Machine Learning

HW1 - Dimensionality Reduction
Spring 2025

1. Write your name below to acknowedge that you should not include any hand-written solution, code, or pictures (unless explicity requested), in your HW solutions. That everything should be typeset and converted to PDF. (1pt)

   Ali Ural

2. Convert the follow features to binary, using their means: (2pts)

$$X = \begin{bmatrix} 2 & 1 \\ 3.5 & 2 \\ 0 & 3 \\ 12 & 4.1 \end{bmatrix}$$

   Answer for Q1:

   Calculating means:

$$\mu_1 = \frac{2 + 3.5 + 0 + 12}{4} = \frac{17.5}{4} = 4.375$$

$$\mu_2 = \frac{1 + 2 + 3 + 4.1}{4} = \frac{10.1}{4} = 2.525$$

   Comparing each column to mean:

$$\begin{bmatrix} 2 < 4.375 & \Rightarrow 0 & 1 < 2.525 & \Rightarrow 0 \\ 3.5 < 4.375 & \Rightarrow 0 & 2 < 2.525 & \Rightarrow 0 \\ 0 < 4.375 & \Rightarrow 0 & 3 > 2.525 & \Rightarrow 1 \\ 12 > 4.375 & \Rightarrow 1 & 4.1 > 2.525 & \Rightarrow 1 \end{bmatrix}$$

   Final matrix:

$$X_{\text{binary}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

3. Given the following enumerated features, create a new *one-hot-encoded* set of features. You may assume that the set of values you see for each feature constitute the entire unique set of enumerations (2pts).

$$X = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Answer for Q2:

Unique values: $\{0, 1, 2\}$

One-hot encoding:

$$\begin{bmatrix} 2 \Rightarrow [0\ 0\ 1] \\ 1 \Rightarrow [0\ 1\ 0] \\ 0 \Rightarrow [1\ 0\ 0] \\ 0 \Rightarrow [1\ 0\ 0] \\ 1 \Rightarrow [0\ 1\ 0] \end{bmatrix}$$

Final matrix:

$$X_{\text{onehot}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

4. Consider the following dataset consisting of 10 observations, each with two features:

$$X = \begin{bmatrix} 0 & -2 \\ 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$$

(a) What are the principal components of the observed data X for use in PCA? (5pts)
All values in the first column are zero, so there is no variance in the first feature. The variance lies entirely in the second feature.

$$\text{First principal component (direction of max variance): } [0\ 1]^T$$

$$\text{Second principal component (orthogonal to the first): } [1\ 0]^T$$

So the principal components are:

$$\text{PC}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \text{PC}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(b) What are the angles of the principal components relative to the original coordinate system? (3pts)

We calculate the angle θ between the first principal component and the x-axis.

$$\theta = \tan^{-1}\left(\frac{1}{0}\right) = 90^\circ$$

So the first principal component is rotated 90° from the x-axis.

(c) If we were to project our data down to 1-D using the principal component, what would the new data matrix X be? (5pts)

We project onto $\text{PC}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$:

$$X \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$

Final projected 1-D data:

$$X_{\text{1D}} = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$