Kernels Homework 1.

April 5, 2025

1. Go through the notebook

   Understanding Kernels

2. Verify numerically that

   $$e^{-\|x-y\|^2/2} \;=\; \mathbb{E}_{\omega,b}\left[\sqrt{2}\cos(\omega^\top x + b)\cdot\sqrt{2}\cos(\omega^\top y + b)\right]$$

   using a Monte-Carlo approximation on the right-hand side, where ω is drawn from $p(\omega) = N(0,1)$ and b is drawn uniformly from $[0, 2\pi]$. how big of a P do you need to make it work?

3. Verify this approximation numerically.

   $$\begin{aligned}\hat{y}(x) \;&=\; k(x)'(zI + \hat{K})^{-1}y \;\approx\; P^{-1/2}f(x)'S'(zI + SS')^{-1}y \\ &=\; P^{-1/2}f(x)'\underbrace{(zI + S'S)^{-1}S'y}_{\hat{\beta}}\end{aligned} \tag{1}$$

   where S are random features from the previous problem.

4. Verify numerically that

   Theorem 1. The Gaussian kernel

   $$K(x,y) \;=\; \exp\Big(-\tfrac{\rho}{1-\rho^2}\,(x-y)^2\Big)$$

   in $L^2(\mathrm{d}\mu)$ when $\mathrm{d}\mu$ is the Gaussian distribution of mean 0 and variance $\frac{1}{2}\,\frac{1+\rho}{1-\rho}$ has f_k as eigenfunctions and

   $$\lambda_k \;=\; (1-\rho)\,\rho^k.$$

   as eigenvalues.

To this end, use Monte-Carlo approximation to verify numerically that

$$\int K(x,\tilde{x}) f_k(\tilde{x}) d\mu(\tilde{x}) \;=\; \lambda_k f_k(x)$$

Reproduce the plot of $f_k(x)$ from the lecture.

5. Check that f_k for large k is "hard to learn" because it "oscillates too much." To this end, fix, say, $n = 500$ observations, simulate 200 Out-of-sample (OOS) observations, and compute OOS MSE of kernel ridge for the simulation $y_i = f_k(x_i)$. What do you see and why? What is the RKHS norm $\|f_k\|_{\mathcal{H}_K}$?

6. Complete the proof of the representer theorem: minimization over $f^*(x) = \sum_{i=1}^n \alpha_i K(x, x_i)$ (i.e., over α_i) gives the kernel ridge.

   Also, prove the $z = 0$ case!