Section 4.

1. $\mu_1 = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{x_1 + \cdots + x_n}{n}$

$\mu_2 = \frac{1}{N-n} \sum_{i=n+1}^{n} x_i = \frac{x_{n+1} + \cdots + x_N}{N-n}.$

2. No.

~~It is hard assignment~~, The distortion function $J = \frac{1}{n} \sum_{i=1}^{n} \|x_i - \mu_k\|^2$, which is not always a convex function. And it does not consider data in other clusters.

3. Assign each data point to the nearest cluster, i.e. to the cluster that gives the smallest squared distance.

The aim function is $\min_{(C_1, \cdots, C_k)} \sum_{i=1}^{N} \|(x_i - \mu_k)^2$ for all the clusters μ_k.

Thus for each data point x_i, the nearest cluster will give the smallest distance $(x_i - \mu_k)^2$.

4. Yes.

Each data considers all the clusters, and can always give the smallest distance.

扫描全能王 创建

GMM

1. C_2.

That means $x_m - \mu_1 < x_m - \mu_2$.

$$\therefore \gamma_{m2} = \frac{(x_m - \mu_2)^2}{(x_m - \mu_1)^2 + (x_m - \mu_2)^2} > \gamma_{m1} = \frac{(x_m - \mu_1)^2}{(x_m - \mu_1)^2 + (x_m - \mu_2)^2}$$

2. Yes.

For hard assignment, $\mathcal{L} = \sum_{i=1}^{N} (x_i - \mu_k)^2$,

For soft assignment, $\mathcal{L} = \sum_{i=1}^{N} (\gamma_{m1} \times (x_m - \mu_1)^2 + \gamma_{m2} \times (x_m - \mu_2)^2)$

$$= 2\sum_{i=1}^{N} \frac{(x_m - \mu_1)(x_m - \mu_2)}{(x_m - \mu_1)^2 + (x_m - \mu_2)^2}$$

扫描全能王 创建