DBSCAN

Point	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	25	80

1 calculate distances

$\varepsilon = 2$

	A	B	C	D	E	F
A	0	1	1.41	8.60	9.21	81.60
B	1	0	1	7.81	8.48	81.32
C	1.41	1	0	7.21	7.81	80.36
D	8.60	7.81	7.21	0	1	74.95
E	9.21	8.48	7.81	1	0	73.97
F	81.60	81.32	80.36	74.95	73.97	0

2 randomly select core point & start clustering

- $\varepsilon = 2$ and $pts_{min} = 3$

Point	# of pts			
A	2	X	⇒	noise (start here)
B	2	X	⇒	noise
C	2	X	⇒	noise
D	1	X	⇒	noise
E	1	X	⇒	noise
F	0	X	⇒	noise

- $\varepsilon = 2$ and $pts_{min} = 2$

Point	# of pts			
A	2	✓	⇒	core (start here)
B	2	✓	⇒	core
C	2	✓	⇒	core
D	1	x	⇒	noise
E	1	x	⇒	noise
F	0	x	⇒	noise

$\varepsilon = 8$

	A	B	C	D	E	F
A	0	1	1.41	8.60	9.21	81.60
B	1	0	1	7.81	8.48	81.32
C	1.41	1	0	7.21	7.81	80.36
D	8.60	7.81	7.21	0	1	74.95
E	9.21	8.48	7.81	1	0	73.97
F	81.60	81.32	80.36	74.95	73.97	0

- $\varepsilon = 8$ and $pts_{min} = 3$

Point	# of pts				
A	2	x	⇒	border	
B	3	✓	⇒	core	
C	4	✓	⇒	core (start here)	A B D E
D	3	✓	⇒	core	
E	2	x	⇒	border	
F	0	x	⇒	noise	

soft clustering

deriving updates for μ, σ^2, p

goal: maximize $P(S_n|\theta) = \prod_{i=1}^{n} p(x_i|\theta)$

$$= \prod_{i=1}^{n} \sum_{j=1}^{K} p_j \, \mathcal{N}(x_i, \mu_j, \sigma_j^2)$$

p_j: each class

assuming 1 class $\therefore p_1 = 1$

$$p(S_n|\theta) = \prod_{i=1}^{n} \mathcal{N}(x_i, \mu, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-1}{2\sigma^2}(x_i - \mu)^2\right)$$

to make this cleaner, log both sides:

$\ln(ab) = \ln(a) + \ln(b)$

$$\ln\left(P(S_n|\theta)\right) = \sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2}(x_i - \mu)^2$$

$\ln\left(\frac{1}{a}\right) = \ln(a^{-1}) = -\ln(a)$

$\ln\left((2\pi\sigma^2)^{-1/2}\right) = -\frac{1}{2}\ln(2\pi\sigma^2)$

$$= \frac{-1}{2}\sum_{i=1}^{n} \ln(2\pi\sigma^2) - \frac{1}{2}\sum_{i=1}^{n} \frac{(x-\mu)^2}{\sigma^2}$$

deriving w.r.t. μ

$$\frac{d}{d\mu} \ln\left(P(S_n|\theta)\right) = 0 + \frac{2}{2} \sum_{i=1}^{n} \frac{(x-\mu)}{\sigma} \quad (-1)$$

$$\frac{1}{\sigma} \sum_{i=1}^{n} (x-\mu) = 0$$

$$\sum_{i=1}^{n} x - \sum_{i=1}^{n} \mu = 0$$

$$(n)(\mu) = \sum_{i=1}^{n} x$$

$$\boxed{\mu = \frac{\sum_{i=1}^{n} x}{n}}$$

deriving w.r.t. σ^2

$$\ln\left(P(S_n|\theta)\right) = -\frac{1}{2} \sum_{i=1}^{n} \ln\left(2\pi\sigma^2\right) - \frac{1}{2} \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2}$$

$$\frac{d}{d\sigma^2} \ln\left(P(S_n|\theta)\right) = -\frac{1}{2} \sum_{i=1}^{n} \frac{2\pi}{2\pi\sigma^2} - \frac{1}{2} \sum_{i=1}^{n} \frac{-(x_i-\mu)^2}{(\sigma^2)^2}$$

$$\frac{d}{dx} \ln(x) = \frac{x'}{x}$$

$$\frac{d}{dx} \frac{1}{x} = \frac{-1}{x^2}$$

$$0 = -\frac{1}{2} \sum_{i=1}^{n} \frac{1}{\sigma^2} - \frac{1}{2} \sum_{i=1}^{n} \frac{-(x_i-\mu)^2}{(\sigma^2)^2}$$

$$0 = \frac{1}{\sigma^2} \sum_{i=1}^{n} 1 - \frac{1}{(\sigma^2)^2} \sum_{i=1}^{n} (x_i-\mu)^2$$

$$0 = \frac{1}{\sigma^2}(n) - \frac{1}{(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

example (gaussian mixture model)

$x = [1 \quad 2.5 \quad 3.5 \quad 4 \quad 5.5]$

$K = 2$

(initial values)

2 clusters

- C_1 : $\mu_1 = 2$ $\quad \sigma_1^2 = 2$ $\quad p_1 = 0.5$
- C_2 : $\mu_2 = 4.5$ $\quad \sigma_2^2 = 2$ $\quad p_2 = 0.5$

E - step compute $p(j|i)$ $\forall$ classes j & $\forall$ datapoints

$$\mathcal{N}(x_i \mid \mu_j, \sigma_j^2) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{1}{2\sigma_j}(x_i - \mu_j)^2\right)$$

$$p(j|i) = \frac{p_j \, \mathcal{N}(x_i \mid \mu_j, \sigma_j^2)}{\sum_{j=1}^{k} p_j \, \mathcal{N}(x_i \mid \mu_j, \sigma_j^2)}$$

do the calculations ...

x_i	$p(C_1 \mid x_i)$	$p(C_2 \mid x_i)$
1	0.963	0.037
2.5	0.689	0.311
3.5	0.311	0.689
4	0.421	0.579
5.5	0.037	0.963

M-step update the parameters μ_j, σ_j^2, p_j

$$\hat{p}_j = \frac{\hat{n}_j}{n} = \frac{\sum_{i=1}^{N} P(j|i)}{n}$$

$\hat{p}_1 = 0.4742$ $\hat{p}_2 = 0.5158$

$$\hat{\mu}_j = \frac{1}{\hat{n}_j} \sum_{i=1}^{n} p(j|i) \cdot x_i$$ compute...

$$\hat{\sigma}_j^2 = \frac{1}{\hat{n}_j} \sum_{i=1}^{n} p(j|i) \| x_i - \mu_j \|_2^2$$ compute...