Introduction to Data Science - Exercises #2

Jun. Prof. Marlene Kretschmer, Dr. Marina Friedel, Dr. Julia Mindlin, Dr. Karin Mora

Oct. 27th, 2025
present your results on Nov 3rd

Exercise 1

Write $\vec{x} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$ as a linear combination of basis vectors $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$.
Hint: basis vectors are like Lego bricks, with which you can create any vector in the corresponding vector space.

Exercise 2

Are the vectors $\vec{x} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $\vec{y} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\vec{z} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ a basis of $\mathbb{R}^3$?

Exercise 3

A company manufactures two products. For 1€ worth of product B, the company spends 0.45€ on materials, 0.25€ on labour, and 0.15€ on overhead. For 1€ worth of product C, the company spends 0.40€ on materials, 0.30€ on labour, and 0.15€ on overhead. The costs for each product can be written as vectors

$$\vec{b} = \begin{pmatrix} 0.45 \\ 0.25 \\ 0.15 \end{pmatrix} \text{ for product B and } \quad \vec{c} = \begin{pmatrix} 0.40 \\ 0.30 \\ 0.15 \end{pmatrix} \text{ for product C.}$$

1. What is the economic interpretation of $100 \cdot \vec{b}$?
   [Hint: Scalar multiplication]

2. The company wishes to manufacture 100€ worth of product B and 200€ worth of product C. What is the total cost for materials, labour, and overhead, respectively?
   [Hint: Scalar multiplication and vector addition]

Exercise 4

For the matrices $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}, B = \begin{pmatrix} -3 & 0 \\ 0 & -2 \end{pmatrix}, C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, D = \begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$ and vector $\vec{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, compute $A \cdot \vec{x}$, $B \cdot \vec{x}, C \cdot \vec{x}$ and $D \cdot \vec{x}$. Draw the vector before and after the matrix multiplication into a coordinate system. What kind of transformation happens to the original vector $\vec{x}$ after the matrix multiplication is applied?

Exercise 5

For the matrices $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}$ and $C = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 1 \end{pmatrix}$, compute all possible pairwise products (i.e. $A \cdot A, A \cdot B, B \cdot A, A \cdot C, \dots$).

Coding Exercise 1

[Learning outcome: writing matrices in python, matrix operations (addition, multiplication)]
A research team sampled three species of fish (sp1, sp2, sp3) at five sites in the Baltic. They performed their measurements once a month during the summer/autumn period (northern hemisphere):

July	sp1	sp2	sp3
Site 1	1	5	35
Site 2	14	2	0
Site 3	0	31	67
Site 4	96	110	78
Site 5	0	0	0

August	sp1	sp2	sp3
Site 1	15	23	10
Site 2	54	96	240
Site 3	0	3	9
Site 4	12	31	27
Site 5	8	14	6

September	sp1	sp2	sp3
Site 1	48	78	170
Site 2	2	0	0
Site 3	0	11	14
Site 4	25	13	12
Site 5	131	96	43

1. Compute the total number of fish (per species) surveyed at each site over three months using matrices?
   Hint: to work with matrices in python use `numpy.array`.

2. The researchers realised that the gear to sample the fish populations is faulty, i.e. the efficiency turned out not to be comparable for the three species. The efficiency of the net was 50% for species 2, and 25% for species 3 of what it was for species 1, respectively. How can the researchers still obtain a reliable estimate of the total fish abundance (for all species together) for

   (a) July per site?

   (b) August per site?

   (c) September per site?

   Hint: use python arrays and matrix-vector multiplication `numpy.matmul`.

Coding Exercise 2

Wind is one of the most extensively studied variables in atmospheric science. Unlike other fields such as temperature or precipitation, which are scalar fields, wind is a vector field (it not only has a speed (magnitude of the vector), it also has a direction).
Vertical motions in the atmosphere are much slower than horizontal motions. Therefore, we can represent wind with a 2D field, which consists of zonal (u) and meridional (v) components. This can be expressed as follows:

$$\vec{V}(x, y) = \vec{u}(x, y)\hat{\imath} + \vec{v}(x, y)\hat{\jmath} \tag{1}$$

Here, $\vec{V}(x, y)$ represents the wind vector field as a function of position (x, y), $\vec{u}(x, y)$ is the zonal wind component (east-west direction), $\vec{v}(x, y)$ is the meridional wind component (north-south direction). $\hat{\imath}$ and $\hat{\jmath}$ are unit vectors in the eastward and northward directions, respectively.
The files `exercise2_u_component_wind.csv` and `exercise2_v_component_wind.csv` contain the monthly average of the zonal and meridional components of wind for the year 2023. The spatial sampling of the latitudes (-90, 90) and longitudes (0, 360) is 2.5 degrees.

1. Plot the wind vectors for each month. Where do you see northerly (north to south) winds? And easterly (east to west) winds?

2. Write a code to find the grid cells where the wind direction is easterly.

Hints: Use operations between numpy arrays. Use np.meshgrid() and plt.pcolormesh to plot the wind map. Use plt.quiver() to plot the wind vectors. Use np.where() for conditional operations.