

Affinity Propagation Algorithm

Jerome Chou

February 1, 2025

Introduction to Affinity Propagation

- ▶ Affinity Propagation is a clustering algorithm that identifies exemplars among data points.
- ▶ It does not require a predefined number of clusters.
- ▶ It uses message passing between data points to determine cluster centers.

Mathematical Formulation of Affinity Propagation

- ▶ Given a similarity matrix $S(i, j)$ that represents how well-suited x_j is as an exemplar for x_i .
- ▶ Two key messages are exchanged:
 - ▶ **Responsibility**: Measures how well-suited point j is to be an exemplar for point i .

$$r(i, j) \leftarrow S(i, j) - \max_{j' \neq j} \{a(i, j') + S(i, j')\} \quad (1)$$

- ▶ **Availability**: Reflects how appropriate it is for point i to choose j as an exemplar.

$$a(i, j) \leftarrow \min \left(0, r(j, j) + \sum_{i' \notin \{i, j\}} \max(0, r(i', j)) \right) \quad (2)$$

- ▶ Iterations continue until convergence.

Affinity Propagation Algorithm Steps

- ▶ Initialize similarity matrix $S(i, j)$.
- ▶ Set diagonal values $S(i, i)$ as preferences for points to become exemplars.
- ▶ Iteratively update responsibility and availability messages.
- ▶ Stop when messages converge or reach a set iteration limit.

Machine Learning Applications of Affinity Propagation

- ▶ Customer segmentation in marketing.
- ▶ Image recognition and pattern detection.
- ▶ Document and text clustering.
- ▶ Bioinformatics, such as gene expression analysis.