Exercises On Vectors, Inner Products, norms, distances

Alexandros Kalousis

University of Applied Sciences, WS

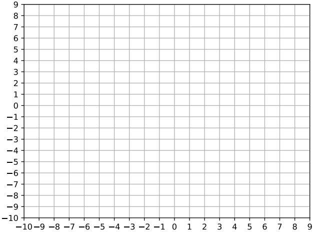
Alexandros.Kalousis@hesge.ch

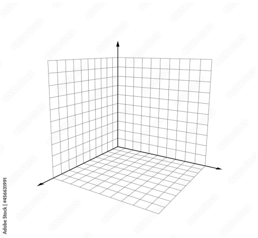
- All programming should be done in puthon
- You should extensively use numpy, numpy arrays, and .dot() to manipulate vectors and matrices.
- Loops are to be avoided at any cost, you should think in terms of operators on vectors and matrices.
- You should do the following exercises: a) b) c) g) h).
- You can work in teams of two.

Vectors I

Exercise a) : Vectors (all exercises in the Vectors Slide Set)

- Given some vector $\boldsymbol{a} \in \mathbb{R}^d$, what is the dimensionality of the vector? what is the difference between $\boldsymbol{a}$ and $\boldsymbol{a}^{\mathsf{T}}$
- Give two examples of vectors for each of the following ds, $d = 1, 2, 3, 4$. Manually visualise these vectors and comment your visualisation.





For your visualisations use grids such as the above when appropriate.

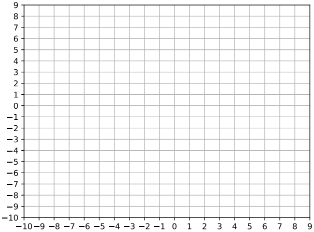
Vectors II

- Given the displacement vector $[2, 3]^{\mathsf{T}}$ draw it at the following starting positions: $\boldsymbol{s}_1 = [0, 0]\mathsf{T}, \boldsymbol{s}_2 = [1, 3]\mathsf{T}, \boldsymbol{s}_1 = [-1, 2]\mathsf{T}$. Comment on the different vectors that you obtain, what is common between them? are they really different?

Vector Operations I

Exercise b) : Vector Operations (all exercises in the Vector Operations Slide Set)

- Perform the following vector operations when possible and visualise each time the original vectors as well as how we manipulate them in order to obtain the final result, use grids such as the following for your visualisations.



Vector Operations II

$$\boldsymbol{a} = \begin{bmatrix}1\\3\end{bmatrix} + \begin{bmatrix}-1\\2\end{bmatrix}, \boldsymbol{b} = \begin{bmatrix}2\\1\end{bmatrix} - \begin{bmatrix}1\\3\end{bmatrix}, \boldsymbol{c} = \begin{bmatrix}2\\3\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix} + \begin{bmatrix}-1\\-2\end{bmatrix},$$

$$\boldsymbol{e} = 2\begin{bmatrix}2\\4\end{bmatrix}, \boldsymbol{e} = -1\begin{bmatrix}2\\4\end{bmatrix}, \boldsymbol{f} = \frac{1}{2}\begin{bmatrix}2\\4\end{bmatrix}, \boldsymbol{g} = \begin{bmatrix}2\\4\end{bmatrix} + \begin{bmatrix}1\\3\\2\end{bmatrix}, \boldsymbol{h} = \begin{bmatrix}2\\4\end{bmatrix} + \begin{bmatrix}1\end{bmatrix}$$

- explain what happens to a vector $\boldsymbol{a} \in \mathbb{R}^d$ when we multiply it with a scalar $b \in \mathbb{R}$, by exploring the following cases: $|b| < 1, |b| > 1$

Vector Operations III

- give examples for the following linear combinations of vectors,
  - three vectors in the two dimensional space, $\mathbb{R}^2$, visualise the original vectors, and how you obtain the final results.
  - three vectors in the three dimensional space, $\mathbb{R}^3$
- Explain what are unit vectors for the $\mathbb{R}^d$ space. Show that any vector $\boldsymbol{a} \in \mathbb{R}^d$ can be written as a linear combinations of the unit vectors $\boldsymbol{e}_i \in \mathbb{R}^d$.

Vector Operation Properties I

Exercise c) : Vector Operation Properties (all exercises in the Vector Operation Properties Slide Set)

- demonstrate that the following properties hold for any vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^d$ and scalars $x, y, z \in \mathbb{R}$:
  - Vector addition is commutative: $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$
  - Scalar multiplication is associative $(xy)\boldsymbol{a} = x(y\boldsymbol{a})$
  - Scalar multiplication is left distributitive $(x + y)\boldsymbol{a} = x\boldsymbol{a} + y\boldsymbol{a}$

Basic vector exercises

Exercise d) : Basic vector exercises (all exercises in the Basic vector exercises slide set)

- do exercises 1.2, 1.5, 1.10, 1.15 from first chapter of the [1] book

Linear combinations of vectors I

Exercise e) : Implementing linear combinations of vectors (all exercises in Linear combinations of vectors slide set)

- Exercise write a python function which will take as arguments
  - two $2-d$ vectors, $\mathbf{a}, \mathbf{b}$
  - two multipliers, one for each vector, contained in the $2-d$ vector $\mathbf{m}$
- and then will:
  - Plot the two vectors
  - Plot their scaled versions
  - Plot how addition happens over the scaled version by moving one vector at the end of the other and then drawing the sum.
- Make sure that the two axes have exactly the same scale.
- Consider using the matplotlib quiver function to draw the vectors.

Linear combinations of vectors II

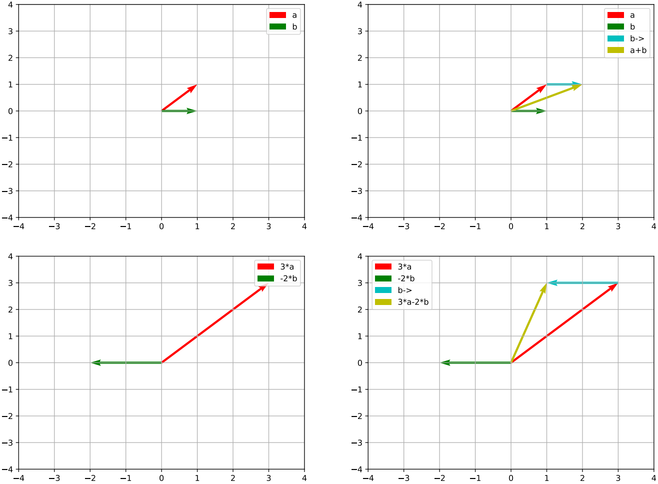
- The signature of the function should be something like:

```
def addABvectors(a=[1,0],b=[0,1],m=[1,1],origin=[0,0],maxes=[3,3],outputFileStem="result"):
    """
    Get two vectors a,b, and two multipliers and:
      1) Plot a and b and stores in
         file outputFileStemVectors.pdf
      2) Plot m[0]*a and m[1]*b and stores in
         file outputFileStemVectorsScaled.pdf
      3) Show how addition is done by placing
         one at the end of the other and stores in
         file outputFileStemVectors...
      4) Plot the sum c[0]*a and c[1]*b and stores in file ...

    Keyword arguments:
    a           -- 2-d vector, the first summand
    b           -- 2-d vector, the second summand
    m           -- a vector with two scalars with which a and b
                   are multiplied
    origin      -- the point from from which the displacement
                   vector starts
    maxes       -- the axes limits, usefull if we are to
                   draw more than one figures and we want the
                   limits to be the same
    """
```

Linear combinations of vectors III

The outputs should look like the following figures for $\mathbf{a} = [1,1]^T, \mathbf{b} = [1,0]^T, \mathbf{m} = [1,1]^T$ first line, and $\mathbf{m} = [3,-2]^T$ second line, and origin=$[0,0]^T$.



Linear combinations of vectors IV

- For each one of the four figures in the previous slide explain what are the vectors that are given in them.
- Visualise the results for $\mathbf{a} = [3,5]^T, \mathbf{b} = a, \mathbf{m} = [1,-1]^T$, explain what happens.
- Given a vector $\mathbf{a}$, e.g. $\mathbf{a} = [3,5]^T$ propose a mathematical formulation that allows you to find vectors $\mathbf{b}$ for which $\mathbf{a}^T\mathbf{b} = 0$. Propose different $\mathbf{b}$ that satisfy the condtion $\mathbf{a}^T\mathbf{b} = 0$ and draw them together with $\mathbf{b}$. Repeat the exercise with different $\mathbf{a}$ vectors. What do you observe?
- In the same graph draw the displacement vector $\mathbf{a} = [3,5]^T$ starting from different points, $\mathbf{o}_1 = [0,0], \mathbf{o}_2 = [2,1], \mathbf{o}_3 = [-2,-1], \mathbf{o}_4 = [-2,1], \mathbf{o}_5 = [2,-1]$. You might need to write a new function to do that. Comment on the figure, what is common between the different vectors? are they really different?

Generic linear combinations

Exercise f) : implementing the general linear combination of vectors case.

- Implement a more generic function that takes as input:
  - a matrix $\mathbf{X}$ of dimensionality : $n \times d$ where the ith line corresponds to the vector $\mathbf{x}_i^T$ of dimensionality d.
  - a vector m of dimensionality n containing one multiplier for each of the n vectors of X.

  and produces as output:
  - $\sum_{i=1}^{i=n} m_i \mathbf{x}_i$

  You are not allowed to use loops.

Inner products, distances, norms I

Exercise g) : Inner product (all exercises in the Inner product slide set)

- Compute explicitly, i.e. show the full development, for the inner product of vectors for each $(\boldsymbol{a}, \boldsymbol{b})$ vector pair bellow

$$([1], [-1]), (\begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}-1\\2\end{bmatrix}), (\begin{bmatrix}2\\1\end{bmatrix}, -\begin{bmatrix}1\\3\end{bmatrix}), (\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}1\\3\end{bmatrix}),$$
$$(\begin{bmatrix}-1\\-2\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}), (\begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}2\\4\\1\end{bmatrix})$$

- Given the $\boldsymbol{a} = [1, 2]^{\mathsf{T}}$ vector, find three different vectors, $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$, for which $\boldsymbol{a}^{\mathsf{T}}\boldsymbol{b}_i = 0$. Draw $\boldsymbol{a}$ and the $\boldsymbol{b}_i$, what do you observe? what is the relation between $\boldsymbol{a}$ and each one of the $\boldsymbol{b}_i$s.

Inner products, distances, norms II

- For the following pairs of vectors find the angle they form and explain how you do it:

$$([1,1]^\mathsf{T},[1,2]^\mathsf{T}),([1,1]^\mathsf{T},[-1,-1]^\mathsf{T}),([1,1]^\mathsf{T},[2,2]^\mathsf{T})$$

- For each of the above vector pairs compute the distances of the involved vectors, visualise the vectors and explain to what does the distance correspond to.
- Compute the norm for the vectors

$$\begin{bmatrix}2\\3\end{bmatrix},\begin{bmatrix}2\\4\\1\end{bmatrix}$$

show explicitly how you compute the norm and explain what does the norm give you?

Inner products, distances, norms III

- Show how can you produce a vector that has the same direction as $[2, 3]^\mathsf{T}$ and a norm of 1.
- Given the vectors: $\boldsymbol{a} = [2, 4]^\mathsf{T}$ and $\boldsymbol{b} = [1, 1]^\mathsf{T}$, explain what is the projection of $\boldsymbol{a}$ on $\boldsymbol{b}$, show how to compute it, compute it, and visualise it.
- Show how you can compute the average of the elements of a vector $[a_1, a_2, \ldots, a_d]^\mathsf{T}$ using only scalar multiplication and inner products.
- What is the norm of the $\frac{\boldsymbol{a}}{||\boldsymbol{a}||}$ vector?

The nearest neighbor classification algorithm I

Exercise h) : the k-NN algorithm
You should implement in python and numpy a simple function, kNN, that will take as input:

- a matrix $\boldsymbol{X}$: $n \times d$ which contains the predictive attributes
- a vector $\boldsymbol{y} \in \mathbb{R}^n$ which contains the class labels
- an instance $\boldsymbol{x} \in \mathbb{R}^d$ that you should classify
- k a scalar that defines the number of neighbors to use for the classification.
- *vis* a boolean that indicates whether you should visualise or not the reults
- your function should contain no loops.

the function should find the k nearest neighbors of $\boldsymbol{x}$and use them to determine the class value for $\boldsymbol{x}$.

The nearest neighbor classification algorithm II

If *vis* $=$ *TRUE* your function should visualise the decision process by doing the following:

- it should operate only over the first two dimensions of $\boldsymbol{X}$ and do the computations over these
- if should plot all instances of $\boldsymbol{X}$ using the circle marker (o) where each class will have a different color
- it should plot the instance $\boldsymbol{x}$ to be classified using the triangle down marker (v)
- it should plot the k-nearest neighbors of $\boldsymbol{x}$ using the triangle up marker $(\hat{})$
- it should plot the lines that correspond to the distances of $\boldsymbol{x}$ from each one of its k nearest neighbors.

The nearest neighbor classification algorithm III

You should demonstrate the use of your kNN function on the iris dataset as follows:

- Randomly select half of the iris dataset to become what we call the training set (this will be your $\boldsymbol{X}$ and $\boldsymbol{y}$), the other half will be your test set and you will use it to select the $\boldsymbol{x}$ vector.
- Randomly select one instance from the test set which will be the $\boldsymbol{x}$ instance to classify
- Classify $\boldsymbol{x}$ using $k = 1, k = 3, k = 5$ nearest neighbors. Compare the answer of your kNN function with the true label of your $\boldsymbol{x}$ instance.
- Visualise the results for $k = 1$ and $k = 3$

Inner product properties

Exercise i) : Inner product properties

- Show the following basic properties of inner products
  - $\mathbf{x}^\mathsf{T}\mathbf{y} = \mathbf{y}^\mathsf{T}\mathbf{x}$
  - $(a\mathbf{x})^\mathsf{T}\mathbf{y} = \mathbf{x}^\mathsf{T}(a\mathbf{y})$
  - $(\mathbf{x}+\mathbf{y})^\mathsf{T}\mathbf{z} = \mathbf{x}^\mathsf{T}\mathbf{z} + \mathbf{y}^\mathsf{T}\mathbf{z}$

  where $a \in R, \mathbf{x}, \mathbf{y}, \mathbf{z} \in R^d$
- use the inner product only (no broadcasting) to compute
  - the average of the elements of a vector $\mathbf{x}$
  - the variance of a vector
  - the weigthed average of the x_2, x_5, x_7 elements where $w_2 = 1, w_5 = 7, w_7 = 2$

The unit circle

Exercise j) : the unit circle

- The unit circle (2-d space) is the circle the center of which is at the origin of the axes and it has a radius of 1.
- What is the property of the points that lie on the unit sphere?
- How to generate a set of points that lie in the unit circle?
- When we move to 3-d we have the unit sphere
- For $d > 3$ we have the unit hypersphere.

What can you say about the $\frac{\mathbf{x}}{||\mathbf{x}||}$

Exercise k) : $\frac{\mathbf{x}}{||\mathbf{x}||}$

- What does the $\frac{\mathbf{x}}{||\mathbf{x}||}$ do?
- What is the result ? and what are the properties of the result? does it have a norm and if yes what is its value? Can you provide the proof of the result?
- Draw the relation between $\mathbf{x}$ and $\frac{\mathbf{x}}{||\mathbf{x}||}$.

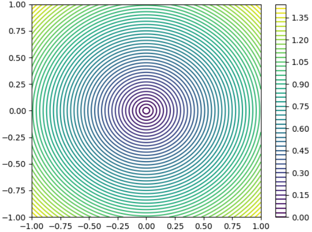
Exercise: Visualising the norms I

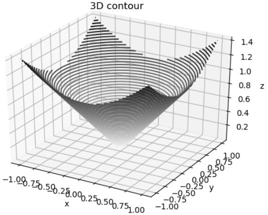
Exercise I) : Norm visualisation

- Write a function *CreateUnitCircleSet(n)* which returns a matrix with dimensionality $n \times 2$ instances which cover in a regular manner the unit circle (no loops are allowed). Plot the returned set of points. Note that we will reuse the function that generates the unit circle points at least once more.
- Draw the level sets of the norm function, i.e. $||\mathbf{x}|| = \mathbf{x}^\mathsf{T}\mathbf{x}, \mathbf{x} \in [-1, 1] \times [-1, 1]$, in a two-dimensional and a three-dimensional graph[1]. Explain the two graphs and their relation with the graph produced in the question above.
- You should think of the norm of a vector just like any other function. Define your own two-dimensional function and draw as above its level sets in two and three dimensions.

Exercise: Visualising the norms II

Your level set plots should like the following for the norm





---

[1] Use functions such as contour and contour3d

Exercise: Distance as an inner product

Exercise m) : distance and inner product

- show how you can compute the distance of two vector $\mathbf{x}_1$ and $\mathbf{x}_2$ with the help of the inner product operation, and show how to do that in numpy.

More on inner products I

Exercise n) : more on inner products

1. Using basic trigonometry show that the signed component (signed length) of the vector **b** along the vector a is given by $d = \frac{\mathbf{a}^T\mathbf{b}}{||\mathbf{a}||}$
2. Let x_i and x_j be two random variables. We are given the random sample $\{x_{i_k}, x_{j_k} | k = 1...n\}$, this is essentially a dataset with two attributes (x_i, x_j) and n instances. Explain the relation between:

$$\rho = \frac{E[(x_i - \mu_i)(x_j - \mu_j)]}{\sigma_{x_i}\sigma_{x_j}} \tag{1}$$

and

$$\hat{\rho} = \frac{\sum_{k=1}^{n}(x_{i_k} - \mu_i)(x_{j_k} - \mu_j)}{\sqrt{\sum_{k=1}^{n}(x_{i_k} - \mu_i)^2}\sqrt{\sum_{k=1}^{n}(x_{j_k} - \mu_j)^2}} \tag{2}$$

why in the denominator of eq 2 we do not have the sample based estimate of the standard deviation, i.e. $\sqrt{\frac{\sum_{k=1}^{n}(x_{j_k} - \mu_j)^2}{n-1}}$

More on inner products II

3. show why $\hat{\rho}$ can be computed as:

$$\begin{aligned} \tilde{\mathbf{x}} &= \mathbf{x} - \mathbf{avg}(\mathbf{x})\mathbf{1} \\ \hat{\rho} &= \frac{\tilde{\mathbf{x}}_i{}^T \tilde{\mathbf{x}}_j}{||\tilde{\mathbf{x}}_i|| ||\tilde{\mathbf{x}}_j||} \end{aligned} \tag{3}$$

where $\mathbf{x}^T = [x_1, x_2, ..., x_n]$. Explain the meaning of eq 3, i.e. ρ can be computed as an inner product of which vectors? under that view what is the interpretation of $\hat{r}ho$?

References

S. Boyd and L. Vandenberghe.
Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares.
Cambridge University Press, 2018.