4

Shape Space and Distances

In this chapter we investigate further geometrical aspects of shape. This chapter and Chapter 5 extend and formalize the material from Chapter 3. We shall consider a briefer version of the material than can be found in the book Dryden and Mardia (1998).

4.1 Shape Space

4.1.1 Introduction

We have already noted that the shape of an object is given by the geometrical information that remains when we filter out translation, rotation and scale information.

A rotation of a configuration is given by post-multiplication of the configuration matrix X by a rotation matrix Γ.

Definition 4.1 *An* $m \times m$ **rotation matrix** *satisfies* $\Gamma^{\mathrm{T}}\Gamma = \Gamma\Gamma^{\mathrm{T}} = I_m$ *and* $|\Gamma| = +1$. *The set of all* $m \times m$ *rotation matrices is known as the special orthogonal group* $SO(m)$.

A translation is obtained by adding a constant m-vector to the coordinates of each point. An isotropic scaling is obtained by multiplying X by a positive real number.

Definition 4.2 *The* **Euclidean similarity transformations** *of a configuration matrix* X *are the set of translated, rotated and isotropically rescaled* X, *i.e.*

$$\{\beta X\Gamma + 1_k\gamma^{\mathrm{T}} : \beta \in \mathbb{R}^+, \Gamma \in SO(m), \gamma \in \mathbb{R}^m\}, \quad (4.1)$$

where $\beta \in \mathbb{R}^+$ *is the scale,* Γ *is a rotation matrix and* γ *is* *a translation m-vector.*

Definition 4.3 *The* **rigid-body transformations** *of a configuration matrix X are the set of translated and rotated X, i.e.*

$$\{X\Gamma + 1_k\gamma^{\mathrm{T}} : \Gamma \in SO(m), \gamma \in \mathbb{R}^m\}, \qquad (4.2)$$

where Γ is a rotation matrix and γ is a translation m-vector.

For $m = 2$ we can use complex notation as in Chapter 3. Consider $k \geq 3$ landmarks in $\mathbb{C}$, $z^o = (z_1^o, \ldots, z_k^o)^{\mathrm{T}}$ which are not all coincident. The Euclidean similarity transformations of z^o are

$$\{\eta z^o + 1_k\xi : \eta = \beta e^{i\theta} \in \mathbb{C}, \xi \in \mathbb{C}\},$$

where $\beta \in \mathbb{R}^+$ is the scale, $0 \leq \theta < 2\pi$ is the rotation angle and $\xi \in \mathbb{C}$ is the translation. Hence, the Euclidean similarity transformations of z^o are the set of the same complex linear transformations applied to each landmark z_j^o. Specifying the Euclidean similarity transformations as complex linear transformations leads to great simplifications in shape analysis for the two dimensional case, as we have seen in the previous chapter.

We could consider the shape of X as the equivalence class of the full set of similarity transformations of a configuration. Alternatively we could filter out the similarity transformations from the configuration in a systematic manner. We shall adopt the latter approach.

If all k points are coincident, then this has a special shape that must be considered as a separate case. The coincident case is not generally of interest.

4.1.2 Filtering translation

In order to represent shape it can be convenient to remove the similarity transformations one at a time. Translation is the easiest to filter from X and can be achieved by considering contrasts of the data, i.e. pre-multiplying by a suitable matrix. We can make a specific choice of contrast by pre-multiplying X with the Helmert sub-matrix of Equation (2.9).

We write

$$X_H = HX \in \mathbb{R}^{(k-1)m} \setminus \{0\} \tag{4.3}$$

(the origin is removed because coincident landmarks are not allowed) and we refer to X_H as the **Helmertized landmarks**.

The centred landmarks are an alternative choice for removing location and are given by

$$X_C = CX. \tag{4.4}$$

We can revert back to the centred landmarks from the Helmertized landmarks by pre-multiplying by H^{T}, as

$$H^{\mathrm{T}}H = I_k - \frac{1}{k}1_k 1_k^{\mathrm{T}} = C$$

and so

$$H^{\mathrm{T}}X_H = H^{\mathrm{T}}HX = CX.$$

4.1.3 Pre-shape

We saw in Section 3.2 that in computing a distance between shapes it is necessary to standardize for size. We standardize for size by dividing through by our notion of size. We choose the centroid size (see Equation (2.2)) which is also given by

$$\|X_H\| = \sqrt{\text{trace}(X^{\mathrm{T}} H^{\mathrm{T}} H X)} = \sqrt{\text{trace}(X^{\mathrm{T}} C X)} = \|CX\| = S(X), \tag{4.5}$$

since $H^{\mathrm{T}} H = C$ is idempotent. Note that $S(X) > 0$ because we do not allow complete coincidence of landmarks. The pre-shape of a configuration matrix X has all information about location and scale removed.

Definition 4.4 *The* **pre-shape** *of a configuration matrix* X *is given by*

$$Z = \frac{X_H}{\|X_H\|} = \frac{HX}{\|HX\|} \tag{4.6}$$

which is invariant under the translation and scaling of the original configuration.

An alternative representation of pre-shape is to initially centre the configuration and then divide by size. The

centred pre-shape is given by

$$Z_C = CX/\|CX\| = H^{\mathrm{T}}Z \qquad (4.7)$$

since $C = H^{\mathrm{T}}H$. Note that Z is a $(k-1) \times m$ matrix whereas Z_C is a $k \times m$ matrix.

Important point: Both pre-shape representations are equally suitable for the pre-shape space which has real dimension $(k-1)m-1$. The advantage in using Z is that it is of full rank and the dimension is less than that of Z_C (although of course they have the same rank). On the other hand, the advantage of working with the centred pre-shape Z_C is that a plot of the Cartesian coordinates gives a correct geometrical view of the shape of the original configuration.

Definition 4.5 *The* **pre-shape space** *is the space of all possible pre-shapes. Formally, the pre-shape space* S_m^k

is the orbit space of the non-coincident k point set configurations in $\mathbb{R}^m$ under the action of translation and isotropic scaling.

The pre-shape space $S_m^k \equiv S^{(k-1)m-1}$ is a hypersphere of unit radius in $(k-1)m$ real dimensions, since $\|Z\| = 1$. The term 'pre-shape' signifies that we are one step away from shape – rotation still has to be removed. The term was coined by Kendall (1984).

4.1.4 Shape

In order to also remove rotation information from the configuration we identify all rotated versions of the pre-shape with each other, and this set or equivalence class is the shape of X. An alternative definition of the shape of X is

Definition 4.6 *The* **shape** *of a configuration matrix X is all the geometrical information about X that is invariant under location, rotation and isotropic scaling (Euclidean similarity transformations). The shape can be represented by the set $[X]$ given by*

$$[X] = \{Z\Gamma : \Gamma \in SO(m)\}, \tag{4.8}$$

where $SO(m)$ is the special orthogonal group of rotations and Z is the pre-shape of X.

Definition 4.7 *The* **shape space** *is the set of all possible shapes. Formally, the shape space Σ_m^k is the orbit space of the non-coincident k point set configurations in $\mathbb{R}^m$ under the action of the Euclidean similarity transformations.*

Important point: The dimension of the shape space is

$$M = km - m - 1 - \frac{m(m-1)}{2},$$

and this can be simply seen as we initially have km coordinates and then must lose m dimensions for location, one dimension for uniform scale and $\frac{1}{2}m(m-1)$ for rotation.

The shape of X is a set – an equivalence class under the action of the group of similarity transformations. In order to visualize shapes it is often convenient to choose a particular member of the shape set $[X]$.

Definition 4.8 *An* **icon** *is a particular member of the shape set* $[X]$ *which is taken as being representative of the shape.*

The word icon means ‘image or likeness’ and it is appropriate as we use the icon to picture a representative figure from the shape equivalence class which has ‘likeness’ to the other members (i.e. the objects of the class are all similar). The term was first used by Goodall (1995).

The centred pre-shape Z_C is a suitable choice of icon.

4.1.5 Size-and-shape: Removing location and rotation

We could change the order of quotienting out the similarity transformations or only remove some of the transformations. For example, if location and rotation are removed but not scale, then we have the size-and-shape of X.

Definition 4.9 *The* **size-and-shape** *of a configuration matrix X is all the geometrical information about X that is invariant under location and rotation (rigid-body transformations), and this can be represented by the set $[X]_S$ given by*

$$[X]_S = \{X_H \Gamma : \Gamma \in SO(m)\}, \tag{4.9}$$

where X_H are the Helmertized coordinates of Equation

(4.3). The space of all size-and-shapes is called the **size-and-shape space** *and is denoted by* $S\Sigma_m^k$*, for* k *points in* m *dimensions. The size-and-shape space is the orbit space of the configuration space under the action of translation and rotation.*

Size-and-shape has also been called the **form**. We discuss size-and-shape in more detail in Chapter 8.

If size is removed from the size-and-shape (e.g. by rescaling to unit centroid size), then we obtain the shape of X,

$$[X] = [X]_S / S(X) = \{Z\Gamma : \Gamma \in SO(m)\},$$

as in Equation (4.8).

4.1.6 Reflection shape

We can also include invariances under reflections for shape or size-and-shape.

Definition 4.10 *The* **reflection shape** *of a configuration matrix X is all the geometrical information that is invariant under the similarity transformations and reflection. The reflection shape can be represented by the set*

$$[X]_R = \{ZR : R \in O(m)\}$$

where $O(m)$ is the set of $m \times m$ orthogonal matrices, satisfying $R^{\mathrm{T}}R = I_m = RR^{\mathrm{T}}$ and $|R| = \pm 1$, and Z is the pre-shape.

Definition 4.11 *The* **reflection size-and-shape** *of a configuration matrix X is all the geometrical information that is invariant under translation, rotation and reflection. The* *reflection size-and-shape can be represented by the set*

$$[X]_{RS} = \{X_H R : R \in O(m)\}$$

where $O(m)$ *is the set of* $m \times m$ *orthogonal matrices and* X_H *are the Helmertized coordinates.*

Important point: With quite a wide variety of terminology used for the different spaces it may be helpful to refer to Figure 38 where we give a diagram indicating the hierarchies of the different spaces.

4.2 Distances

4.2.1 Procrustes distances

A concept of distance between two shapes is required to fully define the non-Euclidean shape metric space. We shall primarily concentrate on the full Procrustes distance, which was introduced for the two dimensional case in

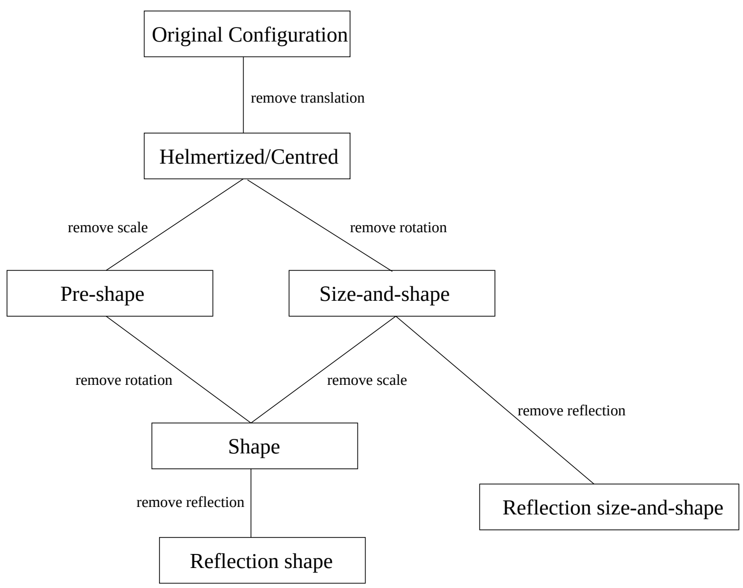

Figure 38 The hierarchies of the various spaces (after Goodall and Mardia, 1992).

Section 3.2.

Consider two configuration matrices from k points in m dimensions X_1 and X_2 with pre-shapes Z_1 and Z_2. We minimize over rotations and scale to find the closest Euclidean distance between Z_1 and Z_2.

Definition 4.12 *The* **full Procrustes distance** *between* X_1 *and* X_2 *is*

$$d_F(X_1, X_2) = \inf_{\Gamma \in SO(m), \beta \in \mathbb{R}} \|Z_2 - \beta Z_1 \Gamma\|, \qquad (4.10)$$

where $Z_r = HX_r / \|HX_r\|, \ r = 1, 2.$

Result 4.1 *The full Procrustes distance is*

$$d_F(X_1, X_2) = \left\{ 1 - \left(\sum_{i=1}^{m} \lambda_i \right)^2 \right\}^{1/2}, \qquad (4.11)$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{m-1} \geq |\lambda_m|$ *are the square roots of the eigenvalues of* $Z_1^{\mathrm{T}} Z_2 Z_2^{\mathrm{T}} Z_1$*, and the smallest* *value* λ_m *is the negative square root iff* $\det(Z_1^{\mathrm{T}} Z_2) < 0$. *The minimizing rotation is given by*

$$\hat{\Gamma} = UV^{\mathrm{T}}, \tag{4.12}$$

where $U, V \in SO(m)$ *and* $Z_2^{\mathrm{T}} Z_1 = V \Lambda U^{\mathrm{T}}$ *with* $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$. *The minimizing scale is*

$$\hat{\beta} = \sum_{i=1}^{m} \lambda_i.$$

Proof:

See Dryden and Mardia (1998, p62).

□

We shall primarily concentrate on using the full Procrustes distance in the shape space, because this is a statistically natural measure of shape distance (see Section

3.2). Note that

$$0 \leq \sum_{i=1}^{m} \lambda_i \leq 1$$

and so

$$0 \leq d_F \leq 1.$$

4.2.2 *Alternative distances*

Alternative distances in shape space could be suggested. In Figure 39 we see a diagrammatic view of the pre-shape sphere. Since the pre-shape sphere is a hypersphere embedded in $\mathbb{R}^{(k-1)m}$ we could consider familiar distances between two points on a sphere, such as the great circle distance or the chordal (Euclidean) distance. Since the shapes of configurations are represented by fibres on the pre-shape sphere, we can define the distance between two shapes as the closest distance between the fibres on the pre-shape sphere. In Figure 39 two minimum distances have been drawn between the fibres (shapes), ρ is the closest great circle distance and d_P the closest chordal distance.

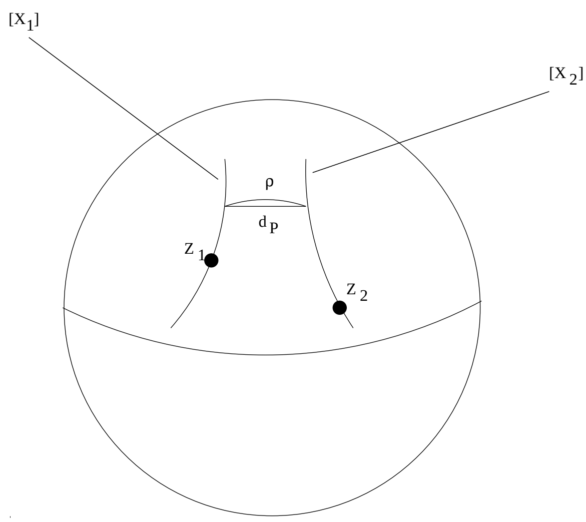

Figure 39 A diagrammatic simplistic view of two fibres $[X_1]$ and $[X_2]$ on the pre-shape sphere, which correspond to the shapes of the original configuration matrices X_1 and X_2 which have pre-shapes Z_1 and Z_2. Also displayed are the smallest great circle ρ and chordal distances d_P between the fibres.

Definition 4.13 *The* **partial Procrustes distance** d_P *is obtained by matching the pre-shapes* Z_1 *and* Z_2 *of* X_1 *and*

X_2 *as closely as possible over rotations, but not scale. So,*

$$d_P(X_1, X_2) = \inf_{\Gamma \in SO(m)} \|Z_2 - Z_1\Gamma\|,$$

where $Z_j = HX_j/\|HX_j\|, \ j = 1, 2.$

Result 4.2 *The partial Procrustes distance is given by*

$$d_P(X_1, X_2) = \sqrt{2}\left(1 - \sum_{i=1}^{m} \lambda_i\right)^{1/2}. \quad (4.13)$$

Proof: By keeping $\beta = 1$ fixed throughout the proof of Result 4.1, and just minimizing over Γ. □

Note the optimal rotation is the same whether or not scaling is in the minimization.

Definition 4.14 *The* **Procrustes distance** *or* **Riemannian metric** $\rho(X_1, X_2)$ *is the closest great circle distance* *between* Z_1 *and* Z_2 *on the pre-shape sphere, where* $Z_j = HX_j/\|HX_j\|,\ j = 1, 2$*. The minimization is carried out over rotations.*

From trigonometry one can see that the Procrustes distance is

$$\rho(X_1, X_2) = 2\text{arcsin}(d_P(X_1, X_2)/2) = \arccos\left(\sum_{i=1}^{m} \lambda_i\right). \tag{4.14}$$

In Figure 40 we see a cross-section of the pre-shape sphere illustrating the relationships between d_F, d_P and ρ. Indeed

$$\begin{aligned} d_F(X_1, X_2) &= \sin \rho, \\ d_P(X_1, X_2) &= 2\sin(\rho/2). \end{aligned}$$

Important point: Note that ρ can be considered as the smallest angle between the complex vectors Z_1 and Z_2 over rotations of Z_1 and Z_2.

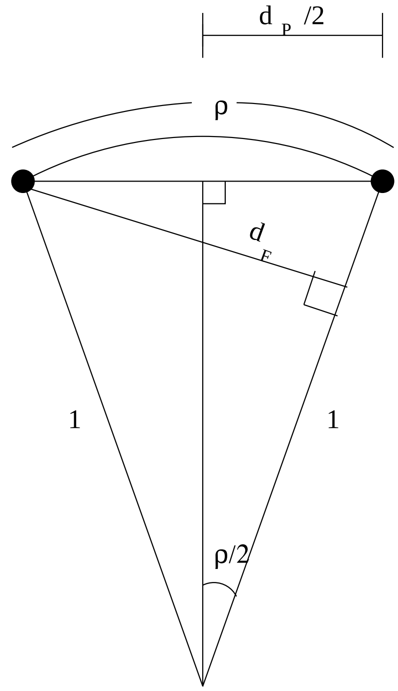

Figure 40 Section of the pre-shape sphere, illustrating the relationship between the Procrustes distances d_F, d_P and ρ.

Important point: For shapes which are close together there is very little difference between the shape distances, since

$$d_P = d_F + O(d_F^3) \ , \ \rho = d_F + O(d_F^3).$$

Consequently for many practical datasets with small variability there is very little difference in the analyses when using different Procrustes distances. However, the distinction between the distances is worth making and the terminology is summarized in Table 1.

Distance	Notation	Formula	Range
Full Procrustes distance	d_F	$\{1 - (\sum_{i=1}^m \lambda_i)^2\}^{1/2}$	$0 \leq d_F \leq 1$
Partial Procrustes distance	d_P	$\sqrt{2}(1 - \sum_{i=1}^m \lambda_i)^{1/2}$	$0 \leq d_P \leq \sqrt{2}$
Procrustes distance	ρ	$\arccos(\sum_{i=1}^m \lambda_i)$	$0 \leq \rho \leq \pi/2$

Table 1 Procrustes distances in the shape space.