

习题 7

7.1 比较经验风险最小化原理和结构风险最小化原理。

解：略。

7.2 VC维的含义是什么？为什么说VC维反映了函数集的学习能力？

解：略。

7.3 为什么说统计学习理论是支持向量机的理论基础？表现在哪些方面？

解：略。

7.4 描述支持向量机的基本思想和数学模型。

解：（1）基本思想略。

（2）数学模型：

最小化泛函：

$$\Phi(w) = \frac{1}{2}(w \cdot w)$$

约束条件为不等式类型：

$$y_i [(x_i \cdot w) - b] \geq 1, i = 1, 2, \dots, l$$

7.5 比较感知机的对偶形式和线性可分支持向量机的对偶性形式。

解：感知机原始模型： $\min_{w,b} L(w,b) = - \sum_{x_i \in M} (y_i (w \cdot x_i + b))$

M 为误分点的集合。等价于

$$\min_{w,b} L(w,b) = \sum_{i=1}^N (-y_i (w \cdot x_i + b))_+$$

对偶形式为： w, b 表示为 x_i, y_i 的线性组合形式，其系数 $w = \sum_{i=1}^N \alpha_i y_i x_i, b = \sum_{i=1}^N \alpha_i y_i$

$$\min_{w,b} L(w,b) = \min_{\alpha_i} L(\alpha_i) = \sum_{i=1}^N (-y_i (\sum_{i=1}^N \alpha_i y_i x_i \cdot x_j + \sum_{i=1}^N \alpha_i y_i))_+$$

线性可分支持向量机原始模型：

$$\min_{w,b} \frac{1}{2} \|w\|^2$$

$$s.t. \quad y_i (w \cdot x_i + b) - 1 \geq 0$$

线性可分支持向量机对偶问题：

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i$$

$$s.t. \quad \sum_{i=1}^N \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C, \quad i=1,2,\dots,N$$

最终 w^*, b^* 可以按下式求出, $w^* = \sum_{i=1}^N \alpha_i^* y_i x_i, b^* = y_j - \sum_{i=1}^N \alpha_i^* (x_i \cdot x_j)$ 可以看出, w^*, b^* 实

质也是 x_i, y_i 的线性组合形式。

7.6 简述非线性支持向量机的核心思想。

解：略。