Week 8 Exercise Sheet

The following exercises have different levels of difficulty indicated by (*), (**), (***). An exercise with (*) is a simple exercise requiring less time or effort to solve compared to an exercise with (***), which is a more complex exercise.

Unsupervised Learning

1. (*) We want to use PCA to reduce dimensionality from 3 to 2. The covariance matrix of the data is

$$\mathbf{C} = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \tag{1}$$

   and the corresponding eigenvectors are

$$\mathbf{w}_1 = \begin{pmatrix} -0.872 \\ 0.466 \\ 0.152 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} -0.390 \\ -0.847 \\ 0.361 \end{pmatrix}, \mathbf{w}_3 = \begin{pmatrix} 0.297 \\ -0.256 \\ 0.920 \end{pmatrix}. \tag{2}$$

   Using the eigenvalue equation $\lambda_i = \mathbf{w}_i \mathbf{C} \mathbf{w}_i$, show that the eigenvalues are $\lambda_1 = 5.070, \lambda_2 = -0.346, \lambda_3 = 2.278$ (to 3 decimal places).

2. (**) Following from 1., which 2 eigenvectors should be used when applying PCA to reduce the dimensionality to 2? If we have 2 datapoints $\mathbf{x}_1 = (2, 3, 3)^T$ and $\mathbf{x}_2 = (4, 1, 0)^T$. Apply the PCA transformation to calculate the transformed datapoints. Show your steps and you assume that the datapoints have already had the mean subtracted.

3. (***) An alternative to derive PCA is to minimise the reconstruction error. Consider the first principal component $\mathbf{u}_1$ such that a transformed data point is $y_{n1} = \mathbf{u}_1^T \mathbf{x}_n$ and the reconstructed data point is $\tilde{\mathbf{x}}_n = \mathbf{u}_1 y_{n1}$. Show that the reconstruction error

$$E = \frac{1}{2N} \sum_{n=1}^{N} |\mathbf{x}_n - \tilde{\mathbf{x}}_n|^2 \tag{3}$$

is equal to

$$E = -\mathbf{u}_1^T \mathbf{C} \mathbf{u}_1 + \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_n^T \mathbf{x}_n, \tag{4}$$

where $\mathbf{C} = \sum_{n=1}^{N} \mathbf{x}\mathbf{x}^T/N$ is the covariance matrix. You will need to use the definition that the square of a vector $\mathbf{a}$ is the inner product $|\mathbf{a}|^2 = \mathbf{a}^T\mathbf{a}$ and that the principal component is normalised $\mathbf{u}_1^T\mathbf{u}_1 =$.