

# Monte Carlo computation of $L_p$ distance between two densities on the unit hypersphere

Kisung You

2022-08-09

## Problem Statement

A  $d$ -dimensional unit hypersphere  $\mathbb{S}^d = \{x \in \mathbb{R}^{d+1} \mid \|x\|_2^2 = \sum_{i=1}^{d+1} x_i^2 = 1\}$  is one of the standard mathematical spaces in the field of [objected-oriented data analysis](#) (Marron and Dryden 2021). Let  $\mathcal{P}(\mathbb{S}^d)$  denote a space of probability densities on  $\mathbb{S}^d$ . For two densities  $f, g \in \mathcal{P}(\mathbb{S}^d)$ , it is frequently needed to measure dissimilarity between the two. Unfortunately, even for the most well-known distributions on the hypersphere, analytic formula for any discrepancy measure is rarely available, leading to require numerical schemes for approximation. Here we focus on  $L_p$  distance between the two densities,

$$L_p(f, g) = \left( \int_{\mathbb{S}^d} |f(x) - g(x)|^p \right)^{1/p} \quad (1)$$

and we show how to combine Monte Carlo way of integration by means of [importance sampling](#) to approximate Equation 1.

## Computation

Importance sampling requires a proposal density. The easiest choice is to use uniform density  $u(x)$  as an importance proposal since sampling from  $u(x)$  is trivial. First, take a random sample from standard normal distribution  $x \sim \mathcal{N}(0, I)$  in  $\mathbb{R}^{d+1}$ . Then, the rest is to take  $L_2$  normalization, i.e.,  $x \leftarrow x/\|x\|_2$ , which makes a sampled vector to have a unit norm. Given the sample generation process, we have the following

$$\begin{aligned} L_p(f, g)^p &= \int_{\mathbb{S}^d} |f(x) - g(x)|^p dx \\ &= \int_{\mathbb{S}^d} \frac{|f(x) - g(x)|^p}{u(x)} u(x) dx \\ &= \mathbb{E}_{u(x)} \left[ \frac{|f(x) - g(x)|^p}{u(x)} \right] \\ &\approx \frac{1}{N} \sum_{n=1}^N \frac{|f(x_n) - g(x_n)|^p}{u(x_n)} \text{ for } x_n \stackrel{iid}{\sim} u(x), \end{aligned}$$

where the last term gets better approximation as  $N \rightarrow \infty$ .

## References

Marron, James Stephen, and I. L. Dryden. 2021. *Object Oriented Data Analysis*. Boca Raton: Taylor & Francis Group, LLC.