HW4

Juan Vila

November 12, 2019

Question 1

Part a

For finding the projection the P onto V, we need to create a matrix that represent the basis. For doing this we $x_1 = x_2 + 2x_3$, this implies that $A = \begin{bmatrix} x_2 + 2x_3 & x_2 + 2x_3 \\ x_2 & x_2 \\ x_3 & x_3 \end{bmatrix}$, then for v_1 we are going to assume $x_3 = 0$ and $x_2 = 1$ and for v_2 we are going to assume that $x_2 = 0$ and $x_3 = 1$ then A:

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

With this A, we are going to calculate the proyection matrix as:

$$P = A(A^T A)^{-1} A^T$$

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \left[\begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \right]^{-1} \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T$$

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T$$

We know that $\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix}$, then:

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{6} \begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T$$

$$P = \frac{1}{6}\begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 5 & -2 \\ 2 & -2 & 2 \end{bmatrix}$$

$$P = \frac{1}{6}\begin{bmatrix} 5 & 1 & 2 \\ 1 & 5 & -2 \\ 2 & -2 & 2 \end{bmatrix}$$

Part b

The rank of the matrix is two, because we can see than $\frac{1}{6}\left(v_1\frac{1}{2} - \frac{1}{2}v_2\right) = \frac{v_3}{6}$, this implies that $v_3$is LD, meaning that the rank is 2.

Part c

now for calculating the distance we apply the following formula:

$$\left\|\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - P\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\|_2$$

$$\left\|\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 5 & 1 & 2 \\ 1 & 5 & -2 \\ 2 & -2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\|_2$$

$$\left\|\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 8/6 \\ 4/6 \\ 2/6 \end{bmatrix}\right\|_2$$

$$\left\|\begin{bmatrix} -1/3 \\ 1/3 \\ 2/3 \end{bmatrix}\right\|_2 = \sqrt{\frac{1}{9} + \frac{1}{9} + \frac{4}{9}} = \sqrt{\frac{2}{3}}$$

Question 2

Part a

The easiest orthonormal basis for X is $[e_1, e_2]$, because as $x_1 = 3e_1$, $x_2 = 3e_1 + 4e_2$ and $x_3 = \sqrt{2}e_2$

Part b

b.i

The projection matrix P onto V will be :

$$P = V(V^T V)^{-1} V^T$$

b.ii

The add up of the distances squared will be for any v is:

$$\|v - Pv\|_2^2 = [v - Pv]^T [v - Pv]$$

$$\|v - Pv\|_2^2 = v^T [I - P]^T [I - P] v$$

we know that $P^T + P = I$ and P^TP=P, then can replace into the equation:

$$\|v - Pv\|_2^2 = v^T \left[I - I + P^T\right]^T \left[I - I + P^T\right] v$$

$$\|v - Pv\|_2^2 = v^T P v$$

Now we replace v for x and we have the same the distance square for every x_i

$$\|x - Px\|_2^2 = \sum_{i=1}^{3} x_i^T P x_i$$

b.iii

$$\begin{bmatrix} 3 & 0 \end{bmatrix} P \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 & 4 \end{bmatrix} P \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 & \sqrt{2} \end{bmatrix} P \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 \end{bmatrix} V(V^T V)^{-1} V^T \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 & 4 \end{bmatrix} V(V^T V)^{-1} V^T \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 & \sqrt{2} \end{bmatrix} V(V^T V)^{-1} V^T \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$$

As we know that $\|v\|_2 = 1$,implies that $V^T V = 1$

$$\begin{bmatrix} 3 & 0 \end{bmatrix} VV^T \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 & 4 \end{bmatrix} VV^T \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 & \sqrt{2} \end{bmatrix} VV^T \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$$

Also, we can convert V into $V = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, and $VV^T = \begin{bmatrix} v_1^2 & v_1v_2 \\ v_1v_2 & v_2^2 \end{bmatrix}$, then:

$$\begin{bmatrix} 3 & 0 \end{bmatrix} \begin{bmatrix} v_1^2 & v_1v_2 \\ v_1v_2 & v_2^2 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} v_1^2 & v_1v_2 \\ v_1v_2 & v_2^2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} v_1^2 & v_1v_2 \\ v_1v_2 & v_2^2 \end{bmatrix} \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$$

$$9v_1^2 + 9v_1^2 + 24v_1v_2 + 16v_2^2 + 2v_2^2$$

$$18v_1^2 + 24v_1v_2 + 18v_2^2$$

We know that $v_1^2 + v_2^2 = 1$ and $v_2^2 = \sqrt{1 - v_1^2}$we replace this fact in the equation and take the gradient of this result:

$$18 + 24v_1\sqrt{1 - v_1^2}$$

$$\{v_1\} : -\frac{2 * 24v_1^2}{2\sqrt{1 - v_1^2}} + 24\sqrt{1 - v_1^2} = 0$$

$$\frac{v_1^2}{\sqrt{1 - v_1^2}} = \sqrt{1 - v_1^2}$$

$$2v_1^2 = 1$$

$$v_1 = \frac{1}{\sqrt{2}} \rightarrow v_2 = \frac{1}{\sqrt{2}}$$

Part c

For building matrix U, we use the result from the previous part we know that $U_1 = [v_1, v_2]^T$, for obtain $U_2 = [a, b]^T$, we know that $U_1^T U_2 = 0$and $U_2^T U_2 = 1$, then we need to solve the following system:

$$\begin{gathered} v_1 a + v_2 b = 0 \\ a^2 + b^2 = 1 \end{gathered}$$

Then from eq 1 we know that $b = -a$, and replacing into the second one we get that $a = \frac{1}{\sqrt{2}} \rightarrow b = -\frac{1}{\sqrt{2}}$, then:

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Now for getting the parameters of Σ we need to calculate the sum of the residuals of projection of U_i onto to X. For doing this we follow the folowing formula: $\sigma_j = \sqrt{\sum_{i=1}^{3} \left(x_i^T U_j\right)^2}$, then:

$$\sigma_1 = \sqrt{\left(\begin{bmatrix} 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right)^2 + \left(\begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right)^2 + \left(\begin{bmatrix} 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right)^2} = \sqrt{30}$$

$$\sigma_2 = \sqrt{\left(\begin{bmatrix} 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}\right)^2 + \left(\begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}\right)^2 + \left(\begin{bmatrix} 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}\right)^2} = \sqrt{6}$$

Then $\Sigma = \begin{bmatrix} \sqrt{30} & 0 \\ 0 & \sqrt{6} \end{bmatrix}$

Question 3

Part a

We can see, it's possible to decompose A into a SDV

$$A = U \sum V^T$$

$$A = \begin{bmatrix} | & & | \\ U_1 & \cdots & U_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_d \end{bmatrix} \begin{bmatrix} - & V_1 & - \\ & \vdots & \\ - & V_d & - \end{bmatrix}$$

We know that the product of $U \sum$ is $:\sum_{i=1} \begin{bmatrix} | \\ U_i \\ | \end{bmatrix} \begin{bmatrix} \sigma_i & 0 & ... & 0 \end{bmatrix}$, where the dimensions are nx1 and 1xd(rank- 1). But as the elements of $\sum$ that are not in the diagonal provokes that much of the n by d matrix are zero. Then taking into consideration we can show that, $U \sum = \begin{bmatrix} | & & | \\ \sigma_1 U_1 & \cdots & \sigma_n U_n \\ | & & | \end{bmatrix}$, then if we use the outer product distribution again we transform the SVD decompomspotion in the following form:

$$A = \sum_{i=1}^{min(n,d)} \sigma_i \begin{bmatrix} | \\ U_i \\ | \end{bmatrix} \begin{bmatrix} - & V_i & - \end{bmatrix}$$

Part b

Since the best basis for the best subspace of the SVM matrix is the one with higher singular value, this imply that is u_1

Part c

1. $X^T = (U\Sigma V^T)^T = V\Sigma^T U^T$
2. $XX^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma^2 U^T$, since V is orthonormal then $V^T V = I$
3. $X^T X = V\Sigma^T U^T U \Sigma V^T = V\Sigma^2 V^T$, since U is othonormal then $U^T U = I$

Part d

d.i

Since Xw=0 we could argue that this already a basis if we cannot find a w vector such Xw=0 where $w! = 0$. Meaning that all columns in X are LI. In the other case, if we find some solution different to the trivial one we could construct one Basis building the basis as part of the other vectors LI of the columns X.

d.ii

We know that the problem in this case is :

$$\widetilde{w} = argmin\,(\widetilde{y} - \Sigma\widetilde{w})^T\,(\widetilde{y} - \Sigma\widetilde{w})$$

We know that $\widetilde{y} = U^T y$ and $\widetilde{w} = V^T w$, if we replace in the problem we get:

$$\widetilde{w} = argmin\left(U^T y - \Sigma V^T w\right)^T\left(U^T y - \Sigma V^T w\right)$$

$$y^T U^T U y - 2V\Sigma^T U^T y w + w^T V \Sigma^T \Sigma V^T w$$

$$y^T y - 2V\Sigma^T U^T y w + w^T V \Sigma^T \Sigma V^T w$$

Then using the fact that $U^T U = I$, $X^T = (U\Sigma V^T)^T = V\Sigma^T U^T$ and $X^T X = V\Sigma^T U^T U \Sigma V^T = V\Sigma^2 V^T$

$$y^T y - 2X^T y w + w^T X^T X w$$

Which is the same as :

$$(y - Xw)^T\,(y - Xw)$$

We can see that using this results the problem given is exactly the same for regular OLS. Now if we continue solving with using the replacements we get:

$$\widetilde{w} = argmin\,(\widetilde{y} - \Sigma\widetilde{w})^T\,(\widetilde{y} - \Sigma\widetilde{w})$$

$$\widetilde{y}^T\widetilde{y} - 2\Sigma^T\widetilde{y}\widetilde{w} + \widetilde{w}^T\Sigma^T\Sigma\widetilde{w}$$

Now if we take FOC for this equation we have:

$$-2\Sigma^T\widetilde{y} + 2\Sigma^T\Sigma\widetilde{w} = 0$$

$$\Sigma^T\Sigma\widetilde{w} = \Sigma^T\widetilde{y}$$

$$\widetilde{w} = \left(\Sigma^T\Sigma\right)^{-1}\Sigma^T\widetilde{y}$$

d.iii

Using the facts that $\widetilde{y} = U^Ty$ and $\widetilde{w} = V^Tw$, if we replace in the problem we get:

$$\widetilde{w} = \left(\Sigma^T\Sigma\right)^{-1}\Sigma^T\widetilde{y}$$

$$V^Tw = \left(\Sigma^T\Sigma\right)^{-1}\Sigma^TU^Ty$$

Now if we multiply by V in both sides we get:

$$VV^Tw = \left(\Sigma^T\Sigma\right)^{-1}V\Sigma^TU^Ty$$

We know that $VV^T = I$ and $X^T = (U\Sigma V^T)^T = V\Sigma^TU^T$

$$w = \left(\Sigma^T\Sigma\right)^{-1}X^Ty$$

Then as we know between $\Sigma^T\Sigma$ are more than one I matrix, and we introduce $U^TU = I$ and $V^TV = I$ and we replace by X^TX

$$w = \left(V\Sigma^TU^TU\Sigma V^T\right)^{-1}X^Ty$$

$$w = \left(X^TX\right)^{-1}X^Ty$$

Then we had show that we can convert $\widetilde{w}$ into the regular OLS solution. Then both solutions are equivalent but in different subspaces.