Kernel Methods in Machine Learning - Course Notes

Hugo Cisneros

1 Kernels and RKHS

1.1 Positive Definite Kernels

A kernel K is a comparison function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$.

With n data point $\{x_1, x_2, ..., x_n\}$ a $n \times n$ matrix $\mathbf{K}$ can be defined by $\mathbf{K}_{ij} = K(x_i, x_j)$.

A kernel K is positive definite (p.d.) if it is symmetric ($K(x, x') = K(x', x)$) and for all sets of a and x

$$\sum_i \sum_j a_i a_j K(x_i, x_j) \geq 0$$

This is equivalent to the kernel matrix being positive semi-definite.

Example: Kernel on $\mathbb{R} \times \mathbb{R}$ defined by $K(x, x') = xx'$ is p.d. ($xx' = x'x$ and $\sum_i \sum_j a_i a_j K(x_i, x_j) = \left(\sum_i a_i x_i\right)^2 \geq 0$).

1.2 Reproducing Kernel Hilbert Spaces (RKHS)

2 Kernel tricks

3 Kernel Methods: Supervised Learning

4 Kernel Methods: Unsupervised Learning

5 The Kernel Jungle

6 Open Problems and Research Topics