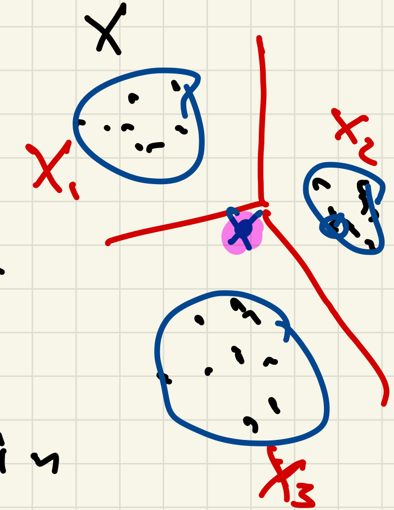
FoDA L24

Mixture of Gaussians

Soft Clustering



Input: Set $X \subset \mathbb{R}^d$, value k

Goal: Decomposition of X in disjoint sets $X_1, X_2 \dots X_k$

$X_j \subset X$ $\quad X_i \cap X_j = \emptyset$

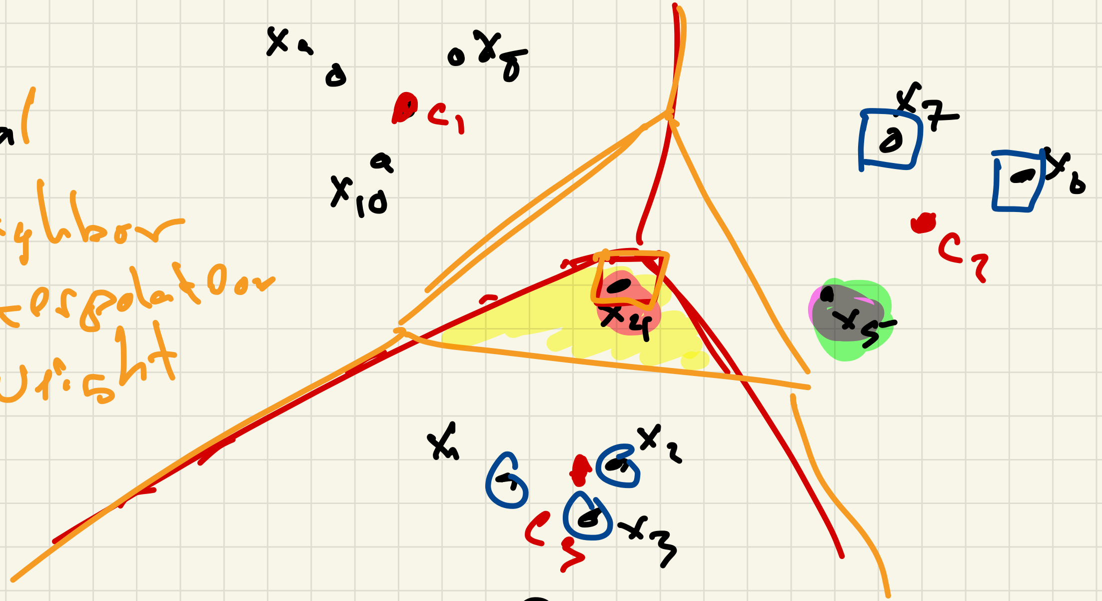
the clusters

Soft Goal: Assignment of each $x \in X$ to each cluster w/ some probability

$$\Pr(x \in X_j) = p_j(x) \in [0,1]$$

Natural

Neighbor Interpolation weight



$$\Pr[x_1 \in X_3] = 1$$

	1	2	3	4	5	6	7	8	9	10
Hard	3	3	3	3	2	2	2	1	1	1
Soft 1	0	0	0	.3	0	0	0	1	1	1
Soft 2	0	0	0	.3	.8	1	1	0	0	0
Soft 3	1	1	1	.4	.2	0	0	0	0	0

$$w_3(x_4) = \Pr[x_4 \in X_3] = 0.4$$

$$w_2(x_4) = 0.3$$

$$w_1(x_4) = 0.3$$

Mixture of Gaussians

Normal 1-d Gaussian

$\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

$(x-\mu)^T(x-\mu)$
$= \|x-\mu\|^2$

Input $X \subset \mathbb{R}^d$, param k

$X = \{x_1, \dots x_n\}$

Goal A set of k Gaussians

$$\text{maximize} \quad \prod_{x \in X} \sum_{j=1}^{k} w_j(x) \underbrace{f_{\mu_j, \Sigma_j}(x)}$$

likelihood of x given Gaussian (μ_j, Σ_j)

d-dim Gaussian pdf

$$f_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{d/2}} \cdot \frac{1}{\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2} \underset{1\times d}{(x-\mu)^T} \underset{d\times d}{\Sigma^{-1}} \underset{d\times 1}{(x-\mu)}\right)$$

$\mu \in \mathbb{R}^d$
$\Sigma \in \mathbb{R}^{d\times d}$

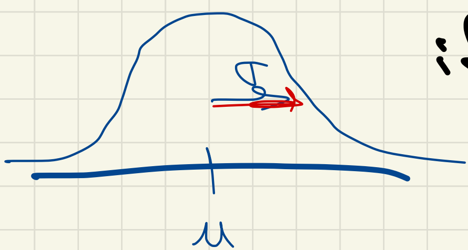
$$d_M(x, \mu) = \sqrt{(x-\mu)^T M (x-\mu)}$$ Mahalanobis distance

$$d_{\Sigma_1^{-1}}(x,\mu)^2 = (x-\mu)^T \Sigma_1^{-1} (x-\mu)$$

if $\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ identity then $d_{\Sigma_1^{-1}}(x,\mu) = \|x-\mu\|$

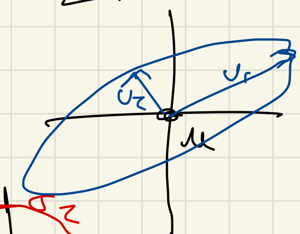
$\|x-\mu\|$

$\Sigma_1 = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$



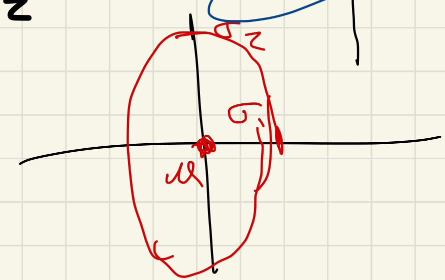
if $\Sigma_1 = \begin{bmatrix} \sigma^2 & & 0 \\ & \sigma^2 & \\ 0 & & \ddots \; \sigma^2 \end{bmatrix}$

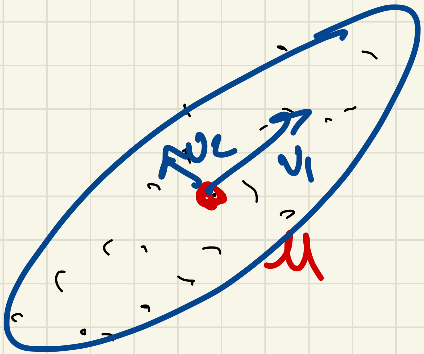
$$d_{\Sigma_1^{-1}}(x,\mu)^2 = \frac{\|x-\mu\|^2}{\sigma^2}$$



if $\Sigma_1 = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots \; \sigma_d^2 \end{bmatrix}$

$d = 2$



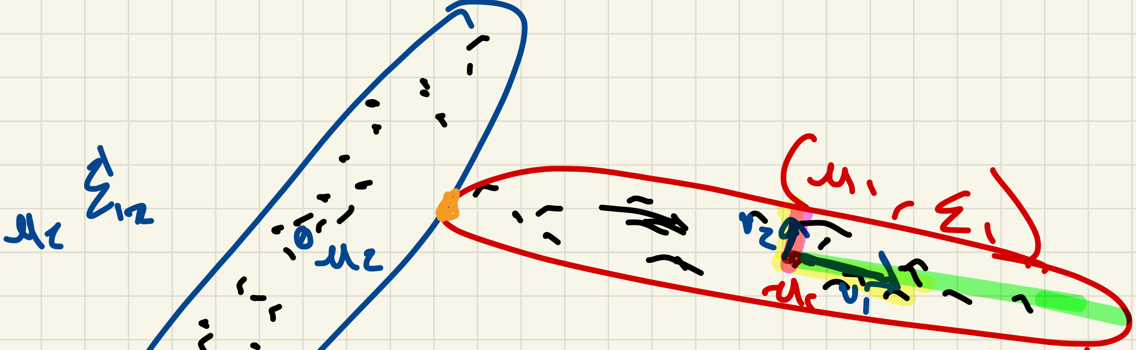


one cluster

Representation of one cluster

μ - center

Σ - covariance



new point

$\Sigma_1 \in \mathbb{R}^{d\times d}$

$\Sigma_1 = V S V^T$ eigen decomp.

$\Sigma_1 = \cancel{U S V^T}$

$\Sigma_1 = X_1^T X_1$

$V^T = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$S = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$

μ_3, Σ_3

EM Algorithm for MoG

0. Init Choose k pts $S \subset X$ $S = \{\mu_1, \mu_2, \dots \mu_k\}$

for all $x \in X$ $w_i(x) = 1$ for $\phi_S(x) = \mu_i$, $w_i(x) = 0$ other i

1. repeat

1a. for all $i \in [1 \dots k]$ ← update model for each cluster i

A $W_i = \sum_{x \in X} w_i(x)$ ← total weight of cluster i

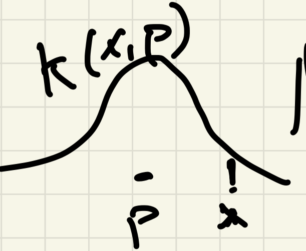
B $\mu_i = \frac{1}{W_i} \sum_{x \in X} w_i(x) x$ ← set mean

C $\Sigma_i' = \frac{1}{W_i} \sum_{x \in X} \underbrace{w_i(x)}_{\text{update weight}} \underbrace{\overbrace{(x - \mu_i)(x - \mu_i)^T}^{d \times d}}_{\text{centering}}$ $\Sigma_i' = \tilde{X}_i \tilde{X}_i^T$

1b. for all $x \in X$

$l_i(x) = f_{\mu_i, \Sigma_i}(x) \quad \forall i \in [1 \dots k]$ $\in (0, \infty)$

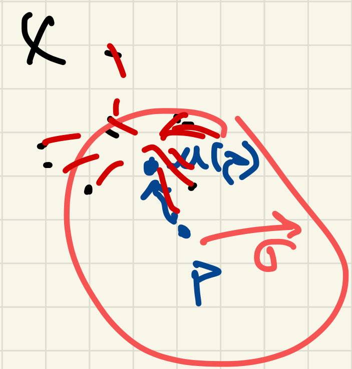
$w_i(x) = l_i(x) \Big/ \left(\sum_{i=1 \dots k} l_i(x)\right)$ $\in (0, 1)$



Mean Shift Clustering

Input $X \subset \mathbb{R}^d$

Kernel $K(p,q) = \exp\left(-\frac{\|p-q\|^2}{2\sigma^2}\right)$



repeat

1. $\forall p \in X$

$$\mu(p) = \frac{\sum'_{x \in X} K(x,p)\, x}{\sum'_{x \in X} K(x,p)}$$

$\forall p \in X$

2. Set $p \leftarrow \mu(p)$