

1 Mean and variance

In order to perform statistics on shapes we must first try to define central statistical concepts on manifolds. In this section we focus on a geodetically complete Riemannian manifold, (M, g) , of dimension n , and present ways of defining the mean, variance and covariance of M -valued random variables. Given an underlying probability space, (Ω, \mathcal{F}, P) , an M -valued random variable is a $\mathcal{F}/\mathcal{B}(M)$ measurable map, $X : \Omega \rightarrow M$, and we denote by $x = X(\omega)$ a realization of X on M .

In order to perform statistics on M we need to construct a measure on M . This measure is induced by the metric g in the following way. Let $x = (x^1, \dots, x^n)$ be representation of $x \in M$ in local coordinates, and let $\frac{\partial}{\partial x} = (\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$ be the corresponding basis of $T_x M$. The metric g is then expressed in this basis by the matrix $G = [g_{ij}(x)]$ where $g_{ij}(x) = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \rangle = g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$. The measure on M is then defined by $dM(x) = \sqrt{|\det G(x)|} dx$ (here dx indicates regular Lebesgue-integration of the coordinate-representation of x in \mathbb{R}^n). X is said to have density p_X w.r.t. dM if

$$P(X \in \mathcal{A}) = \int_{\mathcal{A}} p_X(y) dM(y),$$

holds for all $\mathcal{A} \in \mathcal{B}(M)$ and if the integral over M is equal to 1. Here p_X is a density in the usual sense; a real-valued, positive and integrable function. If π is a chart of the manifold, then $r := \pi(X(\omega))$ defines a random vector with density, ρ_r , w.r.t to the Lebesgue measure given by

$$\rho_r(y) = p_X(\pi^{-1}(y)) \sqrt{|\det G(\pi^{-1}(y))|}.$$

If $\varphi : M \rightarrow \mathbb{R}$ is a $\mathcal{B}(M)/\mathcal{B}(\mathbb{R})$ -measurable map, then $\varphi(X)$ defines a real-valued random variable for which the expectation is

$$\mathbb{E}(\varphi(X)) = \int_M \varphi(y) p_X(y) dM(y).$$

Unfortunately, we cannot define the expectation of M -valued random variables in a similar manner, since the real-valued integral does not generalize to an integral with values on M . Instead we generalize the notion of mean value by first defining the variance of a M -valued random variable and then defining the so-called *Frechet means* as minimizers of the variance (this is just one possible way of generalizing mean points).

Definition 1. Let X be a M -valued random variable with density p_X . Given a point $y \in M$, the *variance* of X is

$$\sigma_X^2(y) = \mathbb{E}(D(y, X)^2) = \int_M D(y, z) p_X(z) dM(z).$$

Here the distance between two points, $D(x, y)$, is the infimum of the lengths of all paths in M from x to y , with the length of a path $c : [0, 1] \rightarrow M$ defined by $L(c) = \int_0^1 g_{c(t, \cdot)}(c_t, c_t) dt$.

Definition 2. Let X be a M -valued random variable with density p_X . If $\sigma_X^2(y)$ is finite for all $y \in M$, we define *Frechet mean points* of X as all points in M minimizing $\sigma_X^2(y)$;

$$\mathbb{E}(X) := \arg \min_{y \in M} \sigma_X^2(y).$$

If a mean point \bar{x} exists, the variance of X is defined by $\sigma^2(X) := \sigma_{\bar{x}}^2(X)$. We can further define the *median points* of X as all minimizers of $\mathbb{E}(\text{dist}(y, X))$.

Note 3

Given a series of measurements x_1, \dots, x_n seen as realizations of X , we define the empirical mean points of X to be the minimizers of

$$\frac{1}{n} \sum_{i=1}^n \text{dist}(y, x_i)^2,$$

and the empirical variance is, as before, defined as the variance of X evaluated at a minimizer. Note that this corresponds to the standard notion of the empirical mean as the minimizer of the sum of squares and the empirical variance as the sum of squared deviations from this mean.

The most apparent question is now whether a mean point exists for X and if it is unique. To give conditions for existence and uniqueness, we follow [?] and define *Riemannian centers of mass* as local minimizers of $\sigma_X^2(y)$. The Riemannian centers of mass have the added benefit of encoding more information about the distribution of X than the mean points, since the centers of mass represent local maxima of the distribution of X (where the mean points only represent the global maximum). The definition of these centers of mass amounts to finding local extrema of $\sigma_X^2(y)$, in which the following theorem due to [?] plays an important part.

Theorem 4. *Let X be a M -valued random variable with density p_X . If $\sigma_X^2(y) < \infty$ and the image measure of the cut locus is zero, $X(P)(C(y)) = 0$, then $\sigma_X^2(y)$ is differentiable with*

$$(\text{grad } \sigma^2)(y) = -2\mathbb{E}(\overrightarrow{y\bar{x}}) = -2 \int_{M/C(y)} \overrightarrow{y\bar{z}} p_X(z) dM(z).$$

The theorem ensures differentiability of $\sigma_X^2(y)$ in points where the cut locus has measure zero, and in this case the extrema of σ^2 are points where $(\text{grad } \sigma^2)(y) = 0$. If the cut locus has positive measure, the variance may still attain an extremum. This leads to the following characterization of Riemannian centers of mass.

Corollary 5. *Let \mathcal{A} be the set of points for which the cut locus has non-zero probability. If $\sigma_X^2(y) < \infty$ for all $y \in M$, then a necessary condition for \bar{x} to be a Riemannian center of mass is $x \in \mathcal{A}$ or $\mathbb{E}(\vec{\bar{x}x}) = 0$ for $\bar{x} \notin \mathcal{A}$.*

Since the Riemannian centers of mass are local minimizers of the variance, the set of mean points is included in the set of centers of mass. If there is only one Riemannian center of mass it therefore follows that it must be the unique mean point. The following corollary gives uniqueness of centers of mass not in \mathcal{A} for a class of manifolds (Hadamard manifolds).

Corollary 6. *Let M be a simply connected, complete manifold with non-positive Riemannian curvature, and let X be a random variable with values in M and finite variance. Then there exists one and only one Riemannian center of mass characterized by $\mathbb{E}(\vec{\bar{x}x}) = 0$. If the cut locus has measure zero everywhere, then this point \bar{x} must be a mean point.*

If we want to assert uniqueness for a larger class of manifolds, we have to make assumptions not only on the curvature of the manifold but on the support of the densities. Compact support of the densities is actually not sufficient - the support of the densities have to be contained in a *regular geodesic ball*.

Definition 7. A ball, $B(x, r) = \{y \in M \mid \text{dist}(x, y) < r\}$, is geodesic if $B(x, r) \cap C(x) = \emptyset$ and it is regular if $2r\sqrt{\kappa} < \pi$ where κ is the maximum of the Riemannian curvature in $B(x, r)$.

Note that the assumption on the curvature of M is no longer global but local. We have thus replaced the global assumption on the curvature with a local one, but have now restricted ourselves to densities with compact support on locally well-behaved parts of the manifold. Under these assumptions the following results hold ([?], [?]), where the case of $\bar{x} \in \mathcal{A}$ is not possible since the balls, on which the densities have support, are geodesic.

Theorem 8. *Let X be a M -valued random variable with density p_X . If the support of p_X is contained in a regular geodesic ball, then there exists one unique Riemannian center of mass on this ball.*

Theorem 9. *Let X be a M -valued random variable with density p_X . If the support of p_X is contained in a regular geodesic ball, and the ball with twice the radius is also a regular and geodesic, then $\sigma_X^2(y)$ is convex and has a unique critical point, \bar{x} , on the ball. This point must be a minimizer and thus the unique mean point of X .*

The preceding corollary and two theorems yield uniqueness and existence statements of mean points for specific cases. One can then wonder if it is possible to relax the global curvature assumption in the corollary or the assumptions on

the well-behaviour of the curvature in the domain of the support of the density. As shown in [?] some assumptions on the curvature is needed even if there is a unique minimizing geodesic joining any two points.

With the introduction of statistical concepts on manifolds in mind we now turn to defining manifolds of curves and equipping these with suitable metrics. These manifolds are infinite-dimensional so the methods introduced in this section cannot be directly applied. Nonetheless we aim to define length of paths in these manifolds which is the first step in trying to perform statistics on the space of shapes.