

## 0.1 Affinity Propagation (AP) 聚类算法

Affinity Propagation (AP) 是一种基于消息传递的聚类算法，通过在数据点之间传递“责任” (Responsibility,  $r(i, k)$ ) 和“可用性” (Availability,  $a(i, k)$ ) 消息，自动确定簇的数量和中心点。算法首先计算数据点之间的相似度  $s(i, k)$ ，然后初始化责任和可用性矩阵。

AP 的目标是通过不断更新以下公式，使得责任和可用性值达到平衡：

$$r(i, k) = s(i, k) - \max_{k' \neq k} \{a(i, k') + s(i, k')\}$$

$$a(i, k) = \min \left( 0, r(k, k) + \sum_{i' \notin \{i, k\}} \max(0, r(i', k)) \right)$$

其中，责任  $r(i, k)$  表示数据点  $i$  作为簇中心  $k$  的适合度，可用性  $a(i, k)$  表示数据点  $i$  选择  $k$  作为簇中心的适合度。最终具有较高“可用性”和“责任值”的点被选择为簇中心。

AP 的优点在于不需要预先指定簇的数量，且能够处理任意形状的簇。缺点在于对相似度度量和偏好值敏感，且计算复杂度较高。AP 算法的时间复杂度为  $O(n^2 \cdot \log(n))$ ，其中  $n$  为样本数量。