K-means

Sewoong Oh

CSE446
University of Washington
Unsupervised learning and K-means

- in unsupervised learning, the dataset consists of only inputs \( \{x_i\}_{i=1}^{n} \)
- and we **do not have any labels**
- objective: find some pattern
- in particular, most popular unsupervised learning task is clustering: cluster the data into \( K \) groups
K-means: an iterative algorithm for clustering

demo: https://www.naftaliharris.com/blog/visualizing-k-means-clustering/
K-means: an iterative algorithm for clustering

**Step 1**
Start with K centers at arbitrary positions.

**Step 2**
Assign each point to its closest center.

**Step 3**
Recalculate the centers each as the center of its cluster.

**Step 4**

**Step 5**

**Step 6**

K-means: an iterative algorithm for clustering

- in this example K-means converged, i.e. it does not change after this point
- will it always converge? Yes
- does it converge to the right answer (whatever that means)? no

- K-means algorithm
  - input: dataset \( \{x_i\}_{i=1}^n \), number of clusters \( K \)
  - output: cluster assignment \( z_i \) for each data point \( x_i \)
  - Initialize each center \( \mu_k \) to a random location for \( k \in \{1,\ldots,K\} \)
  - repeat
    - (assign each point to its nearest cluster-center)
      \[ z_i = \arg\min_k \|x_i - \mu_k\|_2 \text{ for all } i \in \{1,\ldots,n\} \]
    - (recenter each cluster-center)
      - for \( k \in \{1,\ldots,K\} \)
        \[ X_k \leftarrow \{x_i | z_i = k\} \]
        \[ \mu_k \leftarrow \text{mean}(X_k) \]
      - while any \( \mu_k \) changed from previous value
What would we like to do?

- K-means algorithm is trying to minimize the following objective function
  \[
  \text{minimize} \sum_{i=1}^{n} \left\{ \min_{k' \in \{1, \ldots, K\}} \| x_i - \mu_{k'} \|_2^2 \right\}
  \]
  - this can be written in terms of the assignments $z_i$'s as
  \[
  \text{minimize} \sum_{i=1}^{n} \| x_i - \mu_{z_i} \|_2^2
  \]
  - as K-means is alternatively minimizing
    - (the assignment step)
      fix $\mu_k$'s and minimize the objective over $z_i$'s
    - (the re-centering step)
      fix $z_i$'s and minimize the objective over $\mu_k$'s
  - in particular, the objective can only decrease at each step of K-means
Proof of convergence

• there is only a finite set of values that \( \{z_i\}_{i=1}^n \) can take (\( K^n \) is large but finite)

• so there is only finite, \( K^n \) at most, values for cluster-centers also

• each time we update them, we will never increase the objective function
  \[
  \mathcal{L}(z_1, \ldots, z_n, \mu_1, \ldots, \mu_K) = \sum_{i=1}^{n} \|x_i - \mu_{z_i}\|_2^2
  \]

• the objective is lower bounded by zero

• after at most \( K^n \) steps, the algorithm must converge (as the assignments \( z_i \)'s cannot return to previous assignments in the course of K-means iterations)
downsides of K-means

- it requires the number of clusters $K$ to be specified by us
- the final solution depends on the initialization
  (does not find global minimum of the objective)

Initial position of centers  final converged assignment

Trial 1

Trial 2
K-means++: a smart initialization

Smart initialization:
1. Choose first cluster center uniformly at random from data points
2. Repeat \( K-1 \) times
   3. For each data point \( x_i \), compute distance \( d_i \) to nearest cluster center
   4. Choose new cluster center from amongst data points, with probability of \( x_i \) being chosen proportional to \( (d_i)^2 \)

• apply standard K-means after the initialization

\[
\mu_2 = \begin{cases} 
  x_1 = \mu_1 \\
  x_2 \sim \mathcal{P} \left( \frac{\| \mu_1 - x_2 \|^2}{\sum_{i=2}^{N} \| \mu_i - x_i \|^2} \right) \\
  \vdots \\
  x_N 
\end{cases}
\]