AI5031: Machine Learning, exercise sheet 1

1 Notation and math primitives

Compute the resulting value of the following mathematical expressions:

1. $\sum_{i=0}^{10} i$
2. $\prod_{i=1}^{5} i$
3. $\sum_{i=0}^{4} i^2$
4. $\sum_{i=1}^{3} x_i y_i$ if $\vec{x} = (1, 2, 3)^T$ and $\vec{y} = (2, 1, 0)^T$
5. $\sum_{i=1}^{3} (x_i + y_i)$ if $\vec{x} = (1, 2, 3)^T$ and $\vec{y} = (2, 1, 0)^T$
6. $\frac{1}{3} \sum_{i=1}^{3} x_i^2$ if $\vec{x} = (1, 2, 3)^T$

2 Matrix multiplication

We have column vectors $\vec{x} = (1, 2, 3, 0)^T$, $\vec{y} = (2, 3, 4, 5)^T$, $\vec{z} = (2, 3, 0)^T$ and a matrix $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{pmatrix}$. Give the result of the following matrix multiplications if they are legally possible:

1. $A\vec{y}$
2. $A\vec{x}$
3. $\vec{x}^T \vec{y}$ (note: this is another way of writing the scalar product)
4. $\vec{x}A$
5. $A\vec{y}^T$
6. $\vec{x}^T A\vec{y}$ (note: this is called the *Mahalanobis distance* of two vectors given a matrix A)
7. $\vec{x}\vec{y}^T$ (note: this is called the *outer product* of two vectors)
8. $A\vec{z}$
9. $\vec{z}\vec{x}^T$
10. AA
11. AA^T

3 Bonus: Python

Assume we model 2D matrices by row-wise nested lists. That is, a matrix $\begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 4 & 4 \end{pmatrix}$ is represented by a list [[1,1],[2,2],[3,3],[4,4]]. Or a row vector $[1, 2]$ is represented as [[1,2]], and a column vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ as [[1],[2]]. Create a Python function *matmul* which takes two arguments x and y and reproduces the results from the previous exercise. If a matrix multiplication is not possible, it should return *None*. Please not that you can obtain the number of rows of a nested list x as *len(x)*, whereas the nr of columns is obtained as *len(x[0])*. You can do this exercise in the CodeRunner section on our E-Learning site.

4 Functions

The Rectified Linear Unit (ReLU) function $f : x \in \mathbb{R} \mapsto y \in \mathbb{R}$ is defined as $f(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}$ Assuming the quantities from the previous exercise and the matrix $B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, compute:

1. $f(\vec{x})$
2. $f(B\vec{x})$
3. $f(B)$
4. $f(\vec{x}^T\vec{y})$

5 Vector-scalar and vector-vector functions

Let $f(x) =< \vec{x}, (1,1,0,0)^T >$ be a vector-scalar function and $g_i(\vec{x}) = x_i^2$ a vector-vector function. Please compute:

1. $f\left((1,1,1,1)^T\right)$
2. $f(B)$ using the matrix B from the previous exercise
3. $f(B^T)$ using the matrix B from the previous exercise
4. $g\left((1,2,3)^T\right)$
5. $g(B)$ using B from the previous exercise

6 Bonus: A simple machine learning model

A classical problem in machine learning is the OR-problem. We want a ML model which returns 0 (false) or 1 (true), $f : \vec{x} \in \mathbb{R}^2 \mapsto y \in \mathbb{R}$, to learn the logical OR mapping, that is:

$$0, 1 \rightarrow 1 \tag{1}$$
$$1, 0 \rightarrow 1 \tag{2}$$
$$0, 0 \rightarrow 0 \tag{3}$$
$$1, 1 \rightarrow 1 \tag{4}$$

a) Formulate data and target matrices for this problem!
b) Can you find parameters $\vec{w} = (w_1, w_2, w_3)^T$ for the following model function that solve the problem? It may be helpful to draw the data points into a 2D plot.

$$f(\vec{x}) = \begin{cases} 1 & w_1 x_1 + w_2 x_2 + w_3 > 0 \\ 0 & w_1 x_1 + w_2 x_2 + w_3 \leq 0 \end{cases} \tag{5}$$