Numerical Examples in Information Geometry

Massimiliano Ferrara

1 Introduction

Information geometry applies differential geometric methods to the study of probability distributions and statistical inference. This document provides numerical examples to illustrate some concepts in information geometry.

2 Example 1: Geodesics on the Simplex

Consider the set of probability distributions over n outcomes, represented by the **probability simplex** Δ^{n-1}:

$$\Delta^{n-1} = \{(p_1, p_2, \ldots, p_n) \mid p_i \geq 0, \sum_{i=1}^{n} p_i = 1\}$$

2.1 Points in the Simplex

Let's consider two probability distributions P and Q over three outcomes. For example:

$$P = (0.2, 0.5, 0.3)\,, \quad Q = (0.6, 0.3, 0.1)$$

2.2 Riemannian Metric

On the simplex, we can define a Riemannian metric called the Fisher information metric. Given a distribution, the Fisher information $I(P)$ is given by:

$$I(P) = \operatorname{diag}\left(\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{p_3}\right)$$

2.3 Geodesic between Points

A geodesic between P and Q can be represented using the exponential map in the space defined by the Fisher information metric. The geodesic can be parameterized as:

$$\gamma(t) = P^{1-t} Q^{t}, \quad t \in [0, 1]$$

2.4 Evaluate Points on the Geodesic

Calculating some intermediate points on the geodesic: - For $t = 0.5$:

$$\gamma(0.5) = P^{0.5}Q^{0.5} = \left(0.2^{0.5} \cdot 0.6^{0.5}, 0.5^{0.5} \cdot 0.3^{0.5}, 0.3^{0.5} \cdot 0.1^{0.5}\right)$$

Calculating each component gives:

- First component: $\sqrt{0.2} \cdot \sqrt{0.6} \approx 0.348$
- Second component: $\sqrt{0.5} \cdot \sqrt{0.3} \approx 0.387$
- Third component: $\sqrt{0.3} \cdot \sqrt{0.1} \approx 0.173$

Normalizing gives a new distribution.

3 Example 2: Kullback-Leibler Divergence

The Kullback-Leibler (KL) divergence measures how one probability distribution diverges from a second, expected probability distribution. For two distributions P and Q:

$$D_{KL}(P||Q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right)$$

3.1 Example Distributions

Let's use the distributions $P = (0.4, 0.4, 0.2)$ and $Q = (0.5, 0.4, 0.1)$.

3.2 Calculate KL Divergence

$$D_{KL}(P||Q) = 0.4 \log\left(\frac{0.4}{0.5}\right) + 0.4 \log\left(\frac{0.4}{0.4}\right) + 0.2 \log\left(\frac{0.2}{0.1}\right)$$

Calculating each term separately:

- First term: $0.4 \log(0.8) \approx -0.088$
- Second term: $0.4 \cdot 0 = 0.0$
- Third term: 0