COMP9417 - Machine Learning
Math Revision

Introduction The goal of this optional self assessment is to review some important mathematical concepts that are used regularly in machine learning, and which are assumed knowledge for the course. If you find yourself struggling significantly with any aspects of this homework, please reach out to course staff so that we can better help you prepare for the course. Please also note that we have posted some helpful resources under the Week 0 tab on Moodle which may be of use to you for this homework.

Question 1. (Calculus Review)

   (a) Consider the function

$$f(x, y) = a_1 x^2 y^2 + a_4 xy + a_5 x + a_7$$

      compute all first and second order derivatives of f with respect to x and y.

   (b) Consider the function

$$f(x, y) = a_1 x^2 y^2 + a_2 x^2 y + a_3 xy^2 + a_4 xy + a_5 x + a_6 y + a_7$$

      compute all first and second order derivatives of f with respect to x and y.

   (c) Consider the logistic sigmoid:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

      show that $\sigma'(x) = \frac{\partial \sigma}{\partial x} = \sigma(x)(1 - \sigma(x))$

   (d) Consider the following functions:

- $y_1 = 4x^2 - 3x + 3$
- $y_2 = 3x^4 - 2x^3$
- $y_3 = 4x + \sqrt{1 - x}$
- $y_4 = x + x^{-1}$

      Using the second derivative test, find all local maximum and minimum points.

Question 2. (Probability Review)

   (a) A manufacturing company has two retail outlets. It is known that 20% of potential customers buy products from Outlet I alone, 10% buy from both I and II, and 40% buy from neither. Let A denote the event that a potential customer, randomly chosen, buys from outel I, and B the event that the customer buys from outlet II. Compute the following probabilities:

$$P(A), \quad P(B), \quad P(A \cup B), \quad P(\bar{A} \cap \bar{B})$$

      Here, $\bar{A}$ denotes the complement of A.

(b) Let X, Y be two discrete random variables, with joint probability mass function $P(X = x, Y = y)$ displayed in the table below:

		y		
		1	2	3
	1	1/6	1/12	1/12
x	2	1/6	0	1/6
	3	0	r	0

Compute the following quantities:

(i) r

(ii) $P(X = 2, Y = 3)$

(iii) $P(X = 3)$ and $P(X = 3 | Y = 2)$

(iv) $\mathbb{E}[X]$, $\mathbb{E}[Y]$ and $\mathbb{E}[XY]$

(v) $\mathbb{E}[X^2]$, $\mathbb{E}[Y^2]$

(vi) $\text{Cov}(X, Y)$

(vii) $\text{Var}(X)$ and $\text{Var}(Y)$

(viii) $\text{Corr}(X, Y)$

(ix) $\mathbb{E}[X + Y]$, $\mathbb{E}[X + Y^2]$, $\text{Var}(X + Y)$ and $\text{Var}(X + Y^2)$.

Question 3. (Linear Algebra Review)

(a) Write down the dimensions of the following objects:

$$A = \begin{bmatrix} 1 & 3 & 1 & 0 & 2 \\ 1 & 1 & 4 & 1 & 2 \\ 1 & 1 & 1 & 5 & 2 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \\ 3 \\ 2 \end{bmatrix}, \qquad A^T$$

(b) Consider the following objects:

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 2 & 1 \\ 6 & 4 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 7 & 3 & 3 \\ 2 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}, \quad D = \begin{bmatrix} 4 & 2 \\ 4 & 6 \\ 1 & 3 \end{bmatrix}, \quad u = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad v = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}$$

Compute the following:

(i) AB and BA

(ii) AC and CA

(iii) AD and DA

(iv) DC and CD and $D^T C$

(v) Bu and uB

(vi) Au

(vii) Av and vA

(viii) $Av + Cv$

(c) Consider the following objects:

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 2 & 1 \\ 6 & 4 & 3 \end{bmatrix}, \quad u = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad v = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}.$$

Compute the following:

(i) $\|u\|_1, \|u\|_2, \|u\|_2^2, \|u\|_\infty$

(ii) $\|v\|_1, \|v\|_2, \|v\|_2^2, \|v\|_\infty$

(iii) $\|v + w\|_1, \|v + w\|_2, \|v + w\|_\infty$

(iv) $\|Av\|_2, \|A(v - w)\|_\infty$

(d) Consider the following vectors in $\mathbb{R}^2$

$$u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad w = \begin{bmatrix} -1 \\ 1/2 \end{bmatrix}$$

Compute the dot products between all pairs of vectors. Note that the dot product may be written using the following equivalent forms:

$$\langle x, y \rangle = x \cdot y = x^T y.$$

Then compute the angle between the vectors and plot.

(e) Dot products are extremely important in machine learning, explain what it means for a dot product to be zero, positive or negative.

(f) Consider the 2×2 matrix:

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

Compute the inverse of A.

(g) Consider the 2×2 matrix

$$A = \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix}$$

Compute its inverse A^{-1}.

(h) Let X be a matrix (of any dimension), show that $X^T X$ is always symmetric.