The Australian National University
School of Computing

Semester 2, 2024
Theory Assignment 1

COMP3670/6670: Introduction to Machine Learning

Due Date. 5pm, Aug 28, 2024.

Late submissions will be permitted without penalty till 4:59pm Aug 29, 2024.

Late submissions beyond 4:59pm Aug 29, 2024 will encounter a 100% penalty.

College-approved extenuating circumstances permit assessment extensions which need to be filed online here before the assignment deadline, 5pm, Aug 28, 2024. Approval of extension requests are governed by university policies and will take into account any submitted proof of extenuating circumstances.

Maximum Credit. 30 (will be scaled down to a maximum of 10 points in the course by dividing the total credits by 3.)

Credit Breakdown. Each exercise will note the breakdown of credits for each of the parts and subparts within it.

Marking Rubric. The entire assignment will be graded out of a total of 30 credits.

100% You will receive 100% of the available credits for a question if your answer is completely correct **and** you show all the correct steps **and** your steps are clearly laid out.

50% You will receive 50% of the available credits for a question if your answer is not correct **or** you show mostly correct steps **and** your steps are clearly laid out.

0% You will otherwise receive 0% of the available credits for a question.

Sample partial solutions will be released with the grades after the submission deadline.

Exercise 1 **Properties of Matrices** (1+1+1+1+2+2 = 8 credits)

(a) Let $\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 5 & 6 \\ 4 & 6 & 8 \end{bmatrix}$ be a square matrix.

i) Show that $\mathbf{A}$ is symmetric and

ii) determine if it is positive definite.

(b) For the matrix above, $\mathbf{A}$,

i) compute the cube of $\mathbf{A}$ that is $\mathbf{A}^3$, and

ii) show that $\mathbf{A}^3$ is also symmetric.

(c) Let $\mathbf{A}$ be a square matrix and $f(\mathbf{X})$ be an n-th order polynomial, defined by $\sum_{i=0}^{n} a_i \mathbf{X}^i$ where a_i are arbitrary real numbers. Show that $f(\mathbf{A})$ commutes with $\mathbf{A}$, i.e., $f(\mathbf{A})\mathbf{A} = \mathbf{A}f(\mathbf{A})$.

(d) Let $\mathbf{A}$ and $\mathbf{B}$ be rectangular matrices of orders $n \times k$ and $r \times s$, respectively. The matrix of order $nr \times ks$ represented in a block form as

$$\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1k}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2k}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \dots & a_{nk}\mathbf{B} \end{bmatrix}$$

is called the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ of the matrices $\mathbf{A}$ and $\mathbf{B}$.

Find the conditions on a, b, c, d, e, f, g, h such that $\mathbf{X} \otimes \mathbf{Y} = \mathbf{Y} \otimes \mathbf{X}$, where $\mathbf{X}$ and $\mathbf{Y} \in \mathbb{R}^{2\times 2}$ defined as:

$$\mathbf{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

Exercise 2 **Solving Linear Systems** (2+2 = 4 credits)

Find the set $\mathcal{S}$ of all solutions $\mathbf{x}$ of the following linear systems $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{A}$ and $\mathbf{b}$ are defined as follows. Write the solution space $\mathcal{S}$ in parametric form.

(a)

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & -2 & 6 \\ 3 & 6 & -3 & 9 \\ 4 & 8 & -4 & 11 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}$$

(b)

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 & 3 & -1 \\ 2 & -2 & 4 & 6 & -2 \\ 3 & -3 & 6 & 9 & -3 \\ 4 & -4 & 8 & 11 & -4 \\ 5 & -5 & 10 & 14 & -5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

Exercise 3 **Inverses and Rank** (1+1 = 2 credits)

Let $\mathbf{A} \in \mathbb{M}_{n\times n}(\mathbb{R})$ be a matrix such that $\mathbf{A}^2 = \mathbf{A}$.

(a) Show that if $\mathbf{A}$ is invertible, then $\mathbf{A} = \mathbf{I}$.

(b) Show that $\mathbf{rk}(\mathbf{A}) = \mathbf{tr}(\mathbf{A})$, where $\mathbf{rk}(\mathbf{A})$ denotes the rank and $\mathbf{tr}(\mathbf{A})$ denotes the trace of $\mathbf{A}$. The trace of $\mathbf{A}$, $\mathbf{tr}(\mathbf{A})$, is defined as the sum of the diagonal elements of $\mathbf{A}$, i.e., $\mathbf{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$.

Exercise 4 **Subspaces** (1+1+2 = 4 credits)

(a) Which of the following sets are subspaces of $\mathbb{R}^n$? Prove your answer.

(i) $E = \left\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1^2 + x_2^2 + \ldots + x_n^2 \leq 1\right\}$

(ii) $F = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 x_2 \cdots x_n = 0\}$

(b) Let V be an inner product space, and let U and W be subspaces of V. Define $U + W = \{u + w : u \in U, w \in W\}$. Show that $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$.

Exercise 5 **Linear Transformations and Injectivity** (2+2 = 4 credits)

Let V and W be vector spaces, and let $T : V \to W$ be a **linear** transformation.

The *image* of T is defined as:

$$\mathbf{Im}(T) = \{w \in W \mid \exists v \in V \text{ such that } w = T(v)\}.$$

The *kernel* of T is defined as:

$$\mathbf{Ker}(T) = \{v \in V \mid T(v) = 0\}.$$

We say that T is *injective* if for all $u, v \in V$, $T(u) = T(v)$ implies $u = v$.

(a) Prove that T is injective if and only if $\mathbf{Ker}(T) = \{\mathbf{0}\}$.

(b) Consider the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & b \\ c & 0 & 1 \end{bmatrix}$$

with $a, b, c \in \mathbb{R}$. Find the condition on a, b, and c for which this transformation is injective.

Exercise 6 **Linear Transformations and Inner Products** (2+2+2+2 = 8 credits)

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation defined by the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

(a) Show that T is not an orthogonal transformation.

(b) Consider the inner product defined by the matrix

$$\mathbf{D} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Show that this inner product is preserved under T, i.e., for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, we have $(T\mathbf{x})^T\mathbf{D}(T\mathbf{y}) = \mathbf{x}^T\mathbf{D}\mathbf{y}$.

(c) Let $\mathbf{u} = [1, 1]^T$ and $\mathbf{v} = [1, -1]^T$.

i) Compute the angle between $\mathbf{u}$ and $\mathbf{v}$ under the inner product defined by $\mathbf{D}$ and

ii) show that this angle is preserved under the transformation T.