EEL5840 Elements of Machine Intelligence - HW 1

Madhusudan Govindraju 39267182

1

The figure 1 shows the solution for Question 1

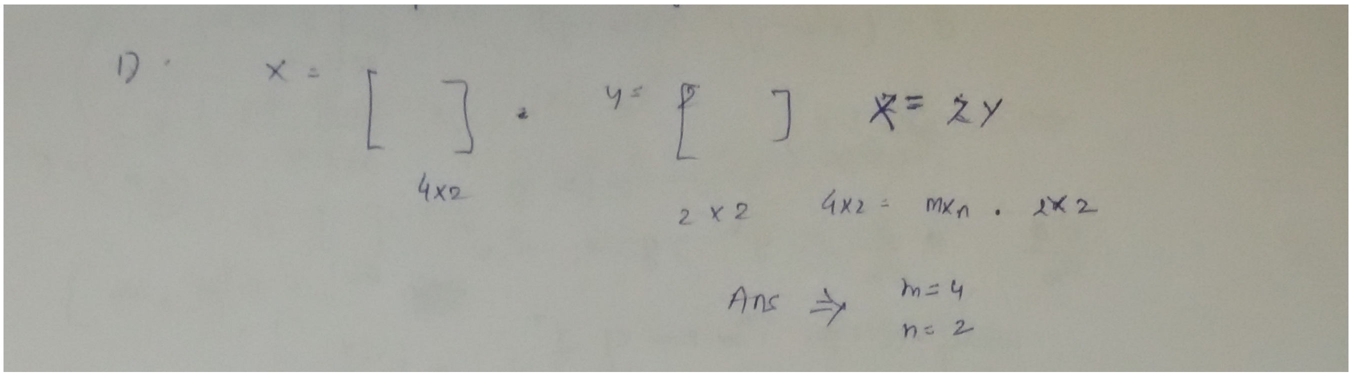
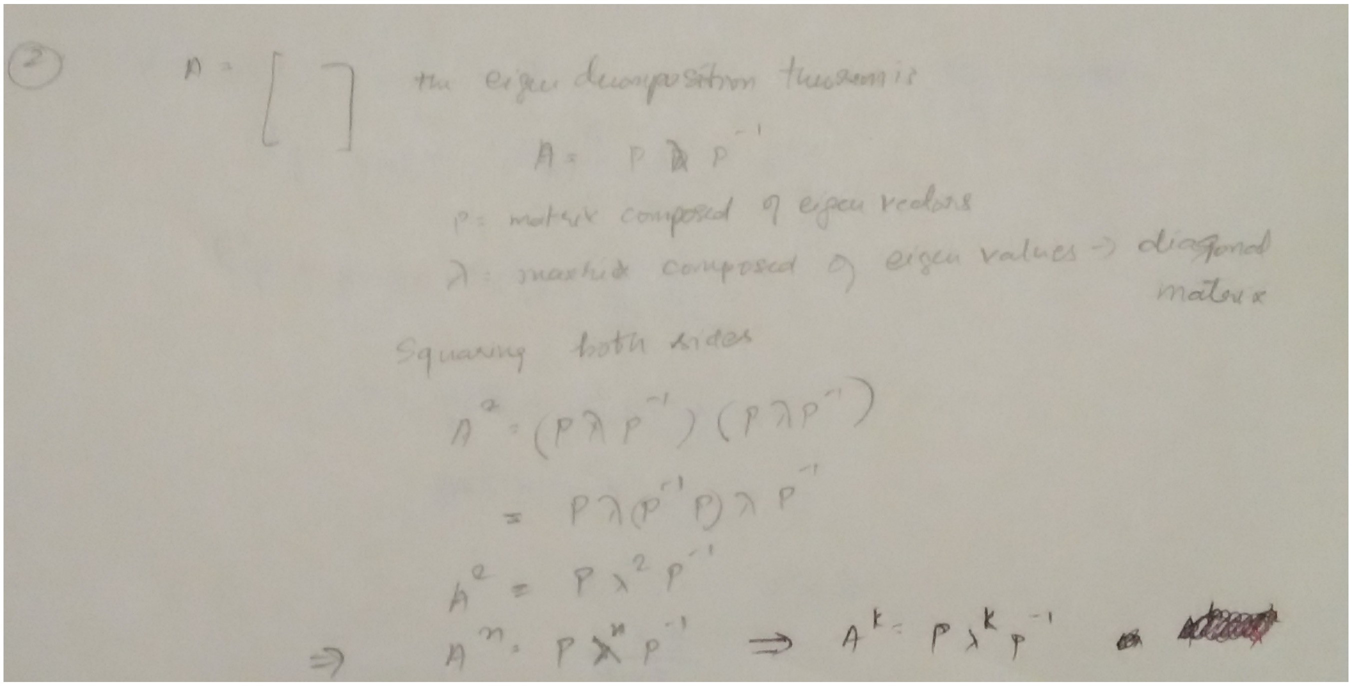


Figure 1: Q 1

2

The figure 2 shows the solution for Question 2



(2) $A = [\ \ ]$ the eigen decomposition theorem is

$$A = P \lambda P^{-1}$$

P = matrix composed of eigen vectors

λ = matrix composed of eigen values → diagonal matrix

Squaring both sides

$$A^2 = (P\lambda P^{-1})(P\lambda P^{-1})$$

$$= P\lambda(P^{-1}P)\lambda P^{-1}$$

$$A^2 = P\lambda^2 P^{-1}$$

$$\Rightarrow A^n = P\lambda^n P^{-1} \Rightarrow A^k = P\lambda^k P^{-1}$$

Figure 2: Q 2

3

The figure 3 shows the solution for Question 3. Using that equation we solve for the given matrices to get the following solution

```
>> X = [ -3 1 1 2 0; 2 0 1 4 1; 1 0 0 1 -1];
>> y = [1; -1; 1; 2; 10];
>> w = ((2.* X)* X')^(-1) * ((2.* X)* y)
w =
   -0.9552
    2.7164
   -7.1791
```

3) $\phi = (x^T w - y)^T (x^T w - y)$

this is similar to $\phi = x^T x$ for which

$$\frac{\partial \phi}{\partial w} = 2x^T \frac{\partial x}{\partial w}$$

where x is $n \times 1$ and x is a function of w

So applying it in our equation we have

$$\frac{\partial \phi}{\partial w} = 2(x^T w - y)^T \frac{\partial (x^T w - y)}{\partial w}$$

$$= 2(w^T x - y^T)\ x^T$$

$$= 2w^T x x^T - 2y^T x^T$$

LHS = 0 to solve the equation

$2w^T x x^T = 2y^T x^T \Rightarrow$ applying transpose to all terms we get

$$\Rightarrow 2xx^T w = 2xy$$

$$w = (2xx^T)^{-1} \cdot (2xy)$$

Figure 3: Q 3

4

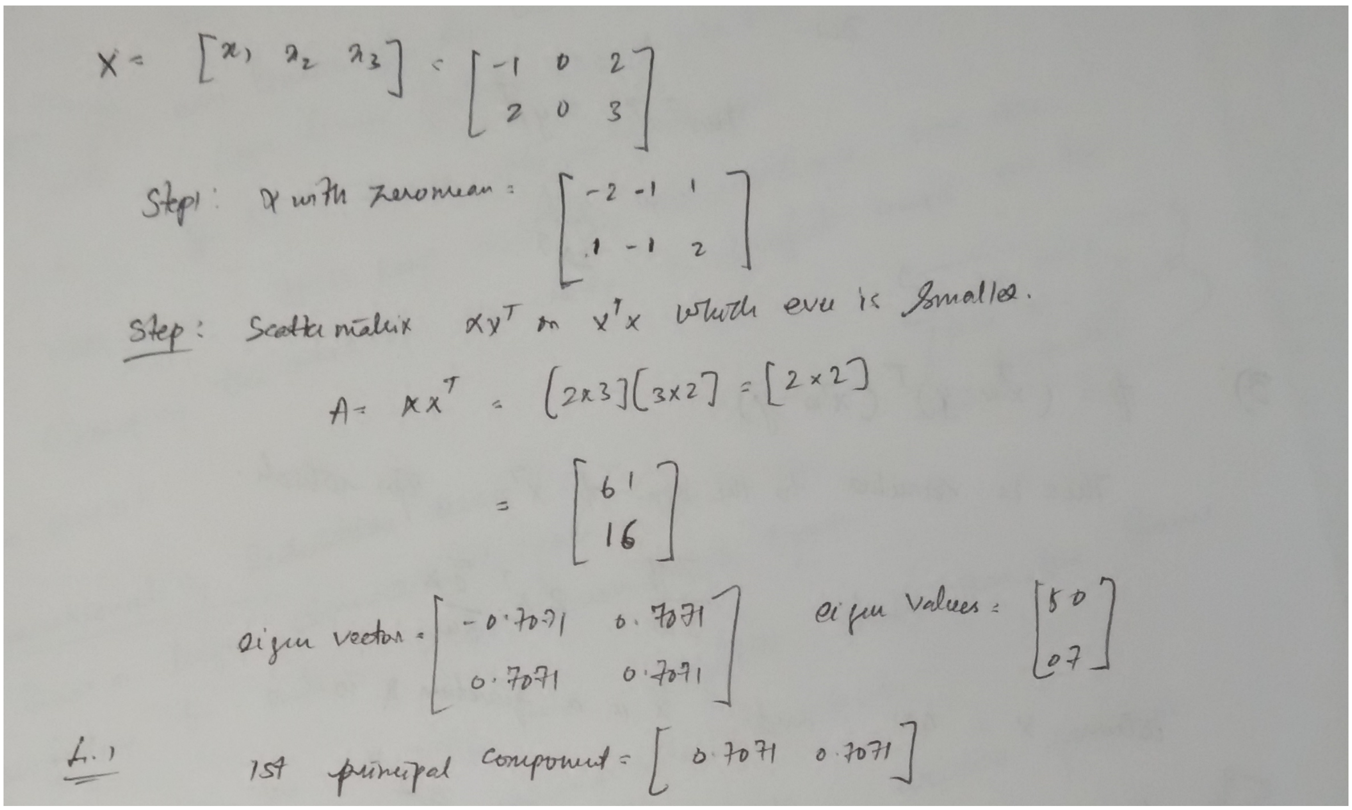
The figures 4 5 6 shows the solution for Question 4. The principal component can be calculated by

```
>>X = [-1 2 ; 0 0; 2 3]
>>pca(X)

ans =

   -0.7071    0.7071
    0.7071    0.7071
```

The column corresponding to the largest eigen value is the 1st principal component. In our case the 1st Principal component is the vector [0.7071 0.7071]
From the high error calculated in fig 6 we can see that the PCA is not that suitable for this dataset.



$X = [x_1 \; x_2 \; x_3] = \begin{bmatrix} -1 & 0 & 2 \\ 2 & 0 & 3 \end{bmatrix}$

Step1: X with zeromean = $\begin{bmatrix} -2 & -1 & 1 \\ 1 & -1 & 2 \end{bmatrix}$

Step: Scatter matrix XX^T or X^TX which ever is smaller.

$A = XX^T = (2\times3)(3\times2) = [2\times2]$

$= \begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}$

eigen vector = $\begin{bmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}$ eigen values = $\begin{bmatrix} 5 & 0 \\ 0 & 7 \end{bmatrix}$

4.1 1st principal component = $[\, 0.7071 \;\; 0.7071 \,]$

Figure 4: Q 4.1

The error in fig 6 is calculated with euclidean distance btw the two points. The matrix is reconstructed into the original 2D subspace with the principal component used to reduce the subspace. The new points are Xbar and the original points are X the distance between the two gives the reconstruction error for the point.

4.2 after projecting it into 1D Subspace.

$$x_1 = [-1\ 2]\begin{bmatrix}0.7071\\0.7071\end{bmatrix} = [0.7071]$$

$$x_2 = [0, 0]\begin{bmatrix}0.7071\\0.7071\end{bmatrix} = [0]$$

$$x_3 = [2, 3]\begin{bmatrix}0.7071\\0.7071\end{bmatrix} = [3.53]$$

with zero mean the new projection in 1D Subspace is

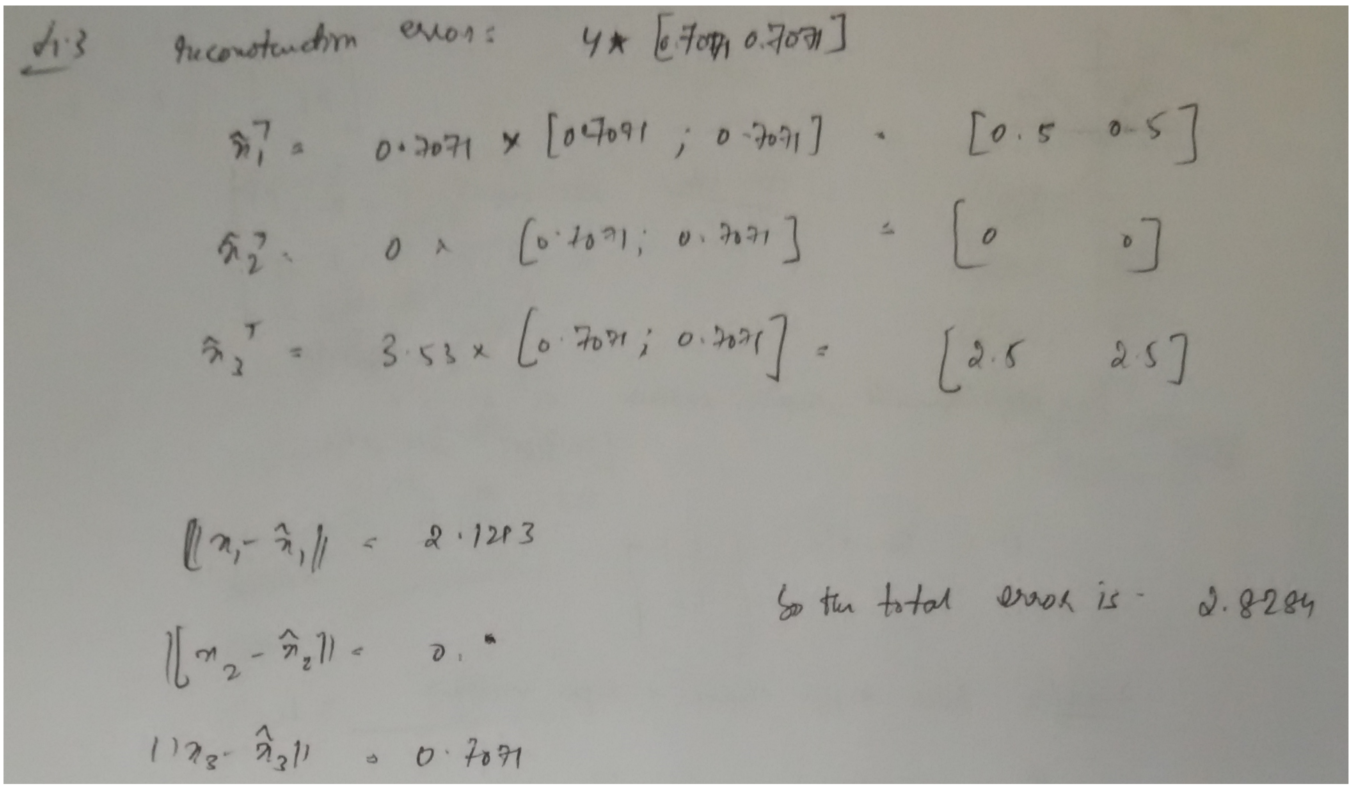
$$\begin{bmatrix}-0.7071\\-1.4124\\2.1176\end{bmatrix}$$

the variance $= \frac{1}{n-1}\sum_{i=1}^{3}\left[Z_i - \text{mean}\right]^2$

$$= \frac{1}{3-1}\left[(-0.7071)^2 + (-1.4124)^2 + (2.1176)^2\right]$$

the variance $= 3.49$

Figure 5: Q 4.2



4.3 reconstruction error: 4 ★ [0.7071 0.7071]

$\hat{x}_1^T = 0.7071 \times [0.7071 ; 0.7071] = [0.5 \quad 0.5]$

$\hat{x}_2^T = 0 \times [0.7071 ; 0.7071] = [0 \quad 0]$

$\hat{x}_3^T = 3.53 \times [0.7071 ; 0.7071] = [2.5 \quad 2.5]$

$\|x_1 - \hat{x}_1\| = 2.1213$

$\|x_2 - \hat{x}_2\| = 0.$

$\|x_3 - \hat{x}_3\| = 0.7071$

So the total error is 2.8284

Figure 6: Q 4.3