Exercise 1: Fisher kernel for univariate Gaussian density

Suppose a univariate Gaussian density model \( p(x) = \mathcal{N}(x; \mu, \sigma^2) \).

(a) Compute the partial derivatives \( \frac{\partial \ln p(x)}{\partial \mu} \) and \( \frac{\partial \ln p(x)}{\partial \sigma} \).

Let \( g(x) \) be the two dimensional gradient vector that concatenates the two partial derivatives.

(b) Compute the Fisher information matrix \( F = \int_x p(x)g(x)g(x)^\top \).

(c) Show that \( \int_x p(x)g(x) = 0 \).

(d) Compute the Fisher vector \( h = F^{-\frac{1}{2}}g \).

Exercise 2: Fisher kernel for univariate Gaussian mixture density

Suppose a univariate Gaussian mixture density model \( p(x) = \sum_{k=1}^{K} w_i \mathcal{N}(x; \mu_i, \sigma^2_i) \).

Where the mixing weights are parameterized as \( w_i = \exp(\alpha_i) / \sum_{j=1}^{K} \exp(\alpha_j) \).

(a) Compute the partial derivatives \( \frac{\partial \ln p(x)}{\partial \mu_i} \), and similar for \( \sigma_i \) and \( \alpha_i \).

Let \( g(x) \) be the 3\( K \) dimensional gradient vector that concatenates these partial derivatives. Denote the Fisher information matrix \( F = \int_x p(x)g(x)g(x)^\top \).

Assume that the posteriors \( p(i|x) = w_i \mathcal{N}(x; \mu_i, \sigma^2_i) / p(x) \) are sharply peaked, i.e. close to one for a single \( i \) and close to zero for all others. Decompose \( F \) into 3 \( \times \) 3 blocks, corresponding to the \( w_i, \mu_i \) and \( \sigma_i \).

(b) Show that \( F \) is block diagonal.
(c) Show that the $\mu$ and $\sigma$ blocks are diagonal, and give the diagonal entries.

Fix $\alpha_K = 0$ to remove a redundant degree of freedom from the $\alpha_i$, and let $\vec{\alpha} = (\alpha_1, \ldots, \alpha_{K-1})$. Let $\vec{g}(x) = \nabla_{\vec{\alpha}} \ln p(x)$ be the gradient with respect to $\vec{\alpha}$, and similarly let $\vec{F}$ be the Fisher information matrix with respect to $\vec{\alpha}$.

(d) Show that the Fisher kernel with respect to $\vec{\alpha}$ can be written as $\vec{g}(x)^\top \vec{F}^{-1} \vec{g}(y) = \phi(x)^\top \phi(y)$ where $\phi(x)$ is a $K$ dimensional vector.

Exercise 3: Positive definite kernels

(a) Which of the following kernels are positive definite? You need to provide a proof each time.

- $K(x, y) = 1/(1 - xy)$ with $\mathcal{X} = (-1, +1)$ (interval excluding -1 and 1).
- $K(x, y) = \max(x, y)$ with $\mathcal{X} = [0, 1]$
- $K(x, y) = \cos(x + y)$ with $\mathcal{X} = \mathbb{R}$
- $K(x, y) = \cos(x - y)$ with $\mathcal{X} = \mathbb{R}$
- $K(x, y) = \text{GCD}(x, y)$ (greatest common divisor) with $\mathcal{X} = \mathbb{N}$.

(b) Show that if $K_1$ and $K_2$ are positive definite, then the product $K(x, y) = K_1(x, y)K_2(x, y)$ is also positive definite.