Fisher information metric on the space of Gaussians

Consider the space of univariate Gaussian distributions parametrised by $(\mu, \sigma) \in \mathbb{R} \times (0, \infty)$. The Gaussian Kullback–Leibler divergence from (μ, σ) to (μ_1, σ_1) is:

$$K(\mu_1, \sigma_1) := \mathrm{KL}(\mathcal{N}(\mu, \sigma^2) : \mathcal{N}(\mu_1, \sigma_1^2)) = \log \frac{\sigma_1}{\sigma} + \frac{(\mu - \mu_1)^2 + \sigma^2 - \sigma_1^2}{2\sigma_1^2}$$

This has a global minimum when the points (μ, σ) and (μ_1, σ_1) coincide, $K(\mu, \sigma) = 0$. We can show this because, at this point, the gradient vanishes:

$$\nabla K(\mu, \sigma) = \begin{bmatrix} \frac{\partial K}{\partial \mu} \\ \frac{\partial K}{\partial \sigma} \end{bmatrix} \Bigg|_{(\sigma, \mu)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and the Hessian (matrix of second derivatives) is positive definite:

$$\nabla^2 K(\mu, \sigma) = \begin{bmatrix} \frac{\partial^2 K}{\partial \mu^2} & \frac{\partial^2 K}{\partial \mu \partial \sigma} \\ \frac{\partial^2 K}{\partial \sigma \partial \mu} & \frac{\partial^2 K}{\partial \sigma^2} \end{bmatrix} \Bigg|_{(\mu, \sigma)} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix}$$

An equivalent way to write this is as a metric tensor

$$g = \frac{\mathrm{d}\mu^2 + 2\,\mathrm{d}\sigma^2}{\sigma^2}$$

so that $g(\vec{u}, \vec{v}) = \vec{u}^T (\nabla^2 K) \vec{v}$ for any vectors $\vec{u}, \vec{v} \in \mathbb{R}^2$.

Under a change of coordinates $\mu = \sqrt{2}x, \sigma = y$, the metric g is (twice) the metric of the Poincaré half-plane model of hyperbolic space.