WTW 801 Assignment 2

Due date: Thursday 18 October, 24:00; email to gusti.vanzyl@up.ac.za Total: 20 marks

1. (7 marks) Generate 1000 data points for the variables x, y, z with the Python 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 nto be entered.)

   Do a PCA and write down the first and second components. (If they are not vectors of length 3, you are doing something wrong.) Check that the components are orthogonal. 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 $\mathbf{y} = (y_1, \ldots, y_n)$:

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

   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.)

3. (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_0? Prove your statement.

4. (1 mark) In PCA a variables $\times$ observations matrix X is transformed to a new variables $\times$ observation matrix Y, by the equation $Y = PX$, where P is a certain matrix for which the columns are orthogonal. Show that this transformation *preserves angles*, i.e. show that $(P\mathbf{x}_1) \cdot (P\mathbf{x}_2) = \mathbf{x}_1 \cdot \mathbf{x}_2$ for any vectors $\mathbf{x}_1, \mathbf{x}_2$ to which P can be applied. In particular, if variables are orthogonal then their transformed counterparts are also orthogonal. Hint: $P^T P = I$.

5. (10 marks) Download 5 years historical daily price data for the Johannesburg Stock Exchange index, three large companies of your choice on the JSE, as well as another factor or two such as relevant exchange rates. 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)?

   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/`