1. (a) Its easy to see that an optimal clustering is not necessary unique. For instance, if our points are in $\mathbb{R}^2$, drawn uniformly from $\{(0,0),(0,1),(1,0),(1,1)\}$m, then the optimal risk for two clusters is clearly $R(C^*) = .25$. However, this risk can be achieved by two different $\hat{C}$, $\{(.5,0),(.5,1)\}$ or $\{(0,.5),(1,.5)\}$, making C^* not unique.

   (b) Let C_k^* achieve the minimal risk $R^{(k)}$. Now let $C_{k+1} = C_k^* \cup \{c_{k+1}\}$ for some c_{k+1} in the space of X. For all X,

$$\Rightarrow \forall c_i \in C_k^* \ : \ \min_{c_j \in C_{k+1}} \|X - c_j\| \leq \|X - c_i\|$$
$$\Rightarrow \min_{c_j \in C_{k+1}} \|X - c_j\| \leq \min_{c_i \in C_k} \|X - c_i\|$$
$$\Rightarrow R(C_k^*) \geq R(C_{k+1})$$

   Since, by definition, $R(C_k^*) = R^{(n)}$ and $R(C_{k+1}) \geq R^{(n+1)}$, we have

$$R^{(n)} = R(C_k^*) \geq R(C_{k+1}) \geq R^{(n+1)}$$

   So $R^{(n)}$ is non increasing in n.