

# CS663 Assignment-5

Saksham Rathi, Kavya Gupta, Shravan Srinivasa Raghavan

Department of Computer Science,  
Indian Institute of Technology Bombay

## Question 5

### Solution

efege

#### 1 Part a

dwefw

#### 2 Part b

The objective is to find the orthonormal matrix  $R$  that minimises the square of the Frobenius norm of the error matrix  $E$  which equals  $P_1 - RP_2$ .

$$\begin{aligned} E(R) &= \|P_1 - RP_2\|_F^2 \\ \Rightarrow E(R) &= \text{Trace}((P_1 - RP_2)^T (P_1 - RP_2)) \\ &= \text{Trace}(P_1^T P_1 - (RP_2)^T P_1 - P_1^T (RP_2) + (RP_2)^T (RP_2)) \\ &= \text{Trace}(P_1^T P_1 - P_2^T R^T P_1 - P_1^T R P_2 + P_2^T R^T (RP_2)) \\ &= \text{Trace}(P_1^T P_1 + P_2^T R^T R P_2 - P_2^T R^T P_1 - P_1^T R P_2) \\ &= \text{Trace}(P_1^T P_1 + P_2^T P_2 - P_2^T R^T P_1 - P_1^T R P_2) \quad (R^T R = R R^T = I \text{ because } R \text{ is orthonormal}) \\ &= \text{Trace}(P_1^T P_1 + P_2^T P_2) - \text{Trace}(P_2^T R^T P_1) - \text{Trace}(P_1^T R P_2) \quad (\text{Because Trace is a linear operator}) \\ &= \text{Trace}(P_1^T P_1 + P_2^T P_2) - \text{Trace}((P_2^T R^T P_1)^T) - \text{Trace}(P_1^T R P_2) \quad (\text{Trace}(A) = \text{Trace}(A^T) \text{ for any square matrix } A) \\ &= \text{Trace}(P_1^T P_1 + P_2^T P_2) - \text{Trace}(P_1^T R P_2) - \text{Trace}(P_1^T R P_2) \\ &= \text{Trace}(P_1^T P_1 + P_2^T P_2) - 2\text{Trace}(P_1^T R P_2) \end{aligned}$$

#### 3 Part c

As shown in the previous part  $E(R) = \text{Trace}(P_1^T P_1 + P_2^T P_2) - 2\text{Trace}(P_1^T R P_2)$ . Clearly for given  $P_1$  and  $P_2$ , the first term is fixed. Therefore to minimise  $E(R)$  wrt  $R$ , the term  $-2\text{Trace}(P_1^T R P_2)$  should be minimised ie,  $\text{Trace}(P_1^T R P_2)$  should be maximised wrt  $R$ .

## 4 Part d

Now the goal is to maximise  $\text{Trace}(P_1^T R P_2)$  wrt  $R$ . We will first show that for any two matrices  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $\text{Trace}(AB) = \text{Trace}(BA)$ .

$$\begin{aligned} \text{Trace}(AB) &= \sum_{i=1}^m C_{ii} \text{ (where } C = AB) \\ \Rightarrow \text{Trace}(AB) &= \sum_{i=1}^m \left( \sum_{j=1}^n A_{ij} B_{ji} \right) \end{aligned} \quad (1)$$

Similarly,

$$\begin{aligned} \text{Trace}(BA) &= \sum_{j=1}^n D_{jj} \text{ (where } D = BA) \\ \Rightarrow \text{Trace}(BA) &= \sum_{j=1}^n \left( \sum_{i=1}^m B_{ji} A_{ij} \right) \end{aligned} \quad (2)$$

Clearly, the two double summations in the RHS of equation 1 and equation 2 are equal and can be obtained by simply reordering the summations. Therefore we have

$$\text{Trace}(AB) = \text{Trace}(BA) \quad (3)$$

We will also use the fact that the SVD of a square matrix  $M = USV^T$  exists with  $U, V$  being orthonormal matrices and  $S$  being a diagonal matrix.

$$\begin{aligned} \text{Trace}(P_1^T R P_2) &= \text{Trace}(R P_2 P_1^T) \text{ (from equation 3)} \\ &= \text{Trace}(R U' S' V'^T) \text{ (where } P_2 P_1^T = U' S' V'^T \text{ is the SVD of } P_2 P_1^T) \\ &= \text{Trace}(S' V'^T R U') \text{ (again using the result from equation 3)} \\ &= \text{Trace}(S' X) \text{ (where } X = V'^T R U' \text{ and } S' \text{ is a diagonal matrix)} \end{aligned}$$

## 5 Part e

Since maximising  $\text{Trace}(P_1^T R P_2)$  is the same as maximising  $\text{Trace}(S' X)$  where  $S'$  is a diagonal matrix and  $X = V'^T R U'$ , we only need to determine the matrix  $X$  such that  $S'_{11} X_{11} + S'_{22} X_{22}$  is maximised where  $S'_{11}$  and  $S'_{22}$  are the singular values of  $P_2 P_1^T$ .

We will now show that  $X$  is also orthonormal.

$$\begin{aligned} XX^T &= (V'^T R U') \cdot (V'^T R U')^T \\ &= (V'^T R U') \cdot (U'^T R^T V') \\ &= V'^T R (U' \cdot U'^T) R^T V' \\ &= V'^T R I R^T V', \text{ (since } U, V \text{ are orthonormal)} \\ &= V'^T R R^T V' \\ &= V'^T I V' \text{ (since } R \text{ is orthonormal)} \\ &= V'^T V' \\ &= I \\ \Rightarrow XX^T &= X^T X = I \\ \Rightarrow X_{11} X_{11} + X_{12} X_{12} &= 1 \text{ \& } X_{11} X_{21} + X_{12} X_{22} = 0 \\ \Rightarrow X_{21} X_{11} + X_{22} X_{12} &= 0 \text{ \& } X_{21} X_{21} + X_{22} X_{22} = 1 \\ \Rightarrow X_{11}^2 + X_{12}^2 &= 1 \\ \Rightarrow X_{21}^2 + X_{22}^2 &= 1 \\ \Rightarrow X_{11} X_{21} + X_{12} X_{22} &= 0 \end{aligned}$$

Since all entries of  $X$  are real numbers,  $0 \leq X_{ij} \leq 1 \forall i, j \in [1, 2]$ . We need to maximise  $S'_{11}X_{11} + S'_{22}X_{22}$  and since  $S'_{11}$  and  $S'_{22}$  are positive values, we only need to maximise the values of  $X_{11}$  and  $X_{22}$ . Since  $X_{11} \leq 1$  and  $X_{22} \leq 1$  the maximum values  $X_{11}$  and  $X_{22}$  can take are 1. Hence,

$$\text{Trace}(S'X) = X_{11}S'_{11} + X_{22}S'_{22} \leq ((1)S'_{11} + (1)S'_{22}) = S'_{11} + S'_{22}$$

Therefore the matrix  $X$  is,

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

since  $x_{11}^2 + x_{12}^2 = 1$  and  $x_{11} = 1 \Rightarrow x_{12} = 0$ , similarly  $x_{21} = 0$ . Hence we get  $X = I_2$  for maximum value of  $\text{Trace}S'X$

## 6 Part f

Since  $X = I_2$  and  $X = V'^T R U'$ , we get

$$\begin{aligned} V'^T R U' &= I \\ \Rightarrow V' V'^T R U' U'^T &= V' I U'^T \\ \Rightarrow I R I &= V' U'^T \text{ (since } V' \text{ and } U' \text{ are orthogonal matrices)} \\ \Rightarrow R &= V' U'^T \end{aligned}$$

If  $X$  is the identity matrix  $R = V' U'^T$ .

## 7 Part g

If  $R$  is also a rotation matrix then  $R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$  for some  $\theta \in \mathbb{R}$ . This gives  $\det(R) = \cos^2 \theta + \sin^2 \theta = 1$ . However if  $R$  is not a rotation matrix then  $\det(R) = 1, -1$ . So the additional constraint imposed is  $\det(R) = 1$ .