

9. 聚类

1. 距离度量

内部

外部

距离计算

$$\left(\sum_{u=1}^n |t_{iu} - t_{ju}|^p \right)^{1/p}$$

原始聚类

k均值

$$E = \sum_{i=1}^k \sum_{x \in G_i} \|x - \mu_i\|_2^2$$

1. k值选择
2. μ_i 初始化

学习向量量化

$$\begin{cases} p' = p_i + \eta (t_j - p_i) & \text{近} \\ p' = p_i - \eta (t_j - p_i) & \text{远} \end{cases}$$

密度聚类

(ϵ , MinPts)

高斯混合聚类 EM

$$p(x) = \sum_{i=1}^k \alpha_i p(x|\mu_i, \sigma_i)$$

参数

层次聚类

$m \rightarrow$ 最近两个合并 $\rightarrow k$

