5) A)

Given: $P_1 = RP_2 + E$

R is orthonormal.

Aim: To find R given P_1, P_2 under constraint R is orthonormal

a) $R = P_1 P_2^T (P_2 P_2^T)^{-1}$ will fail, because

$P_1 P_2^T (P_2 P_2^T)^{-1}$ is not orthonormal

Proof:

consider RR^T

$$= \left(P_1 P_2^T (P_2 P_2^T)^{-1}\right) \cdot \left(P_1 P_2^T (P_2 P_2^T)^{-1}\right)^T$$

$$= \left(P_1 P_2^T (P_2 P_2^T)^{-1}\right) \cdot \left(\left((P_2 P_2^T)^T\right)^{-1} \cdot (P_2^T)^T \cdot P_1^T\right) \quad (\because (ABC)^T = C^T B^T A^T)$$

$$= P_1 P_2^T (P_2 P_2^T)^{-1} \cdot (P_2 \cdot P_2^T)^{-1} \cdot P_2 \cdot P_1^T \quad (\because (P_2 P_2^T)^T = (P_2^T)^T P_2^T = P_2 P_2^T)$$

$$= P_1 P_2^T \left((P_2 P_2^T)^{-1}\right)^2 P_2 P_1^T$$

$\therefore RR^T \neq I$

$\therefore$ R is not orthonormal

So $R = P_1 P_2^T (P_2 P_2^T)^{-1}$ will fail.

b) $E(R) = \|P_1 - RP_2\|_F^2$

$= trace((P_1 - RP_2)^T (P_1 - RP_2))$

$= trace(P_1^T P_1 + P_2^T R^T R P_2 - P_2^T R^T P_1 - P_1^T R P_2)$

As R is orthornormal $\Rightarrow RR^T = I = R^T R$

substitute $\boxed{R^T R = I}$ in the above equation

$\therefore E(R) = \text{trace}(P_1^T P_1 + P_2^T P_2 - P_2^T R^T P_1 - P_1^T R P_2)$ — ①

($\because R^T R = I$, as R is orthonormal)

w.k.t $\text{trace}(A) = \text{trace}(A^T)$

substitute $A = P_1^T R P_2$ in the above equation

$\Rightarrow$ ~~tr~~ $\text{trace}(P_1^T R P_2) = \text{trace}((P_1^T R P_2)^T)$

$= \text{trace}(P_2^T R^T P_1)$ ($\because (ABC)^T = C^T B^T A^T$, $(A^T)^T = A$)

$\Rightarrow \text{trace}(P_2^T R^T P_1) = \text{trace}(P_1^T R P_2)$

substuting it in eqn ① gives

$E(R) = \text{trace}(P_1^T P_1 + P_2^T P_2) - \text{trace}(P_2^T R^T P_1) - \text{trace}(P_1^T R P_2)$

$= \text{trace}(P_1^T P_1 + P_2^T P_2) - \text{trace}(P_1^T R P_2) - \text{trace}(P_1^T R P_2)$

$= \text{trace}(P_1^T P_1 + P_2^T P_2) - 2\,\text{trace}(P_1^T R P_2)$

($\because \text{trace}(P_1^T R P_2) = \text{trace}(P_2^T R^T P_1)$)

c)

we need to Prove miniming $E(R)$ w.r.t R is equalent to maximising $\text{trace}(P_1^T R P_2)$ w.r.t R.

Proof:-

$E(R) = \text{trace}(P_1^T P_1 + P_2^T P_2) - 2\,\text{trace}(P_1^T R P_2)$

$\text{trace}(P_1^T P_1 + P_2^T P_2)$ is independent of R.

So to minimize $E(R)$ w.r.t R, we need to minimize

$-2\,\text{trace}(P_1^T R P_2)$

minimizing $-\text{tr}(P_1^T R P_2) \Rightarrow$ maximizing $\text{trace}(P_1^T R P_2)$

$\therefore$ minimizing $E(R)$ w.r.t $R \equiv$ maximizing $\text{trace}(P_1^T R P_2)$ w.r.t R

d)

$$\operatorname{trace}(AB) = \operatorname{trace}(BA)$$

Proof:-

$$\operatorname{trace}(AB) = \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ji}$$

$$\operatorname{trace}(BA) = \sum_{i=1}^{n} (BA)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij} a_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ji}$$

$$= \operatorname{trace}(AB)$$

$\therefore \operatorname{trace}(AB) = \operatorname{trace}(BA)$

$$\operatorname{trace}(P_1^T R P_2) = \operatorname{trace}((P_1^T)(R P_2))$$

$$= \operatorname{trace}(R P_2 P_1^T) \quad (\because \operatorname{trace}(AB) = \operatorname{trace}(BA))$$

$$= \operatorname{trace}(R U' S' V'^T) \text{ using SVD of } P_1 P_2^T = U' S' V'^T.$$

$$= \operatorname{trace}(S' V'^T R U') = \operatorname{trace}(S'X) \text{ where } X = V'^T R U^T$$

e) we need to find, for matrix x will above expression will be maximised, where

X is a 2×2 matrix with the constraint X is orthonormal

$$X^T X = (V'^T R U'^T)^T (V'^T R U')$$

$$= (U')^T R^T \underbrace{V' V'^T}_{I} R U' \quad (\text{sincee } V', R, U' \text{ are orthonormal})$$

$$\underbrace{}_{I}$$

$$\underbrace{}_{I}$$

$\therefore X^T X = I.$

Claim:- Any ~~not~~ 2×2 orthonormal matrix can be represented in the form

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ (or) } \begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

proof

let matrix be $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & c \\ b & d \end{bmatrix} = I, \quad \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = I$$

$\Rightarrow a^2+b^2=1$, $ac+bd=0$, $ac+bd=0$, $c^2+d^2=1$,

$\Rightarrow a^2+c^2=1$, $ab+cd=0$, $b^2+d^2=1$

$\Rightarrow a=\pm d$, $b=\pm c$, $ac+bd=0$

$\Rightarrow$ case-1 :-

$a=d$, $b=-c$ and $a^2+b^2=1$

take $a=\cos\theta$, $b=\sin\theta$, so matrix is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

case :- 2

$a=-d$, $b=c$, $a^2+b^2=1$

take $d=\cos\theta$, $b=\sin\theta$, we get $\begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

w.k.t S' is a diogonal matrix with non-negative entries

let $S' = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ $\lambda_1, \lambda_2 \geq 0$

trace $(S'x)$ in 1st case is trace $\begin{bmatrix} \lambda_1\cos\theta & -\lambda_1\sin\theta \\ \lambda_2\sin\theta & \lambda_2\cos\theta \end{bmatrix}$

$= (\lambda_1+\lambda_2)\cos\theta$.

it is maximum when $\theta = 0^\circ$

trace$(S'x)$ in case-2 is trace $\begin{bmatrix} -\lambda_1\cos\theta & \lambda_1\sin\theta \\ \lambda_2\sin\theta & \lambda_2\cos\theta \end{bmatrix}$

$= (\lambda_1-\lambda_2)\cos\theta$

it's maximum value $= |\lambda_1-\lambda_2|$

$\therefore$ maximum value of trace$(S'x) = \lambda_1+\lambda_2$, which we will acheive by placing $\theta=0^\circ$ in case-①.

$\therefore X$ which maximizes this equation is $= \begin{bmatrix} \cos(0) & -\sin(0) \\ \sin(0) & \cos(0) \end{bmatrix}$

$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$

f)

w.k.t $X = (V')^T R U'$

$\Rightarrow I = (V')^T R U'$

multiply with V' on left side on both L.H.S and R.H.S

$\Rightarrow V' = V'(V')^T R U'$

multiply with $(U')^T$ on righ side on both L.H.S and R.H.S

$$\Rightarrow V'(U')^T = \overbrace{(V')(V')^T}^{I} R \overbrace{U'(U')^T}^{I}$$

$\Rightarrow V'(U')^T = R$ ($\because$ V', U' are orthonormal matrices)

g)

For R to be specifically a rotation matrix, R should be matrix of type 1st case in previous part

i.e R should be of form $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

If R is of the form $\begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ i.e case-2 it is not specifically a Rotation matrix.

So extra condition that needs to added is $\det(R) = +1$,

Since $\det(R)$ for 1st case is $\cos^2\theta + \sin^2\theta = 1$ but in 2nd case $\det(R) = -\cos^2\theta - \sin^2\theta = 1$

$\therefore$ we need to impose $\det(R) = 1$ as extra condition for R to be specifically a Rotation matrix.