Applied Machine Learning

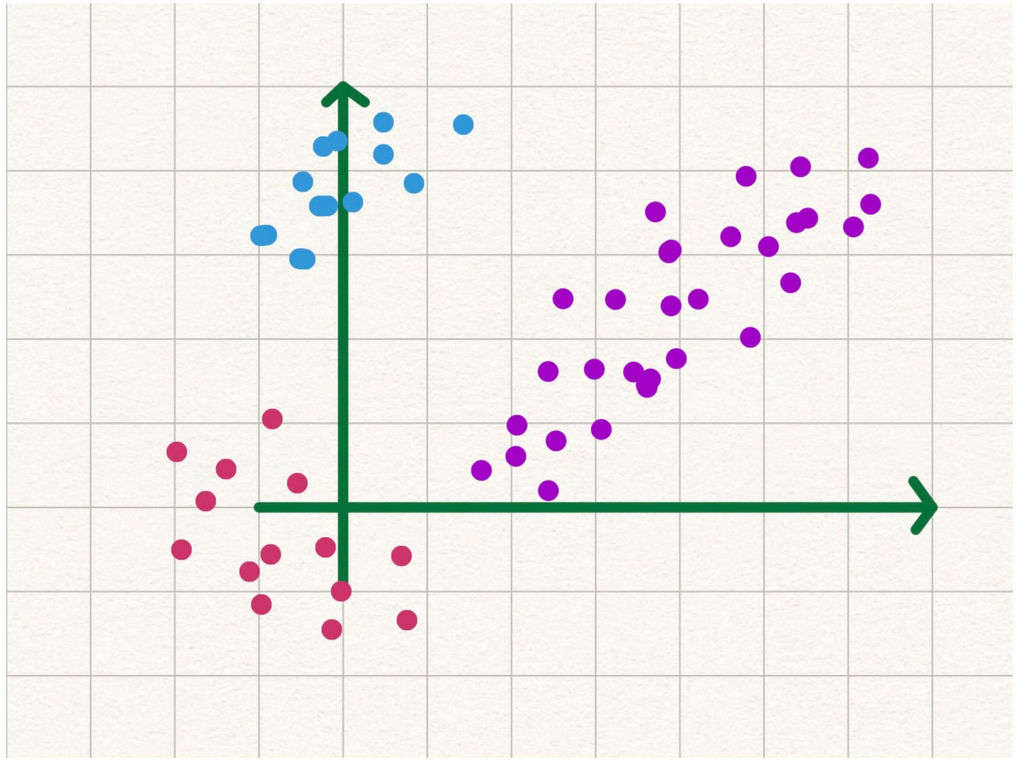
Clustering Using Probability Models

Clustering Using Probability Models

- Clustering review
- Expectation Maximization (EM) for Mixture Models
- General features of EM

k-Means Clustering

1. Initialization: choose k data items as cluster centers $\mathbf{c}_j$
2. While (cluster centers have significant changes)
    1. For each data item $\mathbf{x}_i$
        - closest_center_from($\mathbf{x}_i$).assign($\mathbf{x}_i$)
    2. For each empty cluster center $\mathbf{c}_j$
        - $\mathbf{c}_j$.assign(item_far_from_its_center())
    - For each cluster center $\mathbf{c}_j$
        - $\mathbf{c}_j$.center = $\mathbf{c}_j$.mean()



Expectation Maximization for Mixture Models

- EM
  - Each cluster j
    - probability distribution with parameters θ_j
  - Each item $\mathbf{x}_i$ associated to cluster center j through weight $w_{i,j}$
  - Iterate similarly to k-Means, at each time step n
    - determine weights $w_{i,j}$ to associate each item $\mathbf{x}_i$ to probability distributions j
    - update parameters of probability distributions $\theta^{(n)}$

EM Algorithm

1. Initialize probability distributions
2. While ($\theta^{(n)}$ has not reached convergence)
    1. E-step
        - $p(\delta \mid \theta^{(n)}, \mathbf{x})$
        - $$\begin{aligned} \mathcal{Q}(\theta; \theta^{(n)}) &= \sum_{\delta} \mathcal{L}(\theta; \mathbf{x}, \delta) p(\delta \mid \theta^{(n)}, \mathbf{x}) \\ &= \mathbb{E}_{p(\delta \mid \theta^{(n)}, \mathbf{x})}[\mathcal{L}(\theta; \mathbf{x}, \delta)] \end{aligned}$$
        - $w_{i,j}$ to associate each item $\mathbf{x}_i$ to cluster center j
    2. M-step
        - $\theta^{(n+1)} = \underset{\theta}{\operatorname{argmax}}\, Q(\theta; \theta^{(n)})$

- Number of clusters
- Initialization from k-Means clustering
- Local minima
- Numerical issues
- $$\frac{\pi_k e^{-(\mathbf{x}_i - \mu_k)^\top (\mathbf{x}_i - \mu_k)/2}}{\sum_u \pi_u e^{-(\mathbf{x}_i - \mu_u)^\top (\mathbf{x}_i - \mu_u)/2}}$$

Clustering Using Probability Models

- Clustering review
- Expectation Maximization (EM) for Mixture Models
- General features of EM

Applied Machine Learning

Clustering Using Probability Models