Kernel-Based Learning & Multivariate Modeling
MIRI Master

Lluís A. Belanche
belanche@cs.upc.edu

Problem Set #4, Oct 8, 2019

Problem 1 Polynomial kernel

Let us study the polynomial kernel $k(\boldsymbol{x}, \boldsymbol{x'}) = (\langle \boldsymbol{x}, \boldsymbol{x'} \rangle + c)^q,\ q \in \mathbb{N}, \boldsymbol{x}, \boldsymbol{x'} \in \mathbb{R}^d, c \geq 0 \in \mathbb{R}$.

1. Do the kernel trick in $d = 2$ dimensions for $q = 3$ and give an explicit characterization of the ϕ map in this case.
2. Give an explicit characterization of ϕ in the general case of q, d. Find the dimension of the feature space as a function of q and d (not counting duplicates).

. . . .

Problem 2 Discrete characterization of kernels

Consider a data sample $\{x_1, \cdots, x_N\}$, with $x_n \in \mathbb{R}^d$. You are asked to show that k is a positive semidefinite (PSD) symmetric function in $\mathbb{R}^d$ if and only if it can be expressed as an inner product. Specifically,

1. Prove that if k is a PSD symmetric function in $\mathbb{R}^d$, then k can be expressed as an inner product (hint: use the spectral decomposition of the associated kernel matrix).
2. Prove that if k is an inner product of functions of the data, then k is a kernel in $\mathcal{X}$ (hint: show that the associated kernel matrix is PSD).

. . . .

Problem 3 Basic kernel properties

Prove that, if k is a kernel, the following assertions hold:

1. $k(\boldsymbol{x}, \boldsymbol{x}) \geq 0$
2. $k(\boldsymbol{x}, \boldsymbol{x'}) = k(\boldsymbol{x'}, \boldsymbol{x})$
3. $|k(\boldsymbol{x}, \boldsymbol{x'})| \leq \sqrt{k(\boldsymbol{x}, \boldsymbol{x}) \cdot k(\boldsymbol{x'}, \boldsymbol{x'})}$

hint: express the kernel as an inner product

. . . .

Problem 4 Basic kernel closure properties

Prove that, if k, k' are kernels on the same domain, and $a \geq 0$, then the following functions are also kernels:

1. $k(\boldsymbol{x}, \boldsymbol{x'}) + k'(\boldsymbol{x}, \boldsymbol{x'})$
2. $ak(\boldsymbol{x}, \boldsymbol{x'})$
3. $k(\boldsymbol{x}, \boldsymbol{x'}) \cdot k'(\boldsymbol{x}, \boldsymbol{x'})$

Can you characterize the associated feature maps?

. . . .

Problem 5 Limit of sequences

Let $\{k_n\}_n : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a sequence of kernels and assume that the following limit exists, for all $\boldsymbol{x}, \boldsymbol{x'} \in \mathcal{X}$:

$$k_\infty(\boldsymbol{x}, \boldsymbol{x'}) := \lim_{n \to \infty} k_n(\boldsymbol{x}, \boldsymbol{x'}),\ \forall \boldsymbol{x}, \boldsymbol{x'} \in \mathcal{X}$$

Prove that k_∞ is a valid kernel.

. . . .

Problem 6 Covariance matrices

Let $X_1, X_2, \cdots, X_d$ be real random variables with expected values $\mu_i = \mathbb{E}(X_i)$ and finite second moments (that is, $\mathbb{E}(X_i^2) < \infty$). The covariance matrix of the random vector $X = (X_1, X_2, \cdots, X_d)^\mathrm{T}$ is the matrix $\Sigma = (\sigma_{ij}^2)$. Let us recall that:

- $\mathbb{E}[X] = (\mathbb{E}[X_1], \ldots, \mathbb{E}[X_d])^\mathrm{T} = \boldsymbol{\mu}$
- $\mathbb{E}[(X - \boldsymbol{\mu})(X - \boldsymbol{\mu})^\mathrm{T}] = \Sigma$

Show that Σ is PSD.

. . . .

Problem 7 Two friends arguing ...

Two friends are discussing about positive linear combinations of kernels. Friend #1 argues that the feature space is essentially the same (in the sensen that the new features are linear combinations of the original ones). Friend #2 is suspicious that something different may be happening. Given k_1, k_2 two kernels, $a, b, c \geq 0$, find the feature map of the kernel $a \cdot k_1 + b \cdot k_2 + c$.

. . . .

Problem 8 Polynomial kernels revisited

Prove that the polynomial kernel $k(\boldsymbol{x}, \boldsymbol{x'}) = (a \cdot \langle \boldsymbol{x}, \boldsymbol{x'} \rangle + c)^q, a > 0, c \geq 0,\ q \in \mathbb{N}$ is a kernel. Extend this result to polynomial-based kernels of the form $p(k(\boldsymbol{x}, \boldsymbol{x'}))$, where p is a finite polynomial with non-negative coefficients. Can you devise an infinite-dimensional polynomial kernel?

. . . .

Problem 9 The exponential is everywhere

Given k a kernel, and a real $\gamma > 0$, prove that $\exp(\gamma k(\boldsymbol{x}, \boldsymbol{x'}))$ is a kernel; prove that $\exp(\gamma[2k(\boldsymbol{x}, \boldsymbol{x'}) - k(\boldsymbol{x}, \boldsymbol{x}) - k(\boldsymbol{x'}, \boldsymbol{x'})])$ is a kernel.

. . . .

Problem 10 The RBF kernel

Consider the function:

$$k(\boldsymbol{x}, \boldsymbol{x'}) = \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x'}\|^2}{2\sigma^2}\right), \boldsymbol{x}, \boldsymbol{x'} \in \mathbb{R}^d$$

popularly known as the RBF kernel. Prove that it is a valid kernel. Hint: expand the square and express the kernel as the product of three terms.

. . . .

Problem 11 Normalizing kernels

Prove that, if k is a kernel, then so is:

$$k_n(\boldsymbol{x}, \boldsymbol{x'}) = \frac{k(\boldsymbol{x}, \boldsymbol{x'})}{\sqrt{k(\boldsymbol{x}, \boldsymbol{x})}\sqrt{k(\boldsymbol{x'}, \boldsymbol{x'})}}$$

Find $k_n(\boldsymbol{x}, \boldsymbol{x})$. Prove that $|k_n(\boldsymbol{x}, \boldsymbol{x'})| \leq 1$. Show the feature map for k_n.

. . . .

Problem 12 Kernels from arbitrary functions

Our two friends are again at it, this time discussing about how to generate valid kernels from *arbitrary* functions. Friend #1 argues that she can take any $f : \mathcal{X} \to \mathbb{R}^D$, an arbitrary function of the data, and define $k_f(\boldsymbol{x}, \boldsymbol{x'}) := f(\boldsymbol{x})^{\mathrm{T}} f(\boldsymbol{x'})$. She even claims that in D could be infinite! Friend #2 says the validity of the kernel would depend on the precise form of f. Can you give a formal answer to this argument?

. . . .

Problem 13 Conditionally positive semi-definite kernels

Let $k(\boldsymbol{x}, \boldsymbol{x'}) = -\|\boldsymbol{x} - \boldsymbol{x'}\|^2, \boldsymbol{x}, \boldsymbol{x'} \in \mathbb{R}^d$. Show that this function is CPSD but not PSD. Try to find a way of making it PSD. Hint: you can use your result to obtain yet another validity proof for the RBF kernel.

. . . .