

## 1. K-Means

(a)

data: {1, 2, 3}, {4, 9, 12, 6, 10, 9}

second iteration: cluster centers:  $(1+2+3)/3 = 2$ ,  $(4+9+12+6+10+9)/6 = 8.3333$

data assignment: {1, 2, 3, 4}, {9, 12, 6, 10, 9}

third iteration: cluster centers:  $(1+2+3+4)/4 = 2.5$ ,  $(9+12+6+10+9)/5 = 9.2$

data assignment: {1, 2, 3, 4}, {9, 12, 6, 10, 9}

(b) Algorithm has converged, since re-calculating distances, re-assigning cases to clusters results in no change.

## 2. K-Means and Variance

(a) As we increase K, the variance decreases. Data is partitioned into more concentrated clusters that have smaller variance.

(b)  $k = n$  to get 0 variance. Each cluster only contains one data point which itself is the mean, so the variance is 0.

## 3. Reinforcement Learning I

$$r_{t+1} = 0$$

$$r_{t+2} = 0$$

$$r_{t+3} = 0 \quad \longrightarrow \quad \text{expected total reward is } R_t = 1 \text{ when exit}$$

$$r_{t+4} = 1 \text{ (exit)}$$

The learning agent received the same rewards no matter how many steps taken.

$$r_{t+1} = -1$$

$$r_{t+2} = -1$$

$$r_{t+3} = -1 \quad \longrightarrow \quad \text{expected total reward is } R_t = -k + 1 \text{ in } k \text{ steps when exit}$$

$$r_{t+4} = 1 \text{ (exit)}$$

The learning agent learned to maximize  $R_t$  by escaping in minimum number of steps, and achieved the highest reward of  $R_t = 1$ .

## 4. Reinforcement Learning II (Extra Credit)

(a) only intervals between rewards are important.

(b) Proof:

$$\text{We know } R_t = \sum_{k=1}^{\infty} \gamma^k r_{t+k+1}$$

Then we add a constant  $C$  to  $r_{t+k+1}$

$$R'_t = \sum_{k=1}^{\infty} \gamma^k (r_{t+k+1} + C) = \sum_{k=1}^{\infty} \gamma^k r_{t+k+1} + \sum_{k=1}^{\infty} \gamma^k C = R_t + C \sum_{k=1}^{\infty} \gamma^k$$

Therefore, only intervals between rewards are important.

$$(c) K = C \sum_{k=1}^{\infty} \gamma^k$$

## Part II: Image Segmentation (Extra Credit)

