1.

First, expand $\widetilde{WCSS}(C_j)$:

$$\widetilde{WCSS}(C_j) = \frac{1}{|C_j|} \sum_{a,b \in C_j} |x_a - x_b|^2$$

Expanding the squared norm:

$$= \frac{1}{|C_j|} \sum_{a,b \in C_j} \left(2|x_a|^2 - 2x_a^T x_b\right)$$

$$= \frac{2}{|C_j|} \left[|C_j| \sum_{i \in C_j} |x_i|^2 - 2 \left(\sum_{i \in C_j} x_i \right)^T \left(\sum_{i \in C_j} x_i \right) \right]$$

$$= 2 \sum_{i \in C_j} |x_i|^2 - \frac{2}{|C_j|} \left| \sum_{i \in C_j} x_i \right|^2$$

Now, for $WCSS(C_j)$:

$$WCSS(C_j) = \sum_{i \in C_j} |x_i - \mu_j|^2 = \sum_{i \in C_j} |x_i|^2 - \frac{1}{|C_j|} \left| \sum_{i \in C_j} x_i \right|^2$$

Therefore:

$$\widetilde{WCSS}(C_j) = 2 \cdot WCSS(C_j)$$

Since minimizing $f(x)$ is equivalent to minimizing $2f(x)$, minimizing $\sum_j WCSS(C_j)$ is equivalent to minimizing $\sum_j \widetilde{WCSS}(C_j)$.

2.

Expand the squared norm:

$$J(u; D, C) = \sum_{i \in C} (x_i - u)^T (x_i - u)$$

$$= \sum_{i \in C} \left(x_i^T x_i - 2x_i^T u + u^T u \right)$$

$$= \sum_{i \in C} x_i^T x_i - 2u^T \sum_{i \in C} x_i + |C| u^T u$$

Take the gradient with respect to u and set to zero:

$$\frac{\partial J}{\partial u} = -2 \sum_{i \in C} x_i + 2|C| u = 0$$

Solving for u:

$$u^* = \frac{1}{|C|} \sum_{i \in C} x_i$$

This is the mean of the points in cluster C, which minimizes the sum of squared distances.

3.

(a)

With initial centroids $\mu_1 = (-1, 0, 0)$ and $\mu_2 = (1, 1, 1)$, and weighted centroid formula:

$$\mu(C) = \frac{\sum_{i \in C} w_i x_i}{\sum_{i \in C} w_i}$$

Iteration 1:

- Centroids: $[[-0.83125, 0.475, -0.6125], [1.38125, 0.165, 0.875]]$

Iteration 2:

- Centroids: $[[-0.83125, 0.475, -0.6125], [1.38125, 0.165, 0.875]]$

The algorithm converges after one iteration.

(b)

Yes. Standard K-means uses unweighted means while weighted K-means uses weighted means. Since the weights are not uniform, the centroids will be different, leading to different cluster boundaries and thus different assignments.

(c)

Yes, weighted K-means always converges in finite steps if all weights are positive and the dataset is finite.

The algorithm minimizes:

$$J = \sum_{j=1}^{k} \sum_{i \in C_j} w_i |x_i - \mu_j|^2$$

where $w_i > 0$ and μ_j is the weighted mean of cluster j.

The algorithm alternates between:

1. Assignment step: Assign each point to nearest centroid (minimizes J for fixed centroids)
2. Update step: Update centroids to weighted means (minimizes J for fixed assignments)

Since:

- J never increases after each step
- J is bounded below by 0 (sum of weighted squared distances)
- The number of possible cluster assignments is finite

The algorithm must eventually reach a state where assignments and centroids don't change, thus converging in finite steps.