

COMP3223: Solutions to Calculus Exercises

October 23, 2020

I Partial derivatives and matrix calculus

1. Using the symbol δ_{ab} , the Kroenecker delta

$$\delta_{ab} = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases},$$

show the following

- (a) for vector \mathbf{v} with components v_i , $\sum_i v_i \delta_{ij} = v_j$;

(

ik;

<https://eduassistpro.github.io/>

$\mathbf{C} = \mathbf{A}\mathbf{B}$ is $\mathbf{C}_{ik} = \sum_j \mathbf{A}_{ij} \mathbf{B}_{jk}$

- (d) the trace of a matrix is $\text{tr}(\mathbf{A}) =$

(e) $\frac{\partial w}{\partial v_b} = \delta_{ab}$

Add WeChat edu_assist_pro

2. For $p \times p$ matrix \mathbf{A} with matrix elements a_{ij} and vector $\mathbf{x} = (x_1, \dots, x_p)^T$ show that:

- (a) the i -th element of vector $(\mathbf{A}\mathbf{x})$ is $(\mathbf{A}\mathbf{x})_i = \sum_j a_{ij} x_j$;

- (b) $\nabla_{\mathbf{x}}(\mathbf{A}\mathbf{x}) := \frac{\partial}{\partial \mathbf{x}}(\mathbf{A}\mathbf{x}) = \mathbf{A}^T$. Write out the indices explicitly:

$$\left(\frac{\partial}{\partial \mathbf{x}}(\mathbf{A}\mathbf{x}) \right)_{ij} = \frac{\partial}{\partial x_i}(\mathbf{A}\mathbf{x})_j = \frac{\partial}{\partial x_i} \sum_k a_{jk} x_k;$$

- (c) the gradient of the scalar quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is

$$\nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = (\mathbf{A} + \mathbf{A}^T) \mathbf{x};$$

hint: the i -th matrix element of the gradient is

$$\frac{\partial}{\partial x_i} \sum_{p,q} x_p a_{p,q} x_q;$$

- (d) the partial derivative of the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ with respect to \mathbf{A} can be evaluated for each matrix element a_{ij} , $1 \leq i, j \leq p$:

$$\frac{\partial}{\partial a_{ij}} \left(\sum_{rs} x_r a_{rs} x_s \right)$$

with the result

$$\nabla_{\mathbf{A}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = \mathbf{x} \mathbf{x}^T;$$

with $\mathbf{x} \mathbf{x}^T$ a $p \times p$ matrix.

2 Solutions

1. (a) The column vector \mathbf{v} has components v_i

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_i \\ \vdots \end{pmatrix}, \quad (1)$$

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu_assist_pro

Therefore,

$$\begin{aligned} \sum_{i=1}^n v_i \delta_{ij} &= v_1 \delta_{1j} + v_2 \delta_{2j} + v_3 \delta_{3j} + \cdots + v_i \delta_{ij} + \cdots + v_n \delta_{nj} = \\ &= (\text{has only non-zero term with } \delta_{ij} = 1 \text{ when } i = j) = v_j \end{aligned} \quad (2)$$

- (b) For $m \times n$ matrix \mathbf{A} with components a_{ij} ,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}, \quad (4)$$

$$\sum_{j=1}^n a_{ij} \delta_{jk} = a_{i1} \delta_{1k} + a_{i2} \delta_{2k} + a_{i3} \delta_{3k} + \cdots + a_{in} \delta_{nk} =$$

$$= (\text{only non-zero term } \delta_{jk} = 1 \text{ occurs when } j = k) = a_{ik} \quad (5)$$

This can also be seen by viewing the Kronecker delta as the matrix elements of the identity matrix $(\mathbb{I})_{jk} = \delta_{jk}$:

$$a_{ik} = (\mathbf{A})_{ik} = (\mathbf{A}\mathbb{I})_{ik} = \sum_{j=1}^n (\mathbf{A})_{ij} (\mathbb{I})_{jk} = \sum_{j=1}^n a_{ij} \delta_{jk}$$

(c) Using Equation (5) we obtain

$$\sum_{jk} a_{ij} b_{jk} \delta_{ik} = \sum_j a_{ij} \sum_k b_{jk} \delta_{ik} = \sum_j a_{ij} b_{ji} = a_{i1} b_{1i} + a_{i2} b_{2i} + \cdots + a_{in} b_{ni}, \quad (6)$$

which is the i^{th} diagonal $(\mathbf{C})_{ii} = (\mathbf{C})_{ij} \delta_{ij}$ of $\mathbf{C} = \mathbf{AB}$. For example

3×3 matrices $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$, the matrix \mathbf{C}

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix},$$

and you should notice that the i th diagonal element is

$$C_{ii} = a_{i1}b_{1i} + a_{i2}b_{2i} + a_{i3}b_{3i}.$$

Thus, the i^{th} diagonal element (for $i = 1, 2, 3$) of \mathbf{C} can be obtained from Equation (5) by setting $n = 3$.

(d) Following Equation (5) for matrix in form (4)

$$\sum_{ij} a_{ij} \delta_{ij} = \sum_i a_{ii} = a_{11} + a_{22} + a_{33} + \cdots + a_{ii} + \cdots + a_{nn} = \text{tr}(\mathbf{A}). \quad (7)$$

(e) For a function $f(x_1, x_2, \cdots, x_n)$ with n independent variables x_i $i = 1, \dots, n$, the partial derivative $\frac{\partial}{\partial x_i} f(x_1, \cdots, x_i, \cdots, x_n)$ is defined as

$$\lim_{h_i \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h_i, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)}{h_i} \quad (8)$$

making the derivatives of $f(w_1, \dots, w_n) = w_a$ with respect to w_a and w_b

$$\lim_{h_a \rightarrow 0} \frac{(w_a + h_a) - w_a}{h_a} = 1, \quad \lim_{h_b \rightarrow 0} \frac{w_a - w_a}{h_b} = 0 \implies \frac{\partial w_a}{\partial w_b} = \delta_{ab} \quad (9)$$

2. (a) For matrix \mathbf{A} shown explicitly in Equation (4) in the case of $p \times p$ and input vector \mathbf{x} ,

$$\mathbf{x} = (x_1, \dots, x_p)^T = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \quad (10)$$

$$\mathbf{Ax} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1p}x_p \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pp}x_p \end{pmatrix}, \quad (11)$$

where the i^{th} element of the output vector $\mathbf{y} = \mathbf{Ax}$ is y_i

$$(\mathbf{y})_i = y_i = (\mathbf{Ax})_i = a_{i1}x_1 + \cdots + a_{ip}x_p = \sum_{j=1}^p a_{ij}x_j. \quad (12)$$

Assignment Project Exam Help

- (b) The derivative of the output y_i with respect to x_j measures how rapidly yourself that the

<https://eduassistpro.github.io/>

Add WeChat edu_assist_pro

$$\left(\frac{\partial}{\partial x_j} (\mathbf{Ax}) \right)_i = \frac{\partial}{\partial x_j} (\mathbf{Ax})_i = \frac{\partial}{\partial x_j} \left(\sum_{k=1}^p a_{ik}x_k \right) = \sum_{k=1}^p a_{ik} \delta_{kj},$$

and using (5) we get

$$\sum_{k=1}^p a_{ik} \delta_{kj} = a_{ji} = (\mathbf{A})_{ij}.$$

For a 3×3 example,

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} (\mathbf{Ax})_1 \\ (\mathbf{Ax})_2 \\ (\mathbf{Ax})_3 \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + x_2 a_{12} + x_3 a_{13} \\ x_1 a_{21} + x_2 a_{22} + x_3 a_{23} \\ x_1 a_{31} + x_2 a_{32} + x_3 a_{33} \end{pmatrix} \quad (13)$$

and so, (keeping track of row and column indices),

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{y}) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} \end{pmatrix} \quad (14)$$

and writing out the terms explicitly we get

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{Ax}) = \begin{pmatrix} \frac{\partial}{\partial x_1} (\mathbf{Ax})_1 & \frac{\partial}{\partial x_2} (\mathbf{Ax})_1 & \frac{\partial}{\partial x_3} (\mathbf{Ax})_1 \\ \frac{\partial}{\partial x_1} (\mathbf{Ax})_2 & \frac{\partial}{\partial x_2} (\mathbf{Ax})_2 & \frac{\partial}{\partial x_3} (\mathbf{Ax})_2 \\ \frac{\partial}{\partial x_1} (\mathbf{Ax})_3 & \frac{\partial}{\partial x_2} (\mathbf{Ax})_3 & \frac{\partial}{\partial x_3} (\mathbf{Ax})_3 \end{pmatrix}.$$

Using the explicit forms in eq. (13) you can verify that

$$\frac{\partial}{\partial \mathbf{x}} (\mathbf{Ax}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \mathbf{A} \quad (15)$$

- (c) The expression $Q \triangleq \sum_{p,q} x_p a_{pq} x_q$ is a number. Its derivative with respect to \mathbf{x} means that there will be a term for each component of the vector \mathbf{x} . The answer should be a vector as well. For a (3×3) case,

$$Q = (x_1 \ x_2 \ x_3) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

which we write out as

$$Q = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{21}x_2x_1 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{31}x_3x_1 + a_{32}x_3x_2 + a_{33}x_3^2$$

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat: edu_assist_pro

For the general case,

$$\begin{aligned} \frac{\partial}{\partial x_i} \sum_{p,q} x_p a_{pq} x_q &= \sum_{p,q} a_{pq} \left(\frac{\partial x_p}{\partial x_i} x_q + x_p \frac{\partial x_q}{\partial x_i} \right) \\ &= \sum_{p,q} a_{pq} (\delta_{pi} x_q + x_p \delta_{qi}) \\ &= \sum_q a_{iq} x_q + \sum_p x_p a_{pi} \\ &= \sum_p a_{ip} x_p + \sum_p a_{pi} x_p \\ &= \sum_p ((\mathbf{A})_{ip} + (\mathbf{A}^T)_{ip}) x_p \\ &= ((\mathbf{A} + \mathbf{A}^T)\mathbf{x})_i. \end{aligned}$$

- (d) The expression $Q \triangleq \sum_{p,q} x_p a_{pq} x_q$ is a number. Its derivative with respect to \mathbf{A} means that there will be a term for each component of the matrix \mathbf{A} . The answer should be a matrix as well. For a (3×3) case, Q is (as before),

$$Q = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + (a_{23} + a_{32})x_2x_3$$

and the matrix of partial derivatives $\nabla_{\mathbf{A}} Q$ is

$$\begin{aligned} \nabla_{\mathbf{A}} Q &= \begin{pmatrix} \frac{\partial Q}{\partial a_{11}} & \frac{\partial Q}{\partial a_{21}} & \frac{\partial Q}{\partial a_{31}} \\ \frac{\partial Q}{\partial a_{12}} & \frac{\partial Q}{\partial a_{22}} & \frac{\partial Q}{\partial a_{32}} \\ \frac{\partial Q}{\partial a_{13}} & \frac{\partial Q}{\partial a_{23}} & \frac{\partial Q}{\partial a_{33}} \end{pmatrix} \\ &= \begin{pmatrix} x_1x_1 & x_1x_2 & x_1x_3 \\ x_2x_1 & x_2x_2 & x_2x_3 \\ x_3x_1 & x_3x_2 & x_3x_3 \end{pmatrix} = \mathbf{x}\mathbf{x}^T. \end{aligned} \quad (16)$$

For the general case,

$$\frac{\partial}{\partial a_{ij}} \left(\sum_{r,s} x_r a_{rs} x_s \right) = \sum_s x_r x_s \left(\frac{\partial a_{rs}}{\partial a_{ij}} \right) = x_r x_s \delta_{ri} \delta_{sj}, \text{ (row and column indices have to match)}$$

Assignment Project Exam Help

<https://eduassistpro.github.io/>

Add WeChat edu_assist_pro