



模式识别

第三章

贝叶斯分类：错误分析与多类





最优分类器

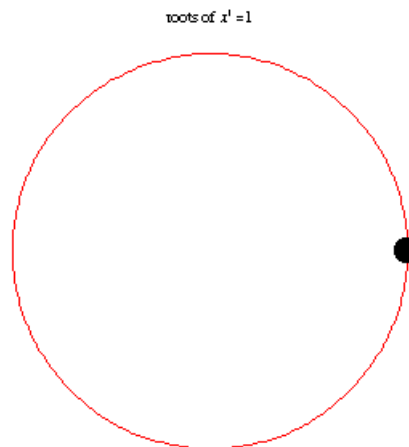




贝叶斯决策与分类

理论上的最优分类器

分类错误率最小





Bayes法则—最小错误贝叶斯分类

对于两类 ω_1 , ω_2 问题, 直观地, 可以根据后验概率做判决:

若 $p(\omega_1|\vec{x}) > p(\omega_2|\vec{x})$ 则 $\vec{x} \in \omega_1$

若 $p(\omega_1|\vec{x}) < p(\omega_2|\vec{x})$ 则 $\vec{x} \in \omega_2$





两个类别的分类

$$p(\omega_1) + p(\omega_2) = 1$$

$$p(\omega_1 | x) + p(\omega_2 | x) = 1$$





两个类别的分类

每类的样本只有两个状态值时

$$p(+|\omega_1) + p(-|\omega_1) = 1$$

$$p(+|\omega_2) + p(-|\omega_2) = 1$$





贝叶斯分类错误率分析

两类问题错误率公式：

$$P(\text{error} | x) = \min(P(\omega_1 | x), P(\omega_2 | x))$$





贝叶斯分类错误率分析

能取得更小的错误率吗？
不可能！！

$$P(\text{error} | x) = \min(P(\omega_1 | x), P(\omega_2 | x))$$





贝叶斯分类错误率分析

两类问题:

$$P(\omega_1 | x) + P(\omega_2 | x) = 1$$

假设 $P(\omega_1 | x)$ 、 $P(\omega_2 | x)$ 的计算值绝对准确，代表了真实值

(1) 样例 x 属于第一与第二类的概率分别为 $P(\omega_1 | x)$ 、 $P(\omega_2 | x)$





贝叶斯分类错误率分析

(2) 分类决策只能选择一个类别（概率最大的类别）！





贝叶斯分类：多类情形

分类规则：Decide ω_i if $P(\omega_i | x) > P(\omega_j | x) \forall j \neq i$

错误率： $P(\text{error} | x) = 1 - \max(P(\omega_1 | x), P(\omega_2 | x) \dots P(\omega_C | x))$





思考

(1) 贝叶斯分类器虽然理论上最优，但并不是万能的

(2) 会存在分类错误, 除非

$$\max(P(\omega_1 | x), P(\omega_2 | x) \dots P(\omega_C | x)) = 1$$

(3) 贝叶斯分类器的错误率是理论上的最低错误率，是由于问题本身的复杂性引起的（越复杂，错误率越高）





更多思考

- (1) 贝叶斯分类器是否完善?
- (2) 贝叶斯分类器应用是否很便捷?





更多思考

基于后验概率的类别可分性分析??

