

# Lecture 4: Alignment Problems

**Radu Balan**

University of Maryland

February 25, 2020

# Alignment Problems

Assume we have two geometric graphs,  $\mathbb{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$  and  $\mathbb{Y} = \{y_1, \dots, y_n\} \subset \mathbb{R}^d$ . Today we discuss how to best align these two sets of points. Specifically we discuss the following *alignment problems*:

- 1 Procrustes problem: Find the rotation transformation that maps one set of points closest to the other set of points;
- 2 Classical Procrustes problem: Find the translation and rotation transformations that map one set of points closest to the other set of points;
- 3 Full alignment problem: Find the translation, rotation and scaling that map optimally one set of points to the other set of points;

We shall not discuss the *graph matching problem*, a related but much harder (NP-hard) problem.

# Alignment Problems

## The Procrustes Problem

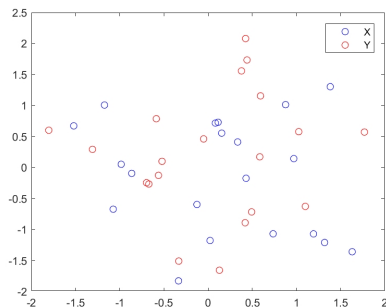
Given matrices  $X, Y \in \mathbb{R}^{d \times n}$  whose columns are the  $n$  points from each set  $\mathbb{X}, \mathbb{Y}$ , find an orthogonal matrix  $Q \in O(d)$  that:

$$\begin{aligned} & \text{minimize} && \|Y - QX\|_F^2 \\ & Q \in O(d) \end{aligned}$$

where

$$\|A\|_F^2 = \text{trace}(A^T A) = \sum_{i,j} |A_{i,j}|^2$$

$$\left\| \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \end{bmatrix} \right\|_F^2 = 11$$

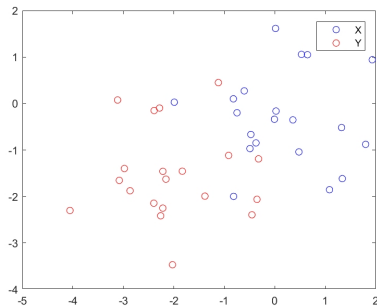


# Alignment Problems

## The Classical Procrustes Problem

Given matrices  $X, Y \in \mathbb{R}^{d \times n}$  whose columns are the  $n$  points from each set  $\mathbb{X}, \mathbb{Y}$ , find an orthogonal matrix  $Q \in O(d)$  and a vector  $z \in \mathbb{R}^d$  that:

$$\begin{aligned} & \text{minimize} && \|Y - Q(X - z\mathbf{1}^T)\|_F^2 \\ & Q \in O(d) \\ & z \in \mathbb{R}^d \end{aligned}$$

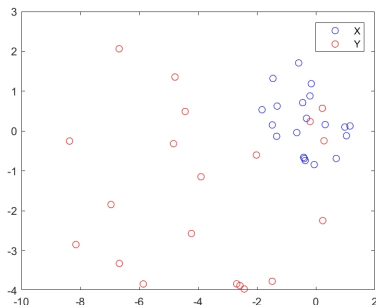


# Alignment Problems

## The Full Alignment Problem

Given matrices  $X, Y \in \mathbb{R}^{d \times n}$  whose columns are the  $n$  points from each set  $\mathbb{X}, \mathbb{Y}$ , find an orthogonal matrix  $Q \in O(d)$ , a vector  $z \in \mathbb{R}^d$  and a positive scalar  $a > 0$  that:

$$\begin{aligned} & \text{minimize} && \|Y - aQ(X - z1^T)\|_F^2 \\ & Q \in O(d) \\ & z \in \mathbb{R}^d \\ & a > 0 \end{aligned}$$



# The Optimization Problem

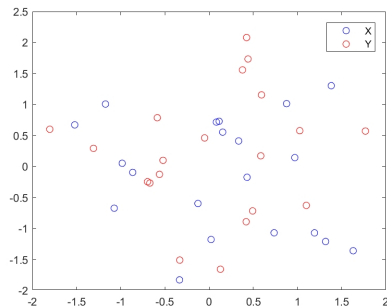
Given matrices  $X, Y \in \mathbb{R}^{d \times n}$  whose columns are the  $n$  points from each set  $\mathbb{X}, \mathbb{Y}$ , find an orthogonal matrix  $Q \in O(d)$  that:

$$\begin{aligned} & \text{minimize} && \|Y - QX\|_F^2 \\ & Q \in O(d) \end{aligned}$$

where

$$\|A\|_F^2 = \text{trace}(A^T A) = \sum_{i,j} |A_{i,j}|^2$$

$$\left\| \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \end{bmatrix} \right\|_F^2 = 11$$



# The solution to the Procrustes problem

## Algorithm (Schönemann 1964)

*Inputs: Matrices  $X, Y \in \mathbb{R}^{d \times n}$ .*

- 1 *Compute the  $d \times d$  matrix  $R = XY^T$ ;*
- 2 *Compute the Singular Value Decomposition (SVD),  $R = U\Sigma V^T$ , where  $U, V \in \mathbb{R}^{d \times d}$  are orthogonal matrices, and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$  is the diagonal matrix with singular values  $\sigma_1, \dots, \sigma_d \geq 0$  on its diagonal;*
- 3 *Compute  $Q = VU^T$ .*

*Output: Orthogonal matrix  $Q \in O(d) \subset \mathbb{R}^{d \times d}$ .*

## Derivation of the solution

The derivation of the solution is as follows. First recall a matrix  $Q \in \mathbb{R}^{d \times d}$  is said *orthogonal* if  $Q^{-1} = Q^T$ . Equivalently,  $Q^T Q = I_d$  or  $Q Q^T = I_d$ . Then note the set of orthogonal matrices  $O(d)$  forms a group: in particular, the product of two orthogonal matrices is still an orthogonal matrix: if  $Q_1, Q_2 \in O(d)$  then

$$(Q_1 Q_2)^T (Q_1 Q_2) = Q_2^T Q_1^T Q_1 Q_2 = Q_2^T I_d Q_2 = Q_2^T Q_2 = I_d.$$

Recall also that  $\text{trace}(AB) = \text{trace}(BA)$ .

1. We start by expanding the objective function:

$$\begin{aligned} \|Y - QX\|_F^2 &= \text{trace}((Y - QX)^T (Y - QX)) = \text{trace}(Y^T Y) - \text{trace}(X^T Q^T Y) - \\ &- \text{trace}(Y^T QX) + \text{trace}(X^T Q^T QX) = \|Y\|_F^2 - \text{trace}(Y^T QX) - \text{trace}(Y^T QX) + \\ &+ \text{trace}(X^T X) = \|Y\|_F^2 - 2\text{trace}(QXY^T) + \|X\|_F^2. \end{aligned}$$



## Derivation of the solution - 2

$$\|Y - QX\|_F^2 = \|Y\|_F^2 - 2\text{trace}(QXY^T) + \|X\|_F^2$$

Then:

$$\begin{aligned} \text{minimize } \|Y - QX\|_F^2 &\Leftrightarrow \text{maximize } \text{trace}(QXY^T) \\ Q \in O(d) & \end{aligned}$$

2. The SVD decomposition of matrix  $R = XY^T$ . One can form two symmetric matrices out of  $R$ :  $R^T R$  and  $RR^T$ . Each is positive semidefinite and diagonalizes:  $R^T R = VDV^T$  and  $RR^T = UEU^T$ , where both  $U$  and  $V$  are orthogonal matrices, and  $D, E$  are diagonal matrices. Fact: The eigenvalues of  $R^T R$  and  $RR^T$  are the same. Furthermore, if  $v$  is an eigenvector for  $R^T R$  then  $Rv$  is an eigenvector for  $RR^T$ . Why: Let  $(v, \sigma^2)$  be an eigenpair for  $R^T R$ :  $R^T Rv = \sigma^2 v$ . Then  $RR^T Rv = \sigma^2 Rv$ . This shows that  $(Rv, \sigma^2)$  is an eigenpair for  $RR^T$ . Similarly, if  $(u, \sigma^2)$  is an eigenpair for  $RR^T$ :  $RR^T u = \sigma^2 u$ , then  $R^T RR^T u = \sigma^2 R^T u$ .

## Derivation of the solution - 3

Thus  $(R^T, \sigma^2)$  is an eigenpair for  $R^T R$ .

Consequence:  $D = E$ . Let  $\Sigma^2 = D = E$  (that is,  $\Sigma = D^{1/2}$ ).

It follows, the SVD decomposition of  $R$  is given by  $R = U\Sigma V^T$ .

The maximization criterion becomes:

$$\text{trace}(QXY^T) = \text{trace}(QR) = \text{trace}(QU\Sigma V^T) = \text{trace}(V^T QU\Sigma)$$

Let  $\tilde{Q} = V^T QU$ . Then we need to maximize

$$\begin{aligned} & \text{maximize} && \text{trace}(\tilde{Q}\Sigma) \\ & \tilde{Q} \in O(d) \end{aligned}$$

Let  $\Sigma = \text{diag}(\sigma_k)_{1 \leq k \leq d}$  and  $\tilde{Q} = (q_{i,j})_{1 \leq i,j \leq d}$ . Then

$$\text{trace}(\tilde{Q}\Sigma) = \sum_{k=1}^d q_{k,k} \sigma_k$$

## Derivation of the solution - 4

Since  $\sum_{j=1}^d |q_{k,j}|^2 = 1$  it follows the maximum of the sum above is achieved when  $q_{1,1} = \dots = q_{d,d} = 1$ . In this case  $\tilde{Q} = I_d$  and  $\text{trace}(\tilde{Q}\Sigma) = \text{trace}(\Sigma)$ .

Thus  $V^T Q U = I_d$  and hence  $Q = V U^T$ .

The alignment error (mismatch) is given by:

$$\text{Err} = Y - QX \quad , \quad \|\text{Err}\|_F^2 = \|X\|_F^2 + \|Y\|_F^2 - 2 \text{trace}(\Sigma)$$

Note  $R = XY^T$  and  $R^T R = V \Sigma^2 V^T$ . Thus  $\Sigma^2$  is the diagonal form of  $R^T R = YX^T XY^T$ . It follows  $\text{trace}(\Sigma) = \text{trace}((R^T R)^{1/2})$  and

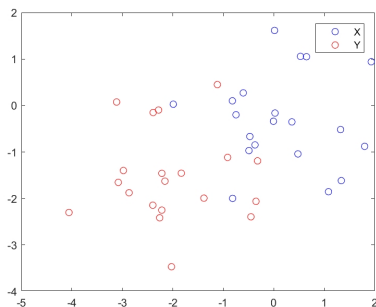
$$\|\text{Err}\|_F^2 = \text{trace}(XX^T) + \text{trace}(YY^T) - 2 \text{trace}((YX^T XY^T)^{1/2})$$

caveat: exponent 1/2 means matrix square root.

# The Optimization Problem

Given matrices  $X, Y \in \mathbb{R}^{d \times n}$  whose columns are the  $n$  points from each set  $\mathbb{X}, \mathbb{Y}$ , find an orthogonal matrix  $Q \in O(d)$  and a vector  $z \in \mathbb{R}^d$  that:

$$\begin{aligned} & \text{minimize} && \|Y - Q(X - z\mathbf{1}^T)\|_F^2 \\ & Q \in O(d) \\ & z \in \mathbb{R}^d \end{aligned}$$



# The solution to the classical Procrustes problem

## Algorithm (Rotation-Translation alignment)

Inputs: Matrices  $X, Y \in \mathbb{R}^{d \times n}$ .

- 1 Compute centers  $\bar{x} = \frac{1}{n}X \cdot \mathbf{1}$ ,  $\bar{y} = \frac{1}{n}Y \cdot \mathbf{1}$  and recenter data  $\tilde{X} = X - \bar{x} \cdot \mathbf{1}^T$ ,  $\tilde{Y} = Y - \bar{y} \cdot \mathbf{1}^T$ .
- 2 Compute the  $d \times d$  matrix  $R = \tilde{X} \tilde{Y}^T$ ;
- 3 Compute the Singular Value Decomposition (SVD),  $R = U \Sigma V^T$ , where  $U, V \in \mathbb{R}^{d \times d}$  are orthogonal matrices, and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$  is the diagonal matrix with singular values  $\sigma_1, \dots, \sigma_d \geq 0$  on its diagonal;
- 4 Compute  $Q = VU^T$  and  $z = \bar{x} - Q^T \bar{y}$ .

Output: Orthogonal matrix  $Q \in O(d) \subset \mathbb{R}^{d \times d}$ , translation vector  $z \in \mathbb{R}^d$ .

## Derivation of the solution

We start by introducing a new vector  $w = \bar{y} - Q(\bar{x} - z)$  so that

$$Y - Q(X - z1^T) = \tilde{Y} - Q\tilde{X} + w1^T$$

Then expand the objective function:

$$\begin{aligned} \|Y - Q(X - z1^T)\|_F^2 &= \|\tilde{Y}\|_F^2 + \|\tilde{X}\|_F^2 + \|w1^T\|_F^2 - 2 \operatorname{trace}(Q\tilde{X}\tilde{Y}^T) + \\ &\quad + 2\operatorname{trace}(w1^T\tilde{Y}) - 2\operatorname{trace}(Q\tilde{X}1w^T) \end{aligned}$$

But:  $\tilde{X}1 = 0$  and  $\tilde{Y}1 = 0$  because of the centering. It follows:

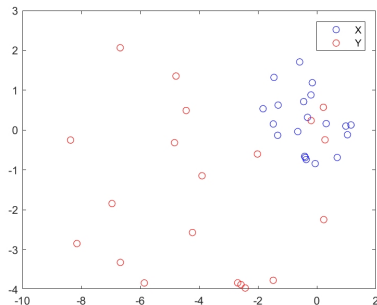
$$\|Y - Q(X - z1^T)\|_F^2 = \|\tilde{Y}\|_F^2 + \|\tilde{X}\|_F^2 + n\|w\|^2 - 2 \operatorname{trace}(QR)$$

where  $R = \tilde{X}\tilde{Y}^T$ . Then the minimum of this criterion is achieved at  $w = 0$  and  $Q$  that maximizes  $\operatorname{trace}(QR)$ , hence the previous algorithm.

# The Optimization Problem

Given matrices  $X, Y \in \mathbb{R}^{d \times n}$  whose columns are the  $n$  points from each set  $\mathbb{X}, \mathbb{Y}$ , find an orthogonal matrix  $Q \in O(d)$ , a vector  $z \in \mathbb{R}^d$  and a positive scalar  $a > 0$  that:

$$\begin{aligned} & \text{minimize} && \|Y - aQ(X - z\mathbf{1}^T)\|_F^2 \\ & Q \in O(d) \\ & z \in \mathbb{R}^d \\ & a > 0 \end{aligned}$$



# The solution to the full alignment problem

## Algorithm (Full alignment)

Inputs: Matrices  $X, Y \in \mathbb{R}^{d \times n}$ .

- 1 Compute centers  $\bar{x} = \frac{1}{n}X \cdot \mathbf{1}$ ,  $\bar{y} = \frac{1}{n}Y \cdot \mathbf{1}$  and recenter data  $\tilde{X} = X - \bar{x} \cdot \mathbf{1}^T$ ,  $\tilde{Y} = Y - \bar{y} \cdot \mathbf{1}^T$ .
- 2 Compute the  $d \times d$  matrix  $R = \tilde{X}\tilde{Y}^T$ ;
- 3 Compute the Singular Value Decomposition (SVD),  $R = U\Sigma V^T$ , where  $U, V \in \mathbb{R}^{d \times d}$  are orthogonal matrices, and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$  is the diagonal matrix with singular values  $\sigma_1, \dots, \sigma_d \geq 0$  on its diagonal;
- 4 Compute  $Q = VU^T$ ,  $a = \frac{\text{trace}(\Sigma)}{\|\tilde{X}\|_F^2}$  and  $z = \bar{x} - \frac{1}{a}Q^T\bar{y}$ .

Output: Orthogonal matrix  $Q \in O(d) \subset \mathbb{R}^{d \times d}$ , translation vector  $z \in \mathbb{R}^d$  and  $a > 0$ .



## Derivation of the solution

We start by introducing a new vector  $w = \bar{y} - aQ(\bar{x} - z)$  so that

$$Y - aQ(X - z1^T) = \tilde{Y} - aQ\tilde{X} + w1^T$$

Then expand the objective function:

$$\begin{aligned} \|Y - aQ(X - z1^T)\|_F^2 &= \|\tilde{Y}\|_F^2 + a^2\|\tilde{X}\|_F^2 + \|w1^T\|_F^2 - 2a \operatorname{trace}(Q\tilde{X}\tilde{Y}^T) + \\ &\quad + 2 \operatorname{trace}(w1^T\tilde{Y}) - 2a \operatorname{trace}(Q\tilde{X}1w^T) \end{aligned}$$

But:  $\tilde{X}1 = 0$  and  $\tilde{Y}1 = 0$  because of the centering. It follows:

$$\|Y - Q(X - z1^T)\|_F^2 = \|\tilde{Y}\|_F^2 + a^2\|\tilde{X}\|_F^2 + n\|w\|^2 - 2a \operatorname{trace}(QR)$$

where  $R = \tilde{X}\tilde{Y}^T$ . Then the minimum of this criterion is achieved at  $w = 0$ ,  $Q$  the orthogonal matrix that maximizes  $\operatorname{trace}(QR)$ , and  $a$  the minimizer of  $\|\tilde{X}\|_F^2 a^2 - 2a \operatorname{trace}(\Sigma)$ . The solution follows.

## References



Wikipedia:

[https://en.wikipedia.org/wiki/Orthogonal\\_Procrustes\\_problem](https://en.wikipedia.org/wiki/Orthogonal_Procrustes_problem)