

有一串向量 $\{x_1, x_2, \dots, x_p\}$

其中 $x_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}$

每个 x_i 属于 C_1 或 C_2

问题: N 个维度有冗余

如何从 N 个维度中选取 M 个维度

($M < N$) 使识别率最高

(特征选择问题)

N 个维度 $\{x_{i1}, x_{i2}, \dots, x_{in}\}$

构造 $\{f_1(x_{i1} \sim x_{in}), f_2(x_{i1} \sim x_{in}), \dots, f_m(x_{i1} \sim x_{in})\}$

(特征提取问题)