

Kernel Method 03 Necessary and Sufficient Conditions

Chen Gong

22 November 2019

在上一小节中，我们描述了正定核的两个定义，并且认为这两个定义之间是相互等价的。下面我们就要证明他们之间的等价性。

1 充分性证明

大家注意到在上一节的描述中，我似乎没有谈到对称性，实际上是因为对称性的证明比较的简单。就没有做过多的解释，那么我重新描述一下我们需要证明的问题。

已知： $K(x, z) = \langle \phi(x), \phi(z) \rangle$ ，证：Gram Matrix 是半正定的，且 $K(x, z)$ 是对称矩阵。

对称性：已知：

$$K(x, z) = \langle \phi(x), \phi(z) \rangle \quad K(z, x) = \langle \phi(z), \phi(x) \rangle \quad (1)$$

又因为，内积运算具有对称性，所以可以得到：

$$\langle \phi(x), \phi(z) \rangle = \langle \phi(z), \phi(x) \rangle \quad (2)$$

所以，我们很容易得到： $K(x, z) = K(z, x)$ ，所以对称性得证。

正定性：我们要证的是 Gram Matrix = $K[K(x_i, x_j)]_{N \times N}$ 是半正定的。那么，对一个矩阵 $A_{N \times N}$ ，我们如何判断这是一个半正定矩阵？大概有两种方法，1. 这个矩阵的所有特征值大于等于 0；2. 对于 $\forall \alpha \in \mathbb{R}^N$ ， $\alpha^T A \alpha \geq 0$ 。这个是充分必要条件。那么，这个问题上我们要使用的方法就是，对于 $\forall \alpha \in \mathbb{R}^N$ ， $\alpha^T A \alpha \geq 0$ 。

$$\alpha^T K \alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_N \end{bmatrix} \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1} & K_{N2} & \cdots & K_{NN} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \quad (3)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j K_{ij} \quad (4)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \langle \phi(x_i), \phi(x_j) \rangle \quad (5)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \phi(x_i)^T \phi(x_j) \quad (6)$$

$$= \sum_{i=1}^N \phi(x_i)^T \sum_{j=1}^N \phi(x_j) \quad (7)$$

$$= \left[\sum_{i=1}^N \phi(x_i) \right]^T \left[\sum_{j=1}^N \phi(x_j) \right] \quad (8)$$

$$= \left\| \sum_{i=1}^N \alpha_i \phi(x_i) \right\|^2 \geq 0 \quad (9)$$

所以，我们可以得到 K 是半正定的，必要性得证。

2 必要性证明

充分性得到证明之后，必要性的证明就会变得很简单了。这个证明可以被我们描述为：

已知：Gram Matrix 是半正定的，且 $K(x, z)$ 是对称矩阵。证：存在一个映射 $\phi: \mathcal{X} \mapsto \mathbb{R}^p$ ，使得 $K(x, z) = \langle \phi(x), \phi(z) \rangle$ 。

对于我们建立的一个映射 $\phi(x) = K(x, \cdot)$ ，我们可以得到 $K(x, \cdot)K(z, \cdot) = K(x, z)$ 。所以有 $K(x, z) = K(x, \cdot)K(z, \cdot) = \phi(x)\phi(z)$ 。我们就得证了，具体的理解可以参考我之前写的关于可再生核希尔伯特空间的理解。

3 小结

讲到这里，我们的核方法就讲完了。核这个东西是真的不好理解，讲到这大家应该有个印象了，这一节配合再生核希尔伯特空间使用食用，效果更佳！下面，我会编写对于 Stein 变分梯度下降论文的详细了解，这个方法真的很难！而这个系列的下一次，我们将进入到概率图模型，贝叶斯方法的大 boss，终于要来了，想想就很激动了！