Rotation Matrix In High Dimensional Space

Zhu Beier
School of Computer Science and Engineering
NTU

Given two points P and Q in high dimensional space $\mathbb{R}^n$, given orignal point O, calculate the rotation matrix R to rotate the vector $\vec{OP}$ to $\vec{OQ}$.

Denote the unit vectors $\boldsymbol{a} = \frac{\vec{OP}}{\|\vec{OP}\|} \in \mathbb{R}^n$ and $\boldsymbol{b} = \frac{\vec{OQ}}{\|\vec{OQ}\|} \in \mathbb{R}^n$, $A = [\boldsymbol{a}, \boldsymbol{b}] \in \mathbb{R}^{n \times 2}$.

Given an abitrary vector $\boldsymbol{x}$, firstly decompose $\boldsymbol{x}$ into two pars:

$$\boldsymbol{x} = \boldsymbol{x}_\perp + \boldsymbol{x}_\parallel \tag{1}$$

where $\boldsymbol{x}_\perp$ is orthogonal to the plane OPQ and $\boldsymbol{x}_\parallel$ is parallel to the plane OPQ.

Derive that $\boldsymbol{x}_\parallel = A(A^TA)^{-1}A^T\boldsymbol{x}$, $\boldsymbol{x}_\perp = (I - A(A^TA)^{-1}A^T)\boldsymbol{x}$.

Because $\boldsymbol{x}_\parallel$ is spaned by $\boldsymbol{a}$ and $\boldsymbol{b}$, it can be represented as: $\boldsymbol{x}_\parallel = \lambda_1\boldsymbol{a} + \lambda_2\boldsymbol{b}$, $\boldsymbol{x}_\perp$ can be represented as $\boldsymbol{x}_\perp = \boldsymbol{x} - (\lambda_1\boldsymbol{a} + \lambda_2\boldsymbol{b})$, where λ_1, λ_2 are subject to

$$\lambda_1, \lambda_2 = \arg\min_{\lambda_1, \lambda 2}\|\boldsymbol{x} - (\lambda_1\boldsymbol{a} + \lambda_2\boldsymbol{b})\|_2^2 \tag{2}$$

$$\mathcal{L} = \|\boldsymbol{x} - (\lambda_1\boldsymbol{a} + \lambda_2\boldsymbol{b})\|_2^2 = \boldsymbol{x}^T\boldsymbol{x} - 2\lambda_1\boldsymbol{x}^T\boldsymbol{b} + 2\lambda_1\lambda_2\boldsymbol{a}^T\boldsymbol{b} + \lambda_1^2 + \lambda_2^2 \tag{3}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_1} = -2\boldsymbol{x}^T\boldsymbol{a} + 2\lambda_1 + 2\lambda_2\boldsymbol{a}^T\boldsymbol{b} = 0 \Rightarrow \lambda_1 + \lambda_2\boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{a}^T\boldsymbol{x} \tag{4}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_2} = -2\boldsymbol{x}^T\boldsymbol{b} + 2\lambda_2 + 2\lambda_1\boldsymbol{a}^T\boldsymbol{b} = 0 \Rightarrow \lambda_1\boldsymbol{a}^T\boldsymbol{b} + \lambda_2 = \boldsymbol{b}^T\boldsymbol{x} \tag{5}$$

$$\Rightarrow \begin{bmatrix} 1 & \boldsymbol{a}^T\boldsymbol{b} \\ \boldsymbol{a}^T\boldsymbol{b} & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}^T\boldsymbol{x} \\ \boldsymbol{b}^T\boldsymbol{x} \end{bmatrix} \tag{6}$$

$$\Rightarrow A^TA\boldsymbol{\lambda} = A\boldsymbol{x} \tag{7}$$

where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2]^T$

$$\Rightarrow \boldsymbol{x}_\parallel = A\boldsymbol{\lambda} = A(A^TA)^{-1}A\boldsymbol{x} \tag{8}$$

$\square$

It is obviously that rotation only affects $\boldsymbol{x}_\parallel$, and $\boldsymbol{x}_\perp$ remains unchanged. Thus the rotated vector $\boldsymbol{x}'$ can be represented as

$$\boldsymbol{x}' = \boldsymbol{x}_\perp + \boldsymbol{x}'_\parallel \tag{9}$$

Use $\boldsymbol{a}, \boldsymbol{b}$ as basis to span a space OPQ, the matrix to rotate $\vec{OP}$ to $\vec{OQ}$ can be represented as:

$$\begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} \tag{10}$$

Preprint. Under review.

In such space $\boldsymbol{a} = A[1,0]^T, \boldsymbol{b} = A[0,1]^T$, suppose the rotate matrix in the OPQ space is $R_{2\times2}$, then $R_{2\times2}[1,0]^T = [0,1]^T$, $R_{2\times2}[0,1]^T = [\lambda_a, \lambda_b]^T$

$$\boldsymbol{b}' = \lambda_a \boldsymbol{a} + \lambda_b \boldsymbol{b} \tag{11}$$

Suppose θ is angle between $\boldsymbol{b}$ and $\boldsymbol{a}$

$$\boldsymbol{b}^T\boldsymbol{b}' = \lambda_a \boldsymbol{b}^T\boldsymbol{a} + \lambda_b \boldsymbol{b}^T\boldsymbol{b} = \lambda_a \boldsymbol{b}^T\boldsymbol{a} + \lambda_b = \cos(\theta) = \boldsymbol{a}\boldsymbol{b}^T \tag{12}$$

$$\boldsymbol{a}^T\boldsymbol{b}' = \lambda_a \boldsymbol{a}^T\boldsymbol{a} + \lambda_b \boldsymbol{a}^T\boldsymbol{b} = \lambda_a + \lambda_b \boldsymbol{a}^T\boldsymbol{b} = \cos(2\theta) = 2\cos(\theta)^2 - 1 = 2(\boldsymbol{a}^T\boldsymbol{b})^2 - 1 \tag{13}$$

$$\Rightarrow \begin{bmatrix} 1 & \boldsymbol{a}^T\boldsymbol{b} \\ \boldsymbol{a}^T\boldsymbol{b} & 1 \end{bmatrix} \begin{bmatrix} \lambda_a \\ \lambda_b \end{bmatrix} = \begin{bmatrix} 2(\boldsymbol{a}^T\boldsymbol{b})^2 - 1 \\ \boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} \tag{14}$$

$$\Rightarrow \begin{bmatrix} \lambda_a \\ \lambda_b \end{bmatrix} = \begin{bmatrix} -1 \\ 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} \tag{15}$$

$$\Rightarrow R_{2\times2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} \tag{16}$$

$$\Rightarrow R_{2\times2} = \begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} \tag{17}$$

We can derive that

$$\boldsymbol{x}'_{\parallel} = A \begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} (A^TA)^{-1}A^T\boldsymbol{x} \tag{18}$$

Thus,

$$\boldsymbol{x}' = A \begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} (A^TA)^{-1}A^T\boldsymbol{x} + (I - A(A^TA)^{-1}A^T)\boldsymbol{x} \tag{19}$$

The rotation matrix R can be expressed as:

$$A \begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} (A^TA)^{-1}A^T + I - A(A^TA)^{-1}A^T \tag{20}$$

$$\Rightarrow I + A(\begin{bmatrix} 0 & -1 \\ 1 & 2\boldsymbol{a}^T\boldsymbol{b} \end{bmatrix} - I_{2\times2})(A^TA)^{-1}A^T \tag{21}$$