1

Introduction

Digital images today play a vital role in science and technology, and also in many aspects of our daily life. This book seeks to advance the analysis of images, especially digitized ones, through the statistical analysis of shapes. Its focus is on the analysis of landmark-based shapes in which a k-ad, that is, a set of k labeled points or landmarks on an object or a scene, is observed in two or three dimensions, usually with expert help, for purposes of identification, discrimination, and diagnostics.

In general, consider the k-ad to lie in $\mathbb{R}^m$ (usually, $m = 2$ or 3) and assume that not all the k points are the same. Then the appropriate shape of the object is taken to be the k-ad modulo a group of transformations.

For example, one may first center the k-ad, by subtracting the mean of the k-ad from each of the k landmarks, to remove the effect of location. The centered k-ad then lies in a hyperplane of dimension $mk - m$, because the sum of each of the m coordinates of the centered k points is zero. Next one may scale the centered k-ad to unit size to remove the effect of scale or size. The scaled, centered k-ad now lies on the unit sphere $S^{m(k-1)-1}$ in a Euclidean space (the hyperplane) of dimension $m(k-1)$ and is now called the *preshape* of the k-ad. Further, to remove the effect of orientation, the scaled, centered k-ad is rotated by means of elements of the (special orthogonal) group $\mathrm{SO}(m)$ of rotations in $\mathbb{R}^m$. The orbit of the preshape under all rotations may then be taken to be the shape of the k-ad. This shape is called a *similarity shape*, and the space of these shapes comprises Kendall's similarity shape space Σ_m^k. While this is a proper choice for many problems in biology and in medical imaging, other notions of shape, such as affine shape and projective shape, are important in machine vision and bioinformatics. The *affine shape* of a k-ad is invariant under all affine transformations; that is, it may be identified as the orbit of the k-ad under (the group of) all affine transformations. This is an appropriate notion of shape of a k-ad based on images taken from far away, for example, from an airplane or a satellite. Here a rectangle may be transformed to a parallelogram. Similarly, *projective shapes* are invariant under

all projective transformations, and these are the appropriate shapes of k-ads recorded, for example, by central projection by a camera, where a line in three dimensions is projected to a point on the image plane, and a three-dimensional object appears as a two-dimensional image. These shape spaces are differentiable manifolds (sometimes after the removal of a small singular set), often with a Riemannian structure allowing one to measure (geodesic) lengths, angles, curvature, and so on. These notions are important in intrinsic analysis as mentioned briefly below and explained in detail in Chapter 5.

For nonparametric analysis of shape distributions Q on a manifold M with a distance d, we focus on the *Fréchet mean* of Q, which is the minimizer (if unique) of the *Fréchet function*, namely, the average squared distance from a point. Sometimes the corresponding minimum average is also considered, called the *variation* of Q. If the distance is the geodesic distance, these parameters, and the corresponding statistical analysis, are said to be *intrinsic*. Often it is more convenient, mathematically as well as from a computational point of view, to embed the manifold in a Euclidean space E (generally of higher dimension than that of M) and use the induced (Euclidean) distance on the image under the embedding j, say. The Fréchet mean and variation, and the statistical analysis based on them, are then said to be *extrinsic*. For $M = S^d$, a d-dimensional unit sphere, the intrinsic, or geodesic, distance between two points is the arc length between them measured on the great circle joining the points. This distance is sometimes referred to as the arc distance. Also, the sphere has a natural embedding into $\mathbb{R}^{d+1}$ by the *inclusion map* j. That is, if we represent S^d as $\{|x| = 1 : x = (x_1, \ldots, x_{d+1}) \in \mathbb{R}^{d+1}\}$, then $j(x) = x$ for $x \in S^d$. The extrinsic distance between two points is the length of the line segment joining the two points, the so-called chord distance.

Suppose one has a sample of k-ads of size n. This yields a sample of n shapes in the appropriate shape space (manifold) M. The common distribution of these n shapes is denoted by Q. The statistical analysis begins by first (1) finding broad conditions for uniqueness of the Fréchet minimizer, and then (2) finding the asymptotic distribution of the corresponding sample Fréchet mean as its estimate. Similarly one estimates the variation. There is no general simplifying procedure for the construction of the intrinsic mean. However, for the extrinsic mean of Q under an embedding j, one first computes the mean, say μ, of Q as a distribution on the ambient Euclidean space $E = \mathbb{R}^D$. The extrinsic mean of Q on the image $j(M)$ of M is given by the point in $j(M)$, if unique, which is at the minimum Euclidean distance from μ, denoted by $P(\mu)$, with P denoting the *projection* operation.

For the statistical problems at hand, consider the particular case of distinguishing between two shape distributions on a manifold. To understand the general nature of a two-sample test for equality of the extrinsic means of two distributions Q_1, Q_2 on M, based on independent samples $\{X_{j_i} : j = 1, \ldots, n_i\}$ of sizes n_i $(i = 1, 2)$, observe that

$$\sqrt{n_i}\,[P(\hat{\mu}_i) - P(\mu_i)] = \sqrt{n_i}\,[d_{\mu_i}P(\hat{\mu}_i - \mu_i)] + o_p(1) \quad (i = 1, 2), \tag{1.1}$$

where P is the projection on $j(M)$ and $\mu_i, \hat{\mu}_i$ are the Euclidean population and sample means of the ith group $(i = 1, 2)$, under the embedding j (i.e., in the ambient Euclidean space $\mathbb{R}^D$). Here $d_{\mu_i}P$ is the differential of the projection map P at μ_i, which is a linear map from the tangent space of $\mathbb{R}^D$ at μ_i, namely, $T_{\mu_i}\mathbb{R}^D = \mathbb{R}^D$, to the tangent space of $\tilde{M}$ at $P(\mu_i)$, that is, $T_{P(\mu_i)}\tilde{M} \subset T_{\mu}\mathbb{R}^D$ (here $\tilde{M} := j(M)$ is the image of M under the embedding j). Note that the Jacobian of $d_{\mu_i}P$ is a singular $D \times D$ matrix whose rank is the same as the dimension d of $\tilde{M}$ (or M). Now, for each i, viewed as a random element of $\mathbb{R}^D$, equation (1.1) converges to a (singular) Normal distribution $N(0, \Sigma_i)$ in $\mathbb{R}^D$, where Σ_i is a $D \times D$ covariance matrix. Under the null hypothesis H_0 of equality of the two extrinsic means, $P(\mu_1) = P(\mu_2) = j(\mu_E) = \tilde{\mu}_E$, say, the tangent spaces $T_{P(\mu_i)}\tilde{M}, i = 1, 2$, are the same. Let $n = n_1 + n_2$ and assume $\frac{n_1}{n} \to p$, $0 < p < 1$. Taking the difference between the two quantities in equation (1.1), one has the convergence in distribution $\sqrt{n}[P(\hat{\mu}_1) - P(\hat{\mu}_2)] \to N(0, p^{-1}\Sigma_1 + (1-p)^{-1}\Sigma_2)$. But the right side of equation (1.1), excluding the term $o_p(1)$, lies in the tangent space $T_{\tilde{\mu}_E}\tilde{M}$ of dimension d and converges to a d-dimensional Normal $N(0, \Gamma)$, where $\Gamma = p^{-1}\Sigma_1 + (1-p)^{-1}\Sigma_2$, with Σ_i being the covariance matrix of the coordinates of $d_{\mu_i}(\tilde{X}_{j_i} - \mu_i)$ $(i = 1, 2)$ with respect to the basis of the tangent space $T_{\tilde{\mu}_E}\tilde{M}$ of the embedded manifold $\tilde{M}$ at the common extrinsic mean $\tilde{\mu}_E$ (under H_0). Here $\tilde{X} = j(X)$. This leads to a chi-squared test. As usual, one replaces Σ_i by the sample estimate $\hat{\Sigma}_i$, obtained as the sample covariance of the coordinates of $d_{\hat{\mu}_i}(\tilde{X}_{j_i} - \hat{\mu}_i)$ $(j = 1, \ldots, n_i)$. For intrinsic analysis, the computation of the intrinsic sample mean begins with a theoretical derivation of the geodesics as well as the Fréchet minimization involving the geodesic distance. Analogous to the extrinsic embedding, one then transfers the population (and sample) distributions to the tangent space at the intrinsic mean by the so-called inverse exponential map described in Chapter 5. The rest of the procedure is similar to that for the extrinsic mean.

It is useful to remember that the larger the group of transformations applied to the k-ads, the larger the orbit under it defining the shape of a k-ad, and the fewer are the details of the numerics of the k-ads preserved in their shapes. In particular, statistical significance (at a given level of significance) in a two-sample test based on a notion of shape invariant under

a larger group is in general a stronger statement than that based on shape invariant under smaller groups. In this context one should note the increasing order of groups of transformations defining Kendall's similarity shapes, reflection similarity shapes, affine shapes, and projective shapes.

Here is a brief outline of the book.

To motivate the reader, Chapter 2 provides an exposition of several data examples, which are analyzed in detail in later chapters.

In Chapter 3, the concepts of Fréchet mean and variation are introduced. The idea is to define the Fréchet function of a probability distribution Q on a metric space M as the integral (with respect to Q) of the squared distance on M, and define the Fréchet mean as the minimizer (if unique) of this Fréchet function and the Fréchet variation as the minimum value (if finite). Conditions are derived for the consistency and asymptotic Normality of the sample estimates of the Fréchet mean. Confidence regions for the population parameters are constructed both by using the asymptotic distribution of the sample estimates and by pivotal bootstrap methods.

Chapter 4 is devoted to extrinsic inference on a differentiable manifold M. Here one embeds M into some higher dimensional Euclidean space and uses the distance induced by this embedding. Because many embeddings are available, one chooses an embedding that is equivariant with respect to a group of transformations large enough to preserve a great deal of the geometry of M. As the results in later chapters show, the corresponding analysis becomes simpler both mathematically and computationally than its intrinsic counterpart. For example, the extrinsic mean is known to exist under fairly broad conditions and in most cases has a closed-form analytic expression (see Chapters 8–12). In particular, the extrinsic mean is the projection of the Euclidean mean of the image of Q on the image manifold under the embedding. Hence there is a unique mean if and only if there is a unique projection. For asymptotic Normality of the sample mean and variation, one requires that this projection map is smooth in a neighborhood of the population mean (which is assumed to exist). The chapter concludes with two-sample nonparametric tests to distinguish between two probability distributions by comparing the sample extrinsic means and variations. Appropriate tests are constructed for both mutually independent and matched pair samples. The numerical examples in Chapter 8 show that in these examples the extrinsic and intrinsic means are very close to each other and the two-sample extrinsic and intrinsic tests yield similar results.

Chapter 5 performs Fréchet analysis on a Riemannian manifold M by using the geodesic distance as the distance metric in the definition of the Fréchet function. The resulting Fréchet parameters are called intrinsic

and the corresponding statistical analysis is called intrinsic analysis on a manifold M. In this chapter, sufficient conditions for existence of a unique intrinsic mean are used to derive the asymptotic Normality of the sample intrinsic mean. Two-sample nonparametric tests are constructed to compare the sample intrinsic means and variations, which can be used to distinguish between two underlying distributions.

Chapter 6 introduces the different notions of shapes treated in this book. They include (direct) similarity shapes, reflection similarity shapes, affine shapes, and projective shapes. For problems in biology such as classifications of species, disease detection, and so on, similarity shape analysis has many uses, while for problems in machine vision and image analysis, affine and projective shape analyses are more appropriate.

In Chapters 7–12, the geometry of each of the shape spaces introduced in Chapter 6 is discussed in detail, and explicit forms of estimates and tests are derived in each case using the methods of Chapters 4 and 5.

In particular, Chapter 7 provides an exposition of the geometry of the (direct) *similarity shape space of* k-ads in m dimensions, or Σ_m^k. The cases of interest include $m = 2$ and 3. This space can be represented as the quotient of the unit sphere with respect to all rotations (in m dimensions), that is, the space of orbits of k-ads under rotations of the preshape sphere, as described at the outset. It is shown that, after removing some singularities, Σ_m^k is a Riemannian manifold. There are no such singularities when $m = 2$. This chapter identifies the tangent space of Σ_m^k, the exponential map, and the geodesic distance on Σ_m^k.

Chapter 8 considers in detail the similarity shape space Σ_2^k obtained when $m = 2$, which is also called the *planar shape space*. This is a compact, connected manifold. This chapter presents the geometry of this space and applies the methods of Chapters 4 and 5 for intrinsic and extrinsic analyses. Analytic expressions for the parameters in the asymptotic distribution of the sample extrinsic mean are derived. This enables one to perform two-sample tests to compare the extrinsic means and variations of two underlying probability distributions. The results of extrinsic and intrinsic analyses on Σ_2^k are applied to two examples.

When $m > 2$, the similarity shape space Σ_m^k fails to be a manifold. After singularities are excluded, the remaining set is a manifold that is not complete. As a consequence, the results from Chapters 4 and 5 cannot be applied to carry out intrinsic and extrinsic analyses. If, instead, one considers the *reflection similarity shape*, which is invariant under all orthogonal transformations, not just rotations, then one can embed the resulting shape space into the vector (or Euclidean) space of symmetric matrices $S(k, \mathbb{R})$ and carry out extrinsic analysis. This is discussed in Chapter 9.

The methods here allow one to extend nonparametric inference on Kendall-type shape spaces from two to higher dimensions. This chapter concludes with an application to a matched pair example.

Chapter 11 focuses on methodologies for nonparametric inference on the affine shape spaces $A\Sigma_m^k$. The *affine shape* of a k-ad x with landmarks in $\mathbb{R}^m$ is defined as the orbit of x under all affine transformations, $x \mapsto Ax + b$ ($A \in \mathrm{GL}(m, \mathbb{R}), b \in \mathbb{R}^m$). The space of affine shapes of all centered k-ads whose columns span $\mathbb{R}^m$ is $A\Sigma_m^k$. Here $\mathrm{GL}(m, \mathbb{R})$ is the so-called general linear space of all $m \times m$ nonsingular real matrices. For extrinsic analysis on $A\Sigma_m^k$, one embeds it into the vector space $S(k, \mathbb{R})$ of all $k \times k$ real symmetric matrices via an equivariant embedding. Using this embedding, an expression for the extrinsic mean and a condition for its uniqueness are derived. The results from Chapter 4 are used to derive the asymptotic distribution for the sample extrinsic mean and variation, and are applied to construct two-sample nonparametric tests to compare two probability distributions.

Chapter 12 presents methodologies for statistical analyses of projective shapes, which are useful in axial analysis and machine vision. For $m = 2$, for example, the projective space $\mathbb{R}P^2$ is the space of all lines in $\mathbb{R}^3$ passing through the origin. The image of a line, or axis, passing through the center of a pin-hole–type camera is recorded as a point on the plane of the camera film. Thus a scene in three dimensions is pictured on the plane of the camera film. A set of k distinct lines originating from the three-dimensional scene then yields a k-ad in $\mathbb{R}P^2$, that is, a point in $(\mathbb{R}P^2)^k$. To define the projective shape of the k-ad, one applies affine transformations $A \in \mathrm{GL}(3, \mathbb{R})$ to the k points in three dimensions. The equivalence class of lines generated from these (or the corresponding equivalence class of points on the camera film) is the *projective shape* of the k-ad.

The last two chapters, Chapters 13 and 14, represent a different theme than the rest of the book. Here we consider functional inference – (a) density estimation, (b) classification and (c) regression – by applying the nonparametric Bayes methodology.

It may be noted that there now exists a substantial literature, mostly in computer science, but also in statistics, on what may be termed continuous shapes. The greater part of this work is in two dimensions, where such shapes may be analyzed by a deformable template of gray levels on a grid of points representing a digitized approximation of the image. Here one uses high-dimensional (parametric) Bayes methodology (see Amit, 2002). The other alternative that has gained popularity in recent years is to consider a two-dimensional shape as given by the actual boundary contour of the object of interest. There is also some work on the three-dimensional shape provided by the boundary surface. Apart from

geometric considerations, much of this work focuses on analytical and computational problems involving matching (or discriminating among) given shapes, or changes in shapes, and not on statistical inference for discrimination among distributions of shapes (see Krim and Yezzi, 2006). For a nonparametric analysis of the latter, one views these spaces of shapes as infinite-dimensional (Hilbert) manifolds. Intrinsic inference here is limited by the fact that establishing the uniqueness of such basic quantities as the intrinsic mean even under restrictive conditions is difficult. In contrast, extrinsic analysis based on equivariant embeddings in vector (Hilbert) spaces holds some promise (see, e.g., the recent work of Ellingson et al., 2011).

Finally, a note on the use of bootstrapping in the text. Efron's *bootstrap* (Efron, 1979) has had a profound impact on modern statistical methodology. Apart from the fulfillment of its original intended goals such as estimating standard errors of rather complicated statistics and avoiding such computations altogether in providing confidence regions, for continuous data a great benefit of the method lies in its substantial edge over central limit theorem (CLT)–based confidence regions in reducing coverage errors. This aspect of the bootstrap's efficacy is established by the method of asymptotic expansions of distributions of smooth statistics (Bhattacharya, 1977; Bhattacharya and Ghosh, 1978) and their application to bootstrapped versions of such statistics. For this theory, see Singh (1981), Babu and Singh (1984), Bhattacharya (1987), Beran (1987), Bhattacharya and Qumsiyeh (1989), Bhattacharya and Denker (1990), and Lahiri (1994). It follows from this, in particular, that a confidence region based on pivotal bootstrapping of an asymptotic chi-squared statistic yields a coverage error of order $O(n^{-2})$, as opposed to the order $O(n^{-1})$ error resulting from the classical chi-squared approximation. For a readable account of the bootstrap's coverage error for symmetric confidence regions and various other matters we refer to Hall (1992). Unfortunately, for the generally high-dimensional shape spaces considered in this monograph, the bootstrapped covariance tends to be singular if the sample size is not very large, thus limiting the usefulness of the bootstrap somewhat.

2

Examples

This chapter collects together, and describes in a simple manner, a number of applications of the theory presented in this book. The examples are based on real data, and, where possible, results of parametric inference in the literature are cited for comparison with the new nonparametric inference theory.

2.1 Data example on S^1: wind and ozone

The wind direction and ozone concentration were observed at a weather station for 19 days. Table 2.1 shows the wind directions in degrees. The data are taken from Johnson and Wehrly (1977) . The data viewed on the unit circle S^1 are plotted in Figure 3.1. We compute the sample extrinsic and intrinsic mean directions, which come out to be 16.71 and 5.68 degrees, respectively. They are displayed in the figure. We use angular coordinates for the data in degrees lying between $[0^\circ, 360^\circ)$ as in Table 2.1. An asymptotic 95% confidence region for the intrinsic mean as obtained in Section 3.7, Chapter 3, turns out to be

$$\{(\cos\theta, \sin\theta) : -0.434 \le \theta \le 0.6324\}.$$

The corresponding end points of this arc are also displayed in Figure 3.1.

Johnson and Wehrly (1977) computed the so-called angular–linear correlation $\rho_{AL} = \max_\alpha\{\rho(\cos(\theta - \alpha), X)\}$, where X is the ozone concentration when the direction of wind is θ. Here ρ denotes the true coefficient of correlation. Based on the sample counterpart r_{AL}, the 95% confidence interval for ρ_{AL} was found to be $(0.32, 1.00)$.

2.2 Data examples on S^2: paleomagnetism

We consider here an application of *directional statistics*, that is, statistics on the unit sphere S^d, with $d = 2$ in the present case, which has an important bearing on a fundamental issue in paleomagnetism. *Paleomagnetism*

Table 2.1 *Wind directions in degrees*

327	91	88	305	344	270	67
21	281	8	204	86	333	18
57	6	11	27	84		

Table 2.2 *Data from Fisher (1953) on remanent magnetism*

D	343.2	62.0	36.9	27.0	359.0	5.7	50.4	357.6	44.0
I	66.1	68.7	70.1	82.1	79.5	73.0	69.3	58.8	51.4

D: Declination; I: Inclination

is the field of earth science that is devoted to the study of fossil magnetism as contained in fossilized rock samples, known as remanent magnetism. It has been theorized for many years that the Earth's magnetic poles have shifted over geological time. This idea is related to the older theory of continental drift, namely, that the continents have changed their relative positions over a period of several hundred million years. If rock samples in different continents dating back to the same period exhibit different magnetic polarities, that would be a confirmation of the theory of continental drift. As pointed out by the geophysicist Irving (1964) in the preface of his book, over the years such confirmations have been achieved with the help of rigorous statistical procedures. In Chapter 4, Section 4.7, a multi-sample nonparametric test for the hypothesis of equality is provided for such purposes. In a seminal paper, Fisher (1953) used a parametric model known as the *Fisher* or *von Mises–Fisher distribution* on the sphere S^2 with a density $f(x;\mu,\tau) = c(\tau)\exp\{\tau x'\mu\}$ with respect to the uniform distribution on the sphere (see Appendix D), where μ is the true direction (given by a point on the unit sphere S^2) and $\tau > 0$ is the concentration parameter. The maximum likelihood estimate (MLE) of the true position μ, based on i.i.d. observations $X_1, \ldots, X_n$ on S^2, is given by $\overline{X}/|\overline{X}|$, assuming $\overline{X} \neq 0$. Thus the MLE is the same as the extrinsic mean of the sample (empirical) distribution on S^2, where μ is the extrinsic mean, as well as the intrinsic mean, of Fisher's distribution.

From the Icelandic lava flow of 1947–1948, nine specimens of remanent magnetism were collected. The data can be viewed as an i.i.d. sample on the manifold S^2 and can be found in Fisher (1953; the data were supplied by J. Hospers). They are displayed in Table 2.2.

The sample extrinsic mean is $\hat{\mu}_E = (.1346, .2984, .9449)$. The sample extrinsic and intrinsic means are very close, namely, at a geodesic distance

of 0.0007 from each other. This estimate of magnetic north is close to the Earth's geographic north pole (0, 0, 1).

Based on his distribution, Fisher obtained a 95% confidence region for the mean direction μ. This region may be expressed as

$$\{p \in S^2 : d_g(\hat{\mu}_E, p) \leq 0.1536\},$$

where d_g denotes the geodesic distance. Fisher's confidence region, and our asymptotic confidence region for the population extrinsic mean derived in Chapter 4, are plotted in Figure 2.1. The former confidence region nearly contains the latter and is considerably larger than it.

To study possible shifts in the positions of Earth's magnetic poles, Fisher also analyzed a second set of data, supplied by Hospers, of remanent magnetism from the early Quaternary period (between 10,000 and one million years ago). The sample estimate (MLE) from this sample of 45 observations turns out to be $\hat{\mu}_E = (.0172, -.2978, -.9545)$, which shows a near reversal of the magnetic poles between the two geological periods. The 95% confidence region for the true direction by Fisher's method is a geodesic ball of radius .1475 around the MLE. Since we were unable to access the original data from the second example in Fisher's paper,

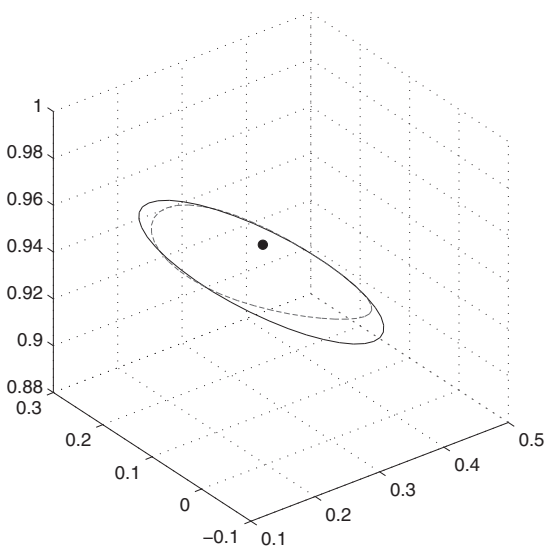

Figure 2.1 Boundaries of the confidence regions for the direction of Earth's magnetic pole, using Fisher's method (*solid*) and the nonparametric extrinsic method (*dashed*), based on data from Fisher (1953).