{\partial a_{1n}x_n}{\partial x_n} \\ \frac{\partial a_{21}x_1}{\partial x_1} & \frac{\partial a_{22}x_2}{\partial x_2} & \frac{\partial a_{23}x_3}{\partial x_3} & ... & \frac{\partial a_{2n}x_n}{\partial x_n} \\ \frac{\partial a_{31}x_1}{\partial x_1} & \frac{\partial a_{32}x_2}{\partial x_2} & \frac{\partial a_{33}x_3}{\partial x_3} & ... & \frac{\partial a_{3n}x_n}{\partial x_n} \\ . & . & . & ... & . \\ . & . & . & ... & . \\ \frac{\partial a_{m1}x_1}{\partial x_1} & \frac{\partial a_{m2}x_2}{\partial x_2} & \frac{\partial a_{m3}x_3}{\partial x_3} & ... & \frac{\partial a_{mn}x_n}{\partial x_n} \end{vmatrix}$$

which gives us matrix A itself. *Hence, Proven.*

1.2.

Consider the nxn square matrix `A` $= \begin{vmatrix} a_{11} & a_{12} & a_{13} & ... & a_{1n} \\ a_{21} & a_{22} & a_{23} & ... & a_{2n} \\ a_{31} & a_{32} & a_{33} & ... & a_{3n} \\ . & . & . & ... & . \\ . & . & . & ... & . \\ a_{n1} & a_{n2} & a_{n3} & ... & a_{nn} \end{vmatrix}$

and `x` $= \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ x_n \end{vmatrix}$ of order n x 1. Hence, $x^T = \begin{vmatrix} x_1 & x_2 & x_3 & ... & x_n \end{vmatrix}$ of order 1 x n.

$x^T A = \begin{vmatrix} \sum_{i=1}^{N}(a_{i1}x_i) & \sum_{i=1}^{N}(a_{i2}x_i) & ... & \sum_{i=1}^{N}(a_{in}x_i) \end{vmatrix}$ of order 1 x n.

Hence, $x^T A x = \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{ij}x_i x_j)$, which is a scalar. Now we differentiate a scalar wrt a vector.

$$\begin{vmatrix} \frac{\partial \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{ij}x_i x_j)}{\partial x_1} & \frac{\partial \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{ij}x_i x_j)}{\partial x_2} & \frac{\partial \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{ij}x_i x_j)}{\partial x_3} & ... & \frac{\partial \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{ij}x_i x_j)}{\partial x_n} \end{vmatrix}$$

$$\begin{vmatrix} \sum_{i=1}^{N}(a_{i1}x_i) + \sum_{j=1}^{N}(a_{1j}x_j) & \sum_{i=1}^{N}(a_{i2}x_i) + \sum_{j=1}^{N}(a_{2j}x_j) & \sum_{i=1}^{N}(a_{i3}x_i) + \sum_{j=1}^{N}(a_{3j}x_j) & ... & \sum_{i=1}^{N}(a_{in}x_i) + \sum_{j=1}^{N}(a_{nj}x_j) \end{vmatrix}$$

This can be split into addition of 2 matrices as:

$$\begin{vmatrix} \sum_{i=1}^{N}(a_{i1}x_i) & \sum_{i=1}^{N}(a_{i2}x_i) & ... & \sum_{i=1}^{N}(a_{in}x_i) \end{vmatrix} + \begin{vmatrix} \sum_{j=1}^{N}(a_{1j}x_j) & \sum_{j=1}^{N}(a_{2j}x_j) & ... & \sum_{j=1}^{N}(a_{nj}x_j) \end{vmatrix}$$

This basically is, $x^T A + x^T A^T = x^T(A + A^T)$. *Hence, Proven.*

2. Problem 2

On differentiating a `mx1` vector by a `kx1` vector we get an `mxk` matrix. Now consider n such `mx1` vectors stacked to form a `m x n` matrix, then the final dimension of the differentiation will be **m x nk**.

3. Problem 3

3.1. Solve

$$\frac{\mathrm{d}\begin{vmatrix} 2sin^2(x)cos(y) \\ x^2+3e^y \end{vmatrix}}{\mathrm{d}\begin{vmatrix} x \\ y \end{vmatrix}}$$

$$= \begin{vmatrix} \frac{\partial(2sin^2(x)cos(y))}{\partial x} & \frac{\partial(2sin^2(x)cos(y))}{\partial y} \\ \frac{\partial(x^2+3e^y)}{\partial x} & \frac{\partial(x^2+3e^y)}{\partial y} \end{vmatrix} = \begin{vmatrix} 2sin(2x)cos(y) & -2sin^2(x)sin(y) \\ 2x & 3e^y \end{vmatrix}$$

3.2. Solve

$$\frac{\mathrm{d}\begin{vmatrix} 3x^2y+xyzw \\ sin(x^2+yw-z \end{vmatrix}}{\mathrm{d}\begin{vmatrix} x \\ y \\ w \\ z \end{vmatrix}}$$

$$= \begin{vmatrix} \frac{\partial(3x^2y+xyzw)}{\partial x} & \frac{\partial(3x^2y+xyzw)}{\partial y} & \frac{\partial(3x^2y+xyzw)}{\partial z} & \frac{\partial(3x^2y+xyzw)}{\partial w} \\ \frac{\partial(sin(x^2+yw-z)}{\partial x} & \frac{\partial(sin(x^2+yw-z)}{\partial y} & \frac{\partial(sin(x^2+yw-z)}{\partial z} & \frac{\partial(sin(x^2+yw-z)}{\partial w} \end{vmatrix}$$

$$= \begin{vmatrix} 6xy+yzw & 3x^2+xzw & xyw & xyz \\ 2xcos(x^2+yw-z) & wcos(x^2+yw-z) & -cos(x^2+yw-z) & ycos(x^2+yw-z) \end{vmatrix}$$

4. Problem 4

Consider $\beta = \begin{vmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ . \\ . \\ \beta_n \end{vmatrix}$ and $\mathtt{x} = \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ x_n \end{vmatrix}$, both vectors of order n x 1.

Then $\beta^T x = \sum_{i=1}^{N}(\beta_i x_i)$. So, $e^{\sum_{i=1}^{N}(\beta_i x_i)}$ is a scalar, which is to be differentiated by the vector $\mathtt{x}$. This gives us the transpose a vector.

$$\left| \begin{array}{ccccc} \frac{\partial e^{\sum_{i=1}^{N}(\beta_i x_i)}}{\partial x_1} & \frac{\partial e^{\sum_{i=1}^{N}(\beta_i x_i)}}{\partial x_2} & \frac{\partial e^{\sum_{i=1}^{N}(\beta_i x_i)}}{\partial x_3} & ... & \frac{\partial e^{\sum_{i=1}^{N}(\beta_i x_i)}}{\partial x_n} \end{array} \right|$$

$$= \left| \begin{array}{ccccc} \beta_1 e^{\sum_{i=1}^{N}(\beta_i x_i)} & \beta_2 e^{\sum_{i=1}^{N}(\beta_i x_i)} & \beta_3 e^{\sum_{i=1}^{N}(\beta_i x_i)} & ... & \beta_n e^{\sum_{i=1}^{N}(\beta_i x_i)} \end{array} \right|$$

$$= \left| \begin{array}{ccccc} \beta_1 e^{\beta^T x} & \beta_2 e^{\beta^T x} & \beta_3 e^{\beta^T x} & ... & \beta_n e^{\beta^T x} \end{array} \right| = e^{\beta^T x} \left| \begin{array}{ccccc} \beta_1 & \beta_2 & \beta_3 & ... & \beta_n \end{array} \right| = \beta^T e^{\beta^T x}$$

Ans) $\beta^T e^{\beta^T x}$