# ENGN/COMP8535 Homework 2

Yuyang Han u7434576

## Q1

(a) P is

```
[[0.         1.         0.         0.        ]
 [0.33333333 0.         0.33333333 0.33333333]
 [0.         0.         0.         1.        ]
 [1.         0.         0.         0.        ]]
```

(b) G is

```
[[0.0375     0.8875     0.0375     0.0375    ]
 [0.32083333 0.0375     0.32083333 0.32083333]
 [0.0375     0.0375     0.0375     0.8875    ]
 [0.8875     0.0375     0.0375     0.0375    ]]
```

(c) Use G and $\alpha$, rank the webpages with fullrank algorithm. Using the power method to compute an approximation of $\pi_\infty$ with 5 iterations from $\pi_0 = [0.25, 0.25, 0.25, 0.25]^T$ gets the following rank results.

```
[[0.32885526]
 [0.31353484]
 [0.11935062]
 [0.23825928]]
```

where the first page has the highest rank.

Q2 (a)

Here is the result of find the svd of X using python numpy library

```
X = np.array([ [3, 2, 2],
               [2, 3, -2] ])

U, s, Vt = np.linalg.svd(X)

print('U:\n', U)

print('S:\n', np.diag(s))

print('V:\n', Vt.T)
```

✓ 0.3s

```
U:
 [[-0.70710678 -0.70710678]
 [-0.70710678  0.70710678]]
S:
 [[5. 0.]
 [0. 3.]]
V:
 [[-7.07106781e-01 -2.35702260e-01 -6.66666667e-01]
 [-7.07106781e-01  2.35702260e-01  6.66666667e-01]
 [-6.47932334e-17 -9.42809042e-01  3.33333333e-01]]
```

continue...

(a) Working out the svd of X by hand

Find SVD of $X = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$

$$X^T = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

the eigenvales and eigenvectors of $X^T X$ can be found by setting the $\det(X^T X - \lambda I) = 0$

where we have $\det(X^T X - \lambda I) = -\lambda(\lambda - 25)(\lambda - 9) = 0$

the eigenvalues of $X^T X$ are $\lambda_1 = 25$ $\lambda_2 = 9$ $\lambda_3 = 0$

Then we can find the eigenvectors $x_1$ $x_2$ $x_3$

- $\lambda_1 = 25$

$(X^T X - 25 I) = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix}$ reduce to row-echlon form is

$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ we have $x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

- $\lambda_2 = 9$

$(X^T X - 9I) x_2 = 0$ $\begin{bmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{bmatrix}$ reduce to row-echlon form is

$\begin{bmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{bmatrix} x_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ solve for $x_2$ we have $x_2 = \begin{bmatrix} 1/4 \\ -1/4 \\ 1 \end{bmatrix}$

- $\lambda_3 = 0$

$(X^T X) x_3 = 0$ $\begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$ reduce to row-echlon form is

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} X_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

solve for $x_3$ we have $X_3 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$

the singular values are the square roots of the eigenvalues

$$\Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

columns of $V$ are the normalized eigenvectors

therefore $V_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix}$ $V_2 = \begin{bmatrix} \sqrt{2}/6 \\ -\sqrt{2}/6 \\ 2\sqrt{2}/3 \end{bmatrix}$ $V_3 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$

$$V = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/6 & -2/3 \\ \sqrt{2}/2 & -\sqrt{2}/6 & 2/3 \\ 0 & 2\sqrt{2}/3 & 1/3 \end{bmatrix}$$

$$u_i = \frac{1}{\sigma_i} X v_i$$

So

$$u_1 = \frac{1}{5}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$$

$$u_2 = \frac{1}{3}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}\begin{bmatrix} \sqrt{2}/6 \\ -\sqrt{2}/6 \\ 2\sqrt{2}/3 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$$

Thus

$$U = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \quad V = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/6 & -2/3 \\ \sqrt{2}/2 & -\sqrt{2}/6 & 2/3 \\ 0 & 2\sqrt{2}/3 & 1/3 \end{bmatrix}$$

such that $X = U \Sigma V^T$

$\blacksquare$

(b) The best rank-one approximation to $X$ is $\sigma_1 u_1 v_1^T$ where $\sigma_1$ is the first singular value, $u_1$ is the first left singular vector, and $v_1^T$ is the first right singular vector of $X$. Below is the reconstructed $X$. using python.

```
[[2.50000000e+00 2.50000000e+00 2.29078674e-16]
 [2.50000000e+00 2.50000000e+00 2.29078674e-16]]
```

The best rank-1 appoximation to $X$ is $\sigma_1 u_1 v_1^T$ where $\sigma_1$ is the first singular value. $u_1$ is the first left singular vector. $v_1$ is the first right singular vector

Hence $X_1$ is given by $\sigma_1 u_1 v_1^T$

$\sigma_1 = 5 \qquad u_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \qquad v_1^T = \left[ \sqrt{2}/2 , \sqrt{2}/2 , 0 \right]$

$$X_1 = \begin{bmatrix} 2.5 & 2.5 & 0 \\ 2.5 & 2.5 & 0 \end{bmatrix}$$

$\blacksquare$

(c)

$$X - X_1 = \begin{bmatrix} 0.5 & -0.5 & 2 \\ -0.5 & 0.5 & -2 \end{bmatrix}$$

$$\|X - X_1\| = \sqrt{0.5^2 + 0.5^2 + 0.5^2 + 0.5^2 + 2^2 + 2^2} = \sqrt{9} = 3$$

therefore ehe approximation error under the Frobenius Norm is 3

$\blacksquare$

Q3. $x_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ $x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $x_3 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ $x_4 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$ $x_5 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

(a)

$$\mu = \frac{1}{5}\begin{bmatrix} 4+0+(-2)+4+2 \\ 3+1+2+(-1)+0 \end{bmatrix} = [8/5, 1]$$

$$\Delta x_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix} - \begin{bmatrix} 1.6 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.4 \\ 2 \end{bmatrix} \quad \Delta x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1.6 \\ 1 \end{bmatrix} = \begin{bmatrix} -1.6 \\ 0 \end{bmatrix} \quad \Delta x_3 = \begin{bmatrix} -2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1.6 \\ 1 \end{bmatrix} = \begin{bmatrix} -3.6 \\ 1 \end{bmatrix}$$

$$\Delta x_4 = \begin{bmatrix} 4 \\ -1 \end{bmatrix} - \begin{bmatrix} 1.6 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.4 \\ -2 \end{bmatrix} \quad \Delta x_5 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1.6 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.4 \\ -1 \end{bmatrix}$$

we form $X = [\Delta x_1, \Delta x_2, \Delta x_3, \Delta x_4, \Delta x_5] \in R^{d \times n}$

where $d = 2$ $n = 5$

then covariance matrix is given by $C = \frac{1}{n} X X^T$

$$C = \frac{1}{5}\begin{bmatrix} 2.4 & -1.6 & -3.6 & 2.4 & 0.4 \\ 2 & 0 & 1 & -2 & -1 \end{bmatrix}\begin{bmatrix} 2.4 & 2 \\ -1.6 & 0 \\ -3.6 & 1 \\ 2.4 & -2 \\ 0.4 & -1 \end{bmatrix}$$

$$= \frac{1}{5}\begin{bmatrix} \frac{136}{5} & -4 \\ -4 & 10 \end{bmatrix}$$

the eigenvalues and eigenvectors of C are

eigenvalues : $[7.0212 \quad 2.2788]$

eigenvectors : $\begin{bmatrix} -0.9764 \\ 0.2159 \end{bmatrix}$ $\begin{bmatrix} 0.2159 \\ 0.9764 \end{bmatrix}$

Then the first principal component is

$$u_1 = \begin{bmatrix} -0.9764 \\ 0.2159 \end{bmatrix}$$ since it is the eigenvector

corresponding to largest eigenvalue $\lambda_1 = 7.0212$

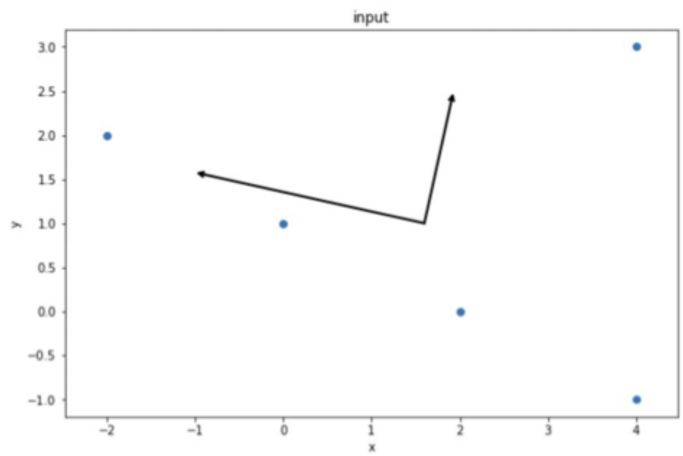
```
pca = PCA(n_components=2)
pca.fit(X)
```
✓ 0.2s

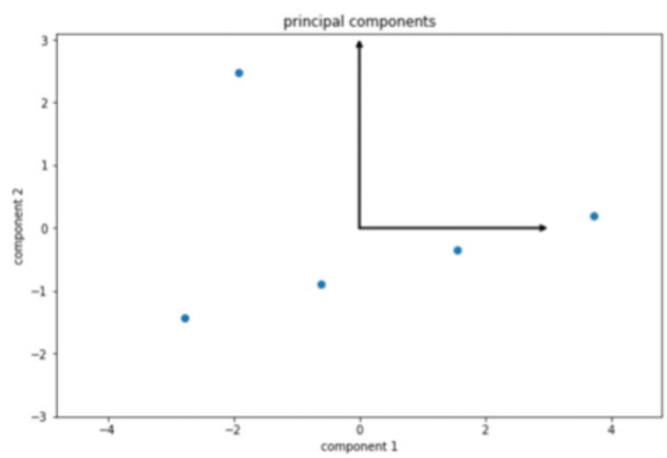
```
PCA(n_components=2)
```

```
print(pca.components_)

print(pca.explained_variance_)
```
✓ 0.2s

```
[[-0.97640184  0.21596167]
 [ 0.21596167  0.97640184]]
[7.02118114 2.27881886]
```



(b) Use PCA to compute lower-dimensional

$$y_j = u_1^T x_j = \begin{bmatrix} -0.9764 & 0.2159 \end{bmatrix} \Delta x_j$$

$y_1 = -1.9114$ $y_2 = 1.5622$ $y_3 = 3.731$

$y_4 = -2.7753$ $y_5 = -0.6065$

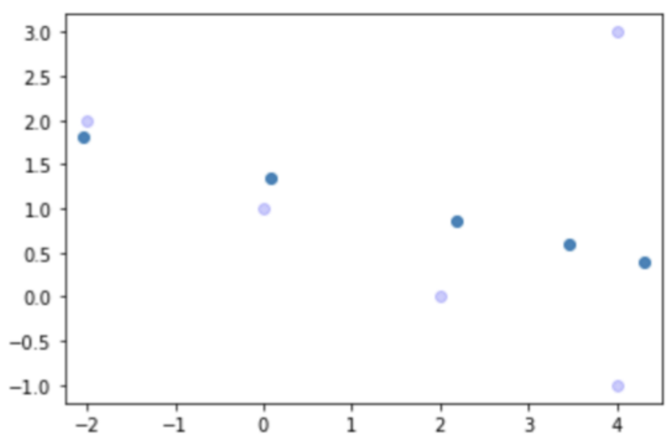
(b)

```python
pca = PCA(n_components=1)
pca.fit(X)
X_1d = pca.transform(X)
print("original shape:   ", X.shape)
print("transformed shape:", X_1d.shape)

print(X_1d)
```

✓ 0.2s

```
original shape:    (5, 2)
transformed shape: (5, 1)
[[-1.91144107]
 [ 1.56224295]
 [ 3.7310083 ]
 [-2.77528776]
 [-0.60652241]]
```



Q4 Given real symmetric $C \in R^{d \times d}$, show

largest eigenvalue $\lambda_1 = \max_{\{u \in R^d : \|u\|_2 = 1\}} u^T C u$ and

$u_1 \in R^d$ corresponding to $\lambda_1$ is

$$u_1 = \underset{\{u \in R^d : \|u\|_2 = 1\}}{\operatorname{argmax}} u^T C u.$$

a) We define a transformed vector $y = Q^T u$

so $u^T C u = y^T \Lambda y$, expanding it we obtain

$y^T \Lambda y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_d y_d^2$

Since $\lambda_1$ is the largest eigenvalue, therefore $\lambda_1 \geq \lambda_i$ for $i = 1, 2 \cdots d$.

therefore we have $y^T \Lambda y \leq \underbrace{\lambda_1 y_1^2 + \lambda_1 y_2^2 + \lambda_1 y_3^2 + \cdots + \lambda_1 y_d^2}$

$$y^T \Lambda y \leq \lambda_1 y^T y$$

Since in the decomposition $Q^{-1} = Q$, $QQ^T = I$

we have $y^T y = u^T Q Q^T u = u^T u$

Hence $y^T \Lambda y \leq \lambda_1 y^T y$, $u^T C u \leq \lambda_1 u^T u$

By definition $\|u\|_2 = 1$ therefore $u^T C u \leq \lambda_1$

$$\lambda_1 = \max_{\{u \in R^d : \|u\|_2 = 1\}} u^T C u$$

b) Real symmetric matrix $C$ can be decomposed as

$C = Q \Lambda Q^T$, where $\Lambda$ is $\operatorname{diag}(\lambda_1, \lambda_2 \ldots \lambda_d)$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_d$ are the eigenvalues of $C$

and $u_1, u_2 \cdots u_d$ are corresponding orthonormal eigenvectors.

Let $u \in R^d$ be a unit vector, $\|u\|_2 = 1$ we can write $u$

as a linear combination of the eigenvectors of $C$

$u = c_1 u_1 + c_2 u_2 + \cdots c_d u_d$ where $c_1, c_2, \cdots\cdots c_d$ are constants.

$$u^T C u = (c_1 u_1 + c_2 u_2 + \cdots\cdots + c_d u_d)^T C (c_1 u_1 + c_2 u_2 + \cdots\cdots + c_d u_d)$$

$$= \underbrace{c_1^2 \lambda_1 + c_2^2 \lambda_2 + \cdots\cdots + c_d^2 \lambda_d}_{(1)}$$

$\|u\|_2 = 1$

$$\underbrace{1 = \|u\|_2 = \sqrt{(c_1^2 + c_2^2 + \cdots + c_d^2)}}_{(2)}$$

therefore maximizing $u^T C u$ is equivalent to maximizing (1)

with respect to (2)

using Lagrange multiplier

Let $L(c, \lambda) = c^T \Lambda c - \lambda (c^T \Lambda c - 1)$

taking $\frac{\partial L}{\partial c} = 0$, we obtain $2\Lambda c - 2\lambda c = 0$

implying $\Lambda c = \lambda c$

this shows $c = \arg\max (1)$ is an eigenvector of $C$

and the max value is $\lambda_1$.

Therefore $\lambda_1 = \max u^T C u$ and it is attained

when $u$ is the eigenvector.

∎

Q5.

$$X = U\Sigma V^T \qquad C = \frac{1}{n} X X^T$$

since datapoints in $X$ has zero mean $\mu = 0$

we have $C = \frac{1}{n}(X-\mu)(X-\mu)^T = \frac{1}{n} X X^T$

columns of $V$ are the eigenvectors of covariance matrix $C$

digonal of $\Sigma$ are the square roots of corresponding eigenvalues.

The first k columns of $V$ gives first k principal components of the data.

The k-dimensional data can be obtained by projecting it onto k principal components

$$y = V_k^T X$$

$\boxed{/\!/\!/}$