UNIVERSITY OF CHINESE ACADEMY OF SCIENCES

CS091M4042H Assignment ——SVM

熊兴旺*
Sno: 2018E8013261007
中国科学院计算技术研究所

November 29, 2018

*email: xingw.xiong@gmail.com; xiongxingwang@ict.ac.cn

1 Problem 1

1.1 Problem Description

Show that, irrespective of the dimensionality of the data space, a data set consisting of just two data points (call them $x^{(1)}$ and $x^{(2)}$, one from each class) is sufficient to determine the maximum-margin hyperplane. Fully explain your answer, including giving an explicit formula for the solution to the hard margin SVM (i.e., w) as a function of $x^{(1)}$ and $x^{(2)}$.

1.2 Solution

可以 SVM 分类超平面 $f(x) = \omega^T x + \omega_0$ 的求解可以转化为下面的最优化问题：

$$\min \frac{\|\omega\|^2}{2} \tag{1.1}$$

$$\textbf{s.t.} \quad y^{(i)}(\omega^T x^{(2)} + \omega_0) \geq 1 \tag{1.2}$$

给定两个不同样本点 $x^{(1)}$，$x^{(2)}$(设 $x^{(1)}$ 为正样本)，那么，这两个样本点分别位于两个不同的超平面上，即满足：

$$\begin{cases} \omega^T x^{(1)} + \omega_0 = +1 \\ \omega^T x^{(2)} + \omega_0 = -1 \end{cases} \tag{1.3a}$$

因此，可得：

$$\omega = \frac{2}{x^{(1)} - x^{(2)}}, \tag{1.4a}$$

$$\omega_0 = -\frac{x^{(1)} + x^{(2)}}{x^{(1)} - x^{(2)}} \tag{1.4b}$$

因此，SVM 分类超平面方程为 $f(x) = \frac{2}{x^{(1)} - x^{(2)}} x - \frac{x^{(1)} + x^{(2)}}{x^{(1)} - x^{(2)}}$

2 Problem 2

2.1 Problem Description

Gaussian kernel takes the form:

$$k(x, x') = \exp(-\left\|x - x'\right\|^2 / 2\sigma^2) \tag{2.1}$$

Try to show that the Gaussian kernel can be expressed as the inner product of an infinite-dimensional feature vector.

Hint: Making use of the following expansion, and then expanding the middle factor as a power series.

$$k(x, x') = \exp(-x^T x / 2\sigma^2) \exp(x^T x' / \sigma^2) \exp(-(x')^T x' / 2\sigma^2) \tag{2.2}$$

2.2 Solution

由于 $k(x,x') = \exp(-x^T x/2\sigma^2)\exp(x^T x'/\sigma^2)\exp(-(x')^T x'/2\sigma^2)$,
又有 $\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \ldots$，故：

$$k(x,x') = \exp(-x^T x/2\sigma^2)\exp(x^T x'/\sigma^2)\exp(-(x')^T x'/2\sigma^2) \tag{2.3}$$

$$= \exp\left(-\frac{x^T x + x'^T x'}{2\sigma^2}\right)\left(1 + \left(\frac{x^T x'}{\sigma^2}\right) + \frac{\left(\frac{x^T x'}{\sigma^2}\right)^2}{2!} + \cdots + \frac{\left(\frac{x^T x'}{\sigma^2}\right)^n}{n!} + \ldots\right) \tag{2.4}$$

$$= \exp\left(-\frac{x^T x + x'^T x'}{2\sigma^2}\right)\left(\sum_{n=0}^{\infty}\frac{x^T}{\sigma^n\sqrt{n!}}\cdot\frac{x'}{\sigma^n\sqrt{n!}}\right) \tag{2.5}$$

令：

$$\Phi(x) = \exp\left(-\frac{x^2}{\sigma^2}\right)\left[1, \frac{x}{\sigma}, \frac{x^2}{\sigma^2\sqrt{2!}}, \ldots, \frac{x^n}{\sigma^n\sqrt{n!}}, \ldots\right] \tag{2.6}$$

那么有：$k(x,x') = \Phi(x)^T \cdot \Phi(x')$，$\Phi(x)$ 表示的向量 x 无穷维特征向量。