

COMP337+527
Data Mining and Visualisation
Problem Set 0

Shagufta Scanlon

Question 1 Consider two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ defined as $\mathbf{x} = (1, 2, -1)^\top$ and $\mathbf{y} = (-1, 0, 1)^\top$. Answer the following questions about these two vectors.

- A. Compute the length (ℓ_2 norm) of \mathbf{x} and \mathbf{y} . (4 marks)

$$\|\mathbf{x}\|_2 = \sqrt{1+4+1} = \sqrt{6} \text{ and } \|\mathbf{y}\|_2 = \sqrt{1+0+1} = \sqrt{2}$$

- B. Compute the inner product between \mathbf{x} and \mathbf{y} . (2 marks)

$$\mathbf{x}^\top \mathbf{y} = -1 + 0 + -1 = -2$$

- C. Compute the cosine of the angle between the two vectors \mathbf{x} and \mathbf{y} . (4 marks)

The definition of cosine similarity is $\frac{\mathbf{x}^\top \mathbf{y}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2}$. Therefore, the required value will be $-2/\sqrt{12}$.

- D. Compute the Euclidean distance between the end points corresponding to the two vectors \mathbf{x} and \mathbf{y} . (4 marks)

The definition of the Euclidean distance is $\sqrt{\sum_i (x_i - y_i)^2}$. Therefore, we get $\sqrt{4+4+4} = 2\sqrt{2}$

- E. For any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ such that $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$ show that the following relationship holds between their cosine similarity $\cos(\mathbf{x}, \mathbf{y})$ and their Euclidean distance $\text{Euc}(\mathbf{x}, \mathbf{y})$. (6 marks)

$$\text{Euc}(\mathbf{x}, \mathbf{y})^2 = 2(1 - \cos(\mathbf{x}, \mathbf{y}))$$

$$\begin{aligned} \text{Euc}(\mathbf{x}, \mathbf{y})^2 &= (\mathbf{x} - \mathbf{y})^\top (\mathbf{x} - \mathbf{y}) \\ &= \mathbf{x}^\top \mathbf{x} + \mathbf{y}^\top \mathbf{y} - 2\mathbf{x}^\top \mathbf{y} \\ &= 1 + 1 - 2\cos(\mathbf{x}, \mathbf{y}) \\ &= 2(1 - \cos(\mathbf{x}, \mathbf{y})) \quad \square \end{aligned}$$

Question 2 Consider a matrix $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ defined as follows:

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Answer the following questions related to \mathbf{A} .

- A. Compute the transpose \mathbf{A}^\top . (2 marks)

For a matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\mathbf{A}^\top = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Therefore, we have

$$\mathbf{A}^\top = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

- B. Compute the determinant $\det(\mathbf{A})$. (2 marks)

$$\det(\mathbf{A}) = ac - bd = 2 \times 2 - 1 \times 1 = 3$$

- C. Compute the inverse \mathbf{A}^{-1} . (4 marks)

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

From which it follows,

$$\mathbf{A}^{-1} = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}.$$

- D. Compute the eigenvalues and eigenvectors of \mathbf{A} . (6 marks)

Eigenvector \mathbf{x} corresponding to the eigenvalue λ satisfies the equation $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. From which it follows that $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$. Therefore, $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. In this case, we get $\det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = 0$. Solving this second-order polynomial equation we get $\lambda = 1, 3$, which are the eigenvalues. Substituting these values separately in the eigenvalue equation we get the eigenvectors corresponding $\lambda = 1$ and $\lambda = 3$ to be respectively $(1, -1)^\top$ and $(1, 1)^\top$, subjected to a scaling factor.