MTH793P **Advanced Machine Learning, Semester B, 2023/24**
Coursework 7

Principal Component Analysis (PCA)

1. Consider the set of points in the plane:

$$x_1 = (4, 3)^T,\ x_2 = (-4, 6)^T,\ x_3 = (7, -2)^T,\ x_4 = (1, 1)^T,\ x_5 = (0, -2)^T$$

   (a) Set up a corresponding data matrix $X \in \mathbb{R}^{2\times 5}$.

   (b) Find the principal components of X (remember to center the data points first).

   (c) Compute the projections $y_1, \ldots, y_5 \in \mathbb{R}^1$ of $x_1, \ldots, x_5 \in \mathbb{R}^2$ on the first principal component (remember to correct for the mean).

   (d) Compute the reconstructions of $x_1, \ldots, x_5$ using the first principal components, denoted $\hat{x}_1, \ldots, \hat{x}_5 \in \mathbb{R}^2$.

   (e) In a 2D axis system, plot the following:
      - The original points $x_1, \ldots, x_5$.
      - The reconstruced points $\hat{x}_1, \ldots, \hat{x}_5$.
      - The principal components (directions).

2. Consider the following points in $\mathbb{R}^5$:

$$x_1 = (2, 3, 1, 0, -1)^T, x_2 = (-4, 2, 4, 4, 1)^T, x_3 = (-1, -1, 3, 9, 0)^T.$$

   Find the projections $y_1, y_2, y_3 \in \mathbb{R}^2$ of these points, that makes them uncorrelated and with maximal variance (remember to center the points first).