In section 14.3.6, we say that expression (14.28) for the within cluster sum of squares

$$W(C) = \frac{1}{2}\sum_{k=1}^{K}\sum_{C(i)=k}\sum_{C(i')=k} d(x_i, x_{i'}).$$

can be written as expression (14.31):

$$\begin{aligned} W(C) &= \frac{1}{2}\sum_{k=1}^{K}\sum_{C(i)=k}\sum_{C(i')=k} ||x_i - x_{i'}||^2 \\ &= \sum_{k=1}^{K} N_k \sum_{C(i)=k} ||x_i - \bar{x}_k||^2. \end{aligned}$$

This is correct, but the latter is not the criterion that is minimized by the usual form of K-means clustering (Algorithm 14.1). Instead, K-means clustering minimizes the criterion

$$W(C) = \sum_{k=1}^{K}\sum_{C(i)=k} ||x_i - \bar{x}_k||^2. \tag{1}$$

That is, there is no cluster size weight N_k, where N_k is the number of observations assigned to cluster k. Intuitively, this occurs because (14.28) effectively gives a weight proportional to N_k^2 to a cluster with N_k members. One could define a modified version of (14.28) that is equivalent to expression (1) by including weights $1/N_k$.

Now in practice one could instead minimize (14.31): this would change the assignment step (2) in Algorithm 14.1, which would no longer simply assign a point to the cluster with the nearest centroid. Qualitatively, this would tend to produce clusters with more equal cluster sizes than the standard algorithm. This could in fact be a useful variation on K-means clustering.

Thanks to Daniela Witten for finding this error.