

Exercises

§12.1-12.6: Kernel Methods

Problem 1 For $\mathbf{x}, \mathbf{z} \in \mathbb{R}^2$, expand $K(\mathbf{x}, \mathbf{z}) = (1 + \mathbf{x}^T \mathbf{z})^2$. Construct an explicit feature map $\phi(\mathbf{x})$ such that $K(\mathbf{x}, \mathbf{z}) = \phi(\mathbf{x})^T \phi(\mathbf{z})$. How many explicit features are required?

Problem 2 Suppose $K(\mathbf{x}, \mathbf{z}) = \phi(\mathbf{x})^T \phi(\mathbf{z})$. For training observations $\mathbf{x}_1, \dots, \mathbf{x}_P$, show that the Gram matrix $G_{pq} = K(\mathbf{x}_p, \mathbf{x}_q)$ satisfies $\mathbf{a}^T G \mathbf{a} \geq 0$ for every $\mathbf{a} \in \mathbb{R}^P$. What does this establish about the eigenvalues of G ?

Problem 3 For $x = (-1, 0, 1)$, compute the Gram matrix for $K(x, z) = (1 + xz)^2$. Then compute its eigenvalues. Verify that the matrix is positive semidefinite.

Problem 4 Use the Gram matrix from Problem 3 and responses $y = (1, 0, 1)^T$. Let $\lambda = 1/3$ and $P = 3$. (a) Solve $(G + P\lambda I)\alpha = \mathbf{y}$. (b) Compute the three training predictions $G\alpha$. (c) Compute the kernel-ridge prediction at $x_* = 2$.

Problem 5 In one dimension, start with $K(x, z) = e^{-\gamma(x-z)^2}$. Show that $K(x, z)$ can be written as $\sum_{m=0}^{\infty} \phi_m(x)\phi_m(z)$ with $\phi_m(x) = e^{-\gamma x^2} \sqrt{\frac{(2\gamma)^m}{m!}} x^m$. Why does this demonstrate that the RBF kernel corresponds to an infinite-dimensional feature map?

Problem 6 For the RBF kernel $K(x, z) = e^{-\gamma(x-z)^2}$, let $x = 0$ and $z = 2$. Compute the kernel value for $\gamma = 0.1, 1, 10$. Describe how increasing γ changes the locality of the similarity measure.

Problem 7 A kernel SVM has support-vector representation $f(\mathbf{x}) = b + \sum_{p=1}^P \alpha_p y_p K(\mathbf{x}_p, \mathbf{x})$. Suppose $b = 0$, two support vectors are $x_1 = -1$ with $y_1 = -1$, and $x_2 = 1$ with $y_2 = 1$, and $\alpha_1 = \alpha_2 = 1$. Use $K(x, z) = (1 + xz)^2$. Compute $f(-2), f(0), f(2)$. What labels are predicted away from the zero-score point?

Problem 8 Describe a leakage-safe 5-fold cross-validation procedure for choosing both γ and λ