Lecture 5: Alignment Problems

Radu Balan

University of Maryland

February 20, 2025

Alignment Problems

Assume we have two geometric graphs, $\mathbb{X} = \{x_1, \cdots, x_n\} \subset \mathbb{R}^d$ and $\mathbb{Y} = \{y_1, \cdots, y_n\} \subset \mathbb{R}^d$. 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

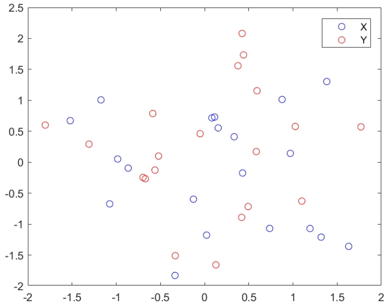
The Procrustes Problem

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:

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

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

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

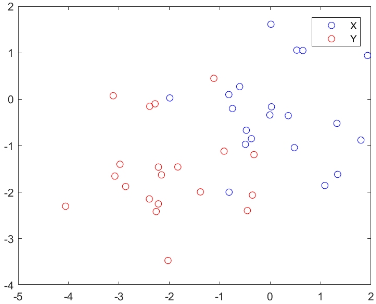


Alignment Problems

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:

$$\underset{\substack{Q\in O(d)\\ z\in\mathbb{R}^d}}{\text{minimize}} \quad \|Y - Q(X - z1^T)\|_F^2$$

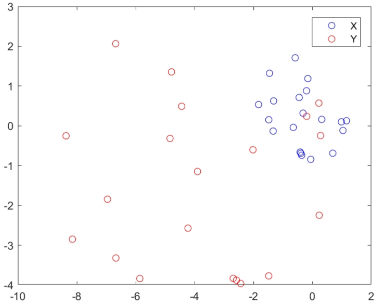


Alignment Problems

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:

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



The Optimization Problem

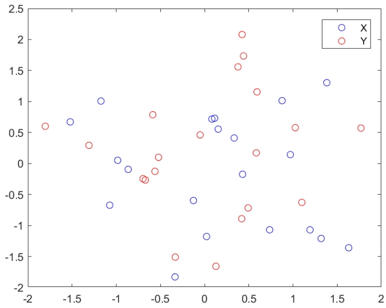
Given matrices $X, Y \in \mathbb{R}^{d \times n}$ whose columns are the n points from each set $\mathbb{X}$, $\mathbb{Y}$, find an orthogonal matrix $Q \in O(d)$ that:

$$\underset{Q \in O(d)}{\text{minimize}} \quad \|Y - QX\|_F^2$$

where

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

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



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* $\Sigma = diag(\sigma_1, \cdots, \sigma_d)$ *is the diagonal matrix with singular values* $\sigma_1, \cdots, \sigma_d \geq 0$ *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 $QQ^T = I_d$. 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 $trace(AB) = trace(BA)$.

1. We start by expanding the objective function:

$$\|Y - QX\|_F^2 = trace((Y-QX)^T(Y-QX)) = trace(Y^T Y) - trace(X^T Q^T Y) -$$

$$-trace(Y^T QX) + trace(X^T Q^T QX) = \|Y\|_F^2 - trace(Y^T QX) - trace(Y^T QX) +$$

$$+trace(X^T X) = \|Y\|_F^2 - 2trace(QXY^T) + \|X\|_F^2.$$

Derivation of the solution - 2

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

Then:

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

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

Derivation of the solution - 3

Thus (R^T, σ^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:

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

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

$$\max_{\tilde{Q} \in O(d)} \quad trace(\tilde{Q}\Sigma)$$

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

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

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 $q_{1,1} = \cdots = q_{d,d} = 1$. In this case $\tilde{Q} = I_d$ and $trace(\tilde{Q}\Sigma) = trace(\Sigma)$.

Thus $V^T Q U = I_d$ and hence $\boxed{Q = VU^T}$.

The alignment error (mismatch) is given by:

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

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

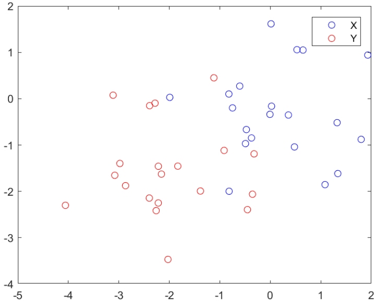
$$\|Err\|_F^2 = trace(XX^T) + trace(YY^T) - 2\, trace((YX^T XY^T)^{1/2})$$

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:

$$\begin{array}{ll} \text{minimize} & \|Y - Q(X - z1^T)\|_F^2 \\ Q \in O(d) & \\ z \in \mathbb{R}^d & \end{array}$$



The solution to the classical Procrustes problem

Algorithm (Rotation-Translation alignment)

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

1. Compute centers $\bar{x} = \frac{1}{n}X\cdot 1$, $\bar{y} = \frac{1}{n}Y\cdot 1$ and recenter data $\tilde{X} = X - \bar{x}\cdot 1^T$, $\tilde{Y} = Y - \bar{y}\cdot 1^T$.
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 $\Sigma = diag(\sigma_1, \cdots, \sigma_d)$ is the diagonal matrix with singular values $\sigma_1, \cdots, \sigma_d \geq 0$ 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:

$$\|Y - Q(X - z1^T)\|_F^2 = \|\tilde{Y}\|_F^2 + \|\tilde{X}\|_F^2 + \|w1^T\|_F^2 - 2\,trace(Q\tilde{X}\tilde{Y}^T) +$$

$$+2trace(w1^T\tilde{Y}) - 2trace(Q\tilde{X}1w^T)$$

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

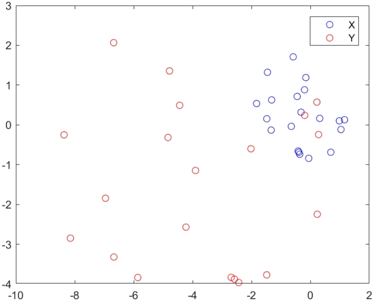
$$\|Y - Q(X - z1^T)\|_F^2 = \|\tilde{Y}\|_F^2 + \|\tilde{X}\|_F^2 + n\|w\|^2 - 2\,trace(QR)$$

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

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



The solution to the full alignment problem

Algorithm (Full alignment)

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

1. Compute centers $\bar{x} = \frac{1}{n}X\cdot 1$, $\bar{y} = \frac{1}{n}Y\cdot 1$ and recenter data $\tilde{X} = X - \bar{x}\cdot 1^T$, $\tilde{Y} = Y - \bar{y}\cdot 1^T$.
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 $\Sigma = diag(\sigma_1, \cdots, \sigma_d)$ is the diagonal matrix with singular values $\sigma_1, \cdots, \sigma_d \geq 0$ on its diagonal;
4. Compute $Q = VU^T$, $a = \frac{trace(\Sigma)}{\|\tilde{X}\|_F^2}$ 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:

$$\|Y - aQ(X - z1^T)\|_F^2 = \|\tilde{Y}\|_F^2 + a^2\|\tilde{X}\|_F^2 + \|w1^T\|_F^2 - 2a\, trace(Q\tilde{X}\tilde{Y}^T) +$$
$$+2\, trace(w1^T\tilde{Y}) - 2a\, trace(Q\tilde{X}1w^T)$$

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

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

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

References

Wikipedia:
https://en.wikipedia.org/wiki/Orthogonal_Procrustes_problem