Let the known set of correspondances be. $P = \{p_1, p_2, \dots, p_n\}$ & $Q = \{q_1, q_2, \dots, q_n\}$ in $\mathbb{R}^d$.

We seek rigid body transformation that optimally aligns the above two sets.

$$(R, t) = \underset{R \in SO(d),\, t \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n} w_i \|(Rp_i + t) - q_i\|^2.$$

$w_i > 0$, are weights for each point pair.

For computing translation t, assume R is fixed.

Let $E(t) = \sum_{i=1}^{n} w_i \|(Rp_i + t) - q_i\|^2$.

For finding optimal translation,

$$\frac{\partial E(t)}{\partial t} = 0 \Rightarrow \sum_{i=1}^{n} w_i \cdot 2\left[(Rp_i + t) - q_i\right] = 0$$

$$\Rightarrow 2t\left(\sum_{i=1}^{n} w_i\right) + 2R\left(\sum_{i=1}^{n} w_i p_i\right) - 2\sum_{i=1}^{n} w_i q_i$$

Let $\bar{p} = \dfrac{\sum_{i=1}^{n} w_i p_i}{\sum_{i=1}^{n} w_i}$ & $\bar{q} = \dfrac{\sum_{i=1}^{n} w_i q_i}{\sum_{i=1}^{n} w_i}$.

$$\Rightarrow \cancel{2}t + \cancel{2}R\bar{p} - \cancel{2}\bar{q} = 0$$

$$\Rightarrow t = \bar{q} - R\bar{p}$$

For the above calculated value of t,

$$(R, \cancel{t}) = \underset{R \in SO(d),\, t \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n} w_i \|(Rp_i + \bar{q} - R\bar{p} - q_i\|^2.$$

$$= \underset{R \in SO(d),\, t \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n} w_i \| R(p_i - \bar{p}) + (\bar{q} - q_i) \|^2$$

$$= \underset{R \in SO(d),\, t \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n} w_i \| R(p_i - \bar{p}) - (q_i - \bar{q}) \|^2$$

②

Let $x_i^\circ := P_i^\circ - \bar{P}$ &

$y_i^\circ := q_i^\circ - \bar{q}$

$$R = \underset{R \in SO(d)}{\arg\min} \sum_{i=1}^{n} w_i^\circ \| R x_i^\circ - y_i^\circ \|^2.$$

For computing rotation

Consider
$$\|R x_i^\circ - y_i^\circ\|^2 = (R x_i^\circ - y_i)^T (R x_i - y_i^\circ)$$
$$= (x_i^T R^T - y_i^T)(R x_i^\circ - y_i)$$
$$= (x_i^T R^T R x_i^\circ - x_i^T R^T y_i^\circ - y_i^T R x_i^\circ + y_i^T y_i^\circ)$$
$$= x_i^T x_i^\circ - x_i^T R^T y_i^\circ - y_i^T R x_i^\circ + y_i^T y_i^\circ. \quad (\because RR^T = R^T R = I)$$

$$x_i^T R^T y_i^\circ = (x_i^T R^T y_i^\circ)^T = y_i^T R x_i^\circ \quad (\because x_i^T R^T y_i^\circ \text{ is scalar}).$$

$$\therefore \|R x_i^\circ - y_i^\circ\|^2 = x_i^T x_i^\circ - 2 y_i^T R x_i^\circ + y_i^T y_i^\circ.$$

$$\therefore R = \underset{R \in SO(d)}{\arg\min} \sum_{i=1}^{n} w_i \| R x_i^\circ - y_i^\circ \|^2$$
$$= \underset{R \in SO(d)}{\arg\min} \sum_{i=1}^{n} w_i^\circ (x_i^T x_i - \underbrace{2 y_i^T R x_i^\circ}_{\text{Only term involving } R.} + y_i^T y_i^\circ).$$

$$\therefore R = \underset{R \in (SO(d))}{\arg\min} \sum_{i=1}^{n} (-2 y_i^T R x_i^\circ) . w_i^\circ$$
$$= 2 \underset{R \in SO(d)}{\arg\max} \sum_{i=1}^{n} y_i^T R x_i^\circ . w_i^\circ \quad \text{——— ①}$$

Let $W = \text{diag}(w_1^\circ, w_2, \dots, w_n)_{n \times n}$.

$Y^T = \begin{bmatrix} y_1^T \\ \vdots \\ y_n^T \end{bmatrix}_{n \times d}$, & $X = [x_1 \ x_2 \ x_3 \ \dots \ x_n]_{d \times n}$

Equation 1 can now be written as,

$$R = \underset{R \in SO(d)}{\operatorname{argmax}} \; tr(WY^T R X)$$

$$tr(WY^T R X) = tr((RX)(WY^T)) = tr(R\underbrace{XWY^T}_{=S})$$

Let $S = XWY^T$.

Now, $\max tr(RXWY^T) = \max tr(RS)$

We find the maximum of RS ~~by~~ from SVD of S.

$SVD(S) = U\Sigma V^T$. U, V are $d \times d$ orthogonal matrices
Σ is $d \times d$ diagonal matrix containing the singular values of S (non-negative) which appear in descending value of their magnitude

$$\therefore tr(RS) = tr(RU\Sigma V^T) = tr(\Sigma V^T R U) \quad (\because tr((RU)(\Sigma V^T)) = tr((\Sigma V^T)(RU)))$$

Let $S' = \cancel{\Sigma}\, V^T R U$

As U, Σ, V are orthogonal so is S.

$$\therefore s_j'^T s_j' = 1 \quad (s_j' \text{ are columns of matrix of } s')$$

$$\therefore \sum_{i=1}^{d} s_{ij}'^2 = 1 \Rightarrow$$

$$\Rightarrow s_{ij}'^2 \le 1 \Rightarrow |s_{ij}'| \le 1.$$

$$\therefore tr(\Sigma V^T R U) = tr(\Sigma S').$$

$$tr(\Sigma S') = \underbrace{\begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_3 & & \\ & & & \ddots & \\ & & & & \sigma_n \end{bmatrix}}_{d \times n} \underbrace{\begin{bmatrix} s_{1j}' \\ s_{2j}' \\ \vdots \\ s_{nj}' \end{bmatrix}}_{n \times d} =$$

$$= \sum_{i=1}^{d} \sigma_i s_{ii}' \le \sum_{i=1}^{d} \sigma_i.$$

$\therefore$ For $tr(\Sigma S')$ to be maximum, $S'_{ii} = 1$.
But as S' is also orthogonal, therefore $S' = I$.

$\therefore \quad S' = I = V^T R U$

$\Rightarrow \quad VI = VV^T R U \Rightarrow V = IRU \Rightarrow V = RU$.

$\therefore \quad R^T V = R^T R U \Rightarrow R^T V = U$

$\Rightarrow R = VU^T$.

If P is a reflection of Q or vice versa, $R_1 = \underset{R \in SO(d)}{\operatorname{argmin}} \sum_{i=1}^{n} w_i \|Rx_i - y$
yields zero energy.

$\therefore$ If $\det(VU^T) = -1$, then it contains reflection.

For maximising $tr(\Sigma S) = \sigma_1 + \sigma_2 + \dots + \sigma_{d-1} - \sigma_d$.
We consider $(-\sigma_d)$ as it is the smallest value among σ_i's.

$\therefore$ If $\det(VU^T) = -1$, then, $R = V \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots & \\ & & & & \det(VU^T) \end{bmatrix} U^T$.

$\therefore \quad R = V.\operatorname{diag}(1, 1, \dots (d-1) \text{ times}, \det(VU^T)). U^T$.

Hence Proved.