APPIAH PRINCE

HOMEWORK 3; STAT 5474

10|15|2021

PART I

$$W(C) = \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k}\sum_{i'\in C_k} \|x_i - x_{i'}\|_2^2$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k}\sum_{i'\in C_k} \|(x_i - \bar{x}_k) + (\bar{x}_k - x_{i'})\|_2^2$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k}\sum_{i'\in C_k} \left[\|x_i - \bar{x}_k\|_2^2 + \|\bar{x}_k - x_{i'}\|_2^2 + 2(x_i - \bar{x}_k)^T(\bar{x}_k - x_{i'})\right]$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k} \left[\sum_{i'\in C_k}\|x_i - \bar{x}_k\|_2^2 + \sum_{i'\in C_k}\|\bar{x}_k - x_{i'}\|_2^2 + 2(x_i - \bar{x}_k)^T\sum_{i'\in C_k}(\bar{x}_k - x_{i'})\right]$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k} \left[n_k\|x_i - \bar{x}_k\|_2^2 + \sum_{i'\in C_k}\|\bar{x}_k - x_{i'}\|_2^2 + 2(x_i - \bar{x}_k)\left(\sum_{i'\in C_k}\bar{x}_k - \sum_{i'\in C_k}x_{i'}\right)\right]$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k} \left[n_k\|x_i - \bar{x}_k\|_2^2 + \sum_{i'\in C_k}\|\bar{x}_k - x_{i'}\|_2^2 + 2(x_i - \bar{x}_k)\left(\overbrace{n_k\bar{x}_k - n_k\bar{x}_k}^{0}\right)\right]$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k} \left[n_k\|x_i - \bar{x}_k\|_2^2 + \sum_{i'\in C_k}\|x_i - \bar{x}_k\|_2^2 + 0\right]$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k} \left[n_k\|x_i - \bar{x}_k\|_2^2 + n_k\|x_i - \bar{x}_k\|_2^2\right]$$

$$= \frac{1}{2}\sum_{k=1}^{K}\sum_{i\in C_k} 2n_k\|x_i - \bar{x}_k\|_2^2 = \sum_{k=1}^{K}\sum_{i\in C_k} n_k\|x_i - \bar{x}_k\|_2^2$$

$$\therefore W(C) = \sum_{k=1}^{K} n_k \sum_{i\in C_k}\|x_i - \bar{x}_k\|_2^2$$ as required.