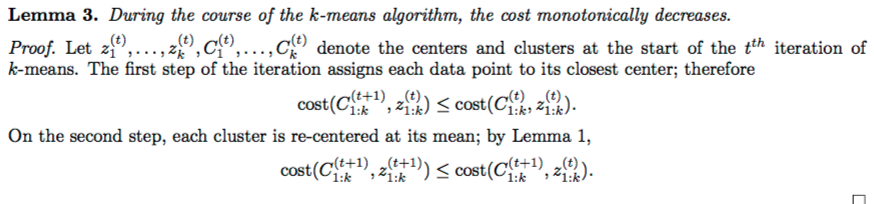


First, RSS decreases in the reassignment step since each vector is assigned to the closest centroid, so the distance it contributes to  decreases. Second, it decreases in the recomputation step because the new centroid is the vector  for which  reaches its minimum.



Thus, we minimize  when the old centroid is replaced with the new centroid. , the sum of the , must then also decrease during recomputation.

Since there is only a finite set of possible clusterings, a monotonically decreasing algorithm will eventually arrive at a (local) minimum.