Graded Assignment 1

Math Basics for Machine Learning

Fall 2022

Instructions

This is the first graded assignment for the Math Basics for Machine Learning course. It contains three tasks. The instructions, as well as links to supplementary material, are given in the task descriptions.

You should give the final answer to every question and submit your solutions by filling in the corresponding Google form. At the end of the form, you will also need to upload the **detailed solutions** to the tasks in this assignment as a single file. The best way to do it is to type them using a Latex editor.

For some tasks, it might be convenient to use Python rather than performing the computations by hand. If you do so, attach your code as well (e.g., save a notebook with code to .pdf or simply make screenshots and add them to the file with other solutions).

The form will be graded automatically, and the scores will be revealed after manual revision. In total, you can earn 8 points for this assignment. This score will contribute to you final score for this course.

You must submit your answers by **Monday, October 17, 23:59 Anywhere on Earth**. *Late submissions will* ***not*** *be accepted.*

It is the idea that you complete this assignment individually. Do not collaborate or copy answers of somebody else.

Have fun!

1 Linear (in)dependence

During the lectures, we discussed the concept of linear (in)dependence and saw examples of linearly (in)dependent sets of vectors.

A general way to check linear independence of a given set of vectors is to combine those vectors in a matrix and bring it to reduced row echelon form (RREF) with the help of elementary row operations.

If you are not yet familiar with this procedure, you can learn more about bringing a matrix to its RREF in this video. There is also another video that demonstrates how we can test a set of vectors for linear independence.

Once you are ready you can solve the questions in this section.

1.1 (1 point) Dimensionality of a subspace

Consider vectors v_1, v_2, v_3 and v_4 defined below:

$$v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix}, \; v_3 = \begin{bmatrix} 2 \\ 4 \\ 1 \\ 3 \end{bmatrix}, \; v_4 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}$$

What is the dimensionality of the subspace $V = span\{v_1, v_2, v_3, v_4\}$ spanned by those vectors?

1.2 (2 points) Basis of a subspace

Consider subspace $V = span\{v_1, v_2, v_3, v_4\}$ from the previous question. Which sets of vectors form a basis in the vector space V?

- ☐ $\{v_1\}$
- ☐ $\{v_2\}$
- ☐ $\{v_3\}$
- ☐ $\{v_4\}$
- ☐ $\{v_1, v_2\}$
- ☐ $\{v_1, v_3\}$
- ☐ $\{v_1, v_4\}$
- ☐ $\{v_2, v_3\}$
- ☐ $\{v_2, v_4\}$
- ☐ $\{v_3, v_4\}$
- ☐ $\{v_1, v_2, v_3\}$
- ☐ $\{v_1, v_2, v_4\}$
- ☐ $\{v_1, v_3, v_4\}$
- ☐ $\{v_2, v_3, v_4\}$
- ☐ $\{v_1, v_2, v_3, v_4\}$

Select all valid bases.

1.3 (1 point) Orthogonalization

Vector set $B = \{u, v\}$, where $u = [1, 2]$ and $v = [6, 2]$, is a basis in $\mathbb{R}^2$.

Using the Gram-Schmidt process, we have transformed B into an orthogonal basis $B' = \{u', v\}$. If $u' = u$, what are the coordinates of v'?

2 Distances and angles

Given a vector a and a set of vectors $X = x_1, ..., x_m$, we say that x_i is the *nearest neighbor* of a if it's the closest vector to a from the set X. In ML, the idea of the nearest neighbour and its generalisation, k-nearest neighbours, is used for solving classification and regression problems (see k-nearest neighbours algorithm).

There are various ways of defining distance, or similarity, between two vectors. The most obvious choice is Euclidean distance. However, angle between the vectors is also used in some applications (see cosine similarity in the Word vectors notebook from the 1st practical session).

In the next three questions, you'll practice the notion of the nearest neighbour.

2.1 (1 point) Nearest neighbour

Find the nearest neighbour of $a = [3, 1, 4]$ from the set of vectors below. Use euclidean distance.

☐ $x_1 = \begin{bmatrix} 4 \\ 3 \\ 6 \end{bmatrix}$

☐ $x_2 = \begin{bmatrix} 3 \\ 1 \\ 9 \end{bmatrix}$

☐ $x_3 = \begin{bmatrix} 1 \\ 4 \\ 10 \end{bmatrix}$

☐ $x_4 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$

If there is more than one nearest neighbour, mark all of them.

2.2 (1 point) Distance

In the previous question, what is the distance from vector a to its nearest neighbour?

2.3 (1 point) Angles

Which of the vector(s) $x_1, ..., x_4$ from above make(s) the largest angle with the vector a?

3 Linear classifier

In the first lecture, we have briefly discussed the problem of *binary classification*, where each example needs to be classified into one of the two classes.

We have seen how how linear classifiers perform this task with a separating hyperplane.

3.1 (1 point) Separating hyperplane

Consider a linear classifier with the separating hyperplane $x_1 + x_2 = 1.5$.

According to the classifier above, which of the following points belong to the same class?

	x_1	x_2
a	0	0
b	1	0
c	1	1
d	2	0
e	2	2

Provide a set of points for each possible class.