Assignment 1 - MVP's

Lindi Li
3460570

Gewei Cao
3461232

December 2022

§ 1 Intruduction

blabla

§ 2 Reproducing Kernel Hilbert Space

§ 2.1 What is Kernel

In the simplest form of machine learning, in order to predict x, the algorithm collects the samples in the training set χ that are similar to x, and then take the weighted value of these samples as the predict value of x. Here comes the questions:

- How to measure the similarity between samples?
- How to weight the value of each sample?

In general, the higher the similarity of the sample to our point of interest x, the more the sampling weights. We set $y_i \in \mathbb{R}$ as dependent variable, and x_i as a $1 \times D$ vector x_i in $\mathbb{R}^D$. Assume that (y_i, x_i) where $i = 1, \ldots, N$ is i.i.d. To evaluate the similarity between two observations, a kernel is defined as a function of two input patterns $k(x_i, y_i)$, mapping onto a real-valued output. For example, the Gaussian kernel is

$$k(x_i, x_j) = e^{\frac{\|x_i - x_j\|}{\sigma^2}},$$

where $\| x_i - x_j \|$ is the Euclidean distance between x_i and x_j, and $\sigma^2 \in \mathbb{R}^+$ is the bandwidth of the kernel function.
We now define that $k : \chi \times \chi \to \mathbb{R}$ is a kernel if

- k is symmetric: $k(x, y) = k(y, x)$.
- k is positive semi-definite, meaning that $\sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) \geq 0, \forall \alpha_i, \alpha_j \in \mathbb{R}, x \in \mathbb{R}^{\mathbb{D}}, D \in \mathbb{Z}^+$.

From the similarity-based point of view, the use of kernels for regression can be described in two stages. We first set a target function $y = f(x)$ and assume that in a space of functions, there exists a function that can estimate $y = f(x)$ well. The target function is represented by

$$f(x) = \sum_{i=1}^{N} c_i k(x, x_i),$$

In the second stage, we utilize regularization to simplify the function. To achieve this purpose, Hilbert space and reproducing kernel Hilbert space will be introduced below.

§ 2.2 Hilbert Space

Recall that an inner product $< a, b >$ can be

- a usual dot product: $< a, b >= a'b = \sum_i a_i b_i$.
- a kernel product: $< a, b >= k(a, b) = \psi(a)'\psi(b)$, where $\psi(a)$ may have infinite dimensions.

We define a Hilbert space an inner product space that....

§ 3 Bayies

blabla

§ 4 Empirical study

blabla