WTW 801 Assignment 2

Due date: Saturday 30 October, 23:30. Upload it to the ClickUP link that will be provided. Total: 20 marks.

Max=20 marks if submitted on time; 17 marks if 3 days late; else 0 marks.

1. (6 marks) Generate 1200 data points for the variables x, y, z with the Python (or otherwise, if you prefer another computer language) commands

   ```
   import numpy as np
   sample=np.random.multivariate_normal...
   ...(np.array([2,3,0]),np.array([[10,7,5],[7,6,4],[5,4,3]]), 1000).T
   xdata=sample[0,:]
   ydata=sample[1,:]
   zdata=sample[2,:]
   ```

   (The ellipsis "..." indicates that the line continues. It should not be entered.)

   Do a PCA and write down the three components and the corresponding principal values. (If they are not vectors of length 3, you are doing something wrong.) Check that the pointwise products of the components are zero (up to machine precision). Write down the dot (i.e. pointwise) product of the first and third components, and compare that with the dot product of xdata and zdata. Show your steps and append your code to the assignment.

2. (1 mark) Recall the dot product between vectors $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$:

   $$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n.$$

   What is the difference, if any, between $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z})$ and $\mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$? Motivate your answer.

3. (1 mark) Assume we have mutually orthogonal vectors $\mathbf{f}_1, \cdots, \mathbf{f}_n$ and a representation

   $$\mathbf{w} = a_1 \mathbf{f}_1 + \cdots + a_m \mathbf{f}_m,$$

   where $a_1, \ldots, a_M$ are scalars. Prove that

   $$a_i = \frac{\mathbf{w} \cdot \mathbf{f}_i}{\mathbf{f}_i \cdot \mathbf{f}_i}$$

   for every $i = 1, \ldots, m$. (Hint: calculate $\mathbf{w} \cdot \mathbf{f}_i$ and simplify.)

4. (1 mark) Suppose that such a representation is not possible but that we instead have

   $$\mathbf{w}_i = a_i + \beta_{i1} \mathbf{f}_1 + \cdots + \beta_{ik} \mathbf{f}_k + \mathbf{g}_i$$

   where $a_i, \beta_{i1}, \ldots, \beta_{ik}$ are scalars and $\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{g}_i$ are mutually orthogonal vectors in $\mathbb{R}^n$ and have mean zero. What is the relationship between the mean of vector $\mathbf{w}_i$ and a_i? Prove your statement.

5. (9+1 marks) Download 5 years historical daily price data for the Johannesburg Stock Exchange index, four large companies of your choice on the JSE, as well as another financial variable of your choice. Convert the prices to daily *returns*.

   - Build a factor model for these returns. Give both the model with the factor loadings and the components (factors) that you include in the model.
   - Which are the main financial assets that contribute most to the first principal component (factor)?
   - Which percentage of total variance is explained by the first principal component (factor)?
   - Can you interpret the first principal component? (1 mark)

   Share price data can be downloaded at among others the link `https://uk.finance.yahoo.com/`

   ZAR exchange rate data can be downloaded among others at the link `https://www.resbank.co.za/Research/Rates/`