Exercise - 1: Von Mises Negative Log-Likelihood

December 2, 2025

Questions

Q1. For a single observation y, write down the log-likelihood $\log p(y \mid \mu, \kappa)$ based on the von Mises distribution, and then express the corresponding negative log-likelihood (NLL).

Q2. Suppose the model predicts $\log \kappa = s$, so that $\kappa = \exp(s)$. Rewrite the NLL in terms of s.

Q3. Briefly explain why predicting $\log \kappa$ instead of κ directly can be advantageous.

Hint

Let $y \in (-\pi, \pi]$. The probability density function of the von Mises distribution is:

$$p(y \mid \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\big(\kappa \cos(y - \mu)\big),$$

where $\kappa \geq 0$ is the shape parameter and $\mu \in (0, 2\pi)$ is the location parameter. The variable y satisfies:

$$0 < y < 2\pi.$$

Here, $I_0(\kappa)$ is the modified Bessel function of order 0, which commonly appears in problems with cylindrical or circular symmetry. It is defined as:

$$I_0(\kappa) = \sum_{k=0}^{\infty} \frac{(\kappa/2)^{2k}}{(k!)^2}.$$

The function $I_0(\kappa)$ normalizes the probability density and is always positive for $\kappa \geq 0$.

Useful points:

- While deriving the log-likelihood and negative log-likelihood (NLL), keep $I_0(\kappa)$ in its original form for simplicity.
- $\log(ab) = \log a + \log b$ and $\log(a^c) = c \log a$.
- $\log(2\pi)$ is constant and can be dropped when optimizing.
- In practice: $\kappa = \text{torch.exp}(s)$, and $\log_0(\kappa) = \text{torch.log(torch.special.i0(kappa))}$.