Kernel Matrix Calculation

Dr Chibisi Chima-Okereke

2020-05-23

The Kernel Matrix

Consider a data matrix X consisting of n vectors $x_1, x_2, \ldots, x_n$ each of size m. The kernel matrix $\mathbf{K}$ is symmetric and calculated from kernel functions of the form $K(x_i, x_j)$ which forms element $K_{i,j}$ of the kernel matrix [1].

Kernel Functions

The following kernel functions are used in the article [2, 3, 4]:

- Linear Kernel (dot product)
$$K(x_i, x_j) = x_i^T x_j \tag{1}$$
- Gaussian
$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{\theta}\right) \tag{2}$$
- Polynomial
$$K(x_i, x_j) = (x_i^T x_j + \text{Offset})^d \tag{3}$$
- Exponential
$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|}{\theta}\right) \tag{4}$$
- Log Kernel
$$K(x_i, x_j) = -\log\left(1 + \|x_i - x_j\|^\beta\right) \tag{5}$$
- Cauchy
$$K(x_i, x_j) = \frac{1}{1 + \frac{\|x_i - x_j\|^2}{\theta}} \tag{6}$$

- Power

$$K(x_i, x_j) = -\|x_i - x_j\|^\beta \tag{7}$$

- Wave

$$K(x_i, x_j) = \frac{\theta}{\|x_i - x_j\|} \sin\left(\frac{\|x_i - x_j\|}{\theta}\right) \tag{8}$$

- Sigmoid

$$K(x_i, x_j) = \tanh\left(\beta_0 x_i^T x_j + \beta_1\right) \tag{9}$$

The kernel matrix calculation scales $O(m\ n^2)$ with the dimensions of the data matrix X.

References

[1] An introduction to support vector machines and other kernel-based learning methods, Nello Cristianini & John Shawe-Taylor

[2] Data Variant Kernel Analysis, Yuichi Motai

[3] Classes of Kernels for Machine Learning: A Statistics Perspective, Marc G. Genton.

[4] Conditionally Positive Definite Kernels for SVM Based Image Recognition, S. Boughorbel, J. -. Tarel and N. Boujemaa, 2005 IEEE International Conference on Multimedia and Expo, Amsterdam, 2005, pp. 113-116