Machine learning – Ex3

Matthis Brocheton

Problem 1

Dataset:

$A1 = (2,\ 3)$
$A2\ = (7,\ 0.5)$
$A3 = (11,\ 1)$
$A4 = (4,\ \ -0.5)$
$A5 = (10,\ 2.5)$

A6= $(3,\ 2)$

$A7 = (7,\ 0)$

$A8 = (9,\ 1.5)$

$A9 = (4,\ 2.5)$

$A10 = (10,\ 1)$

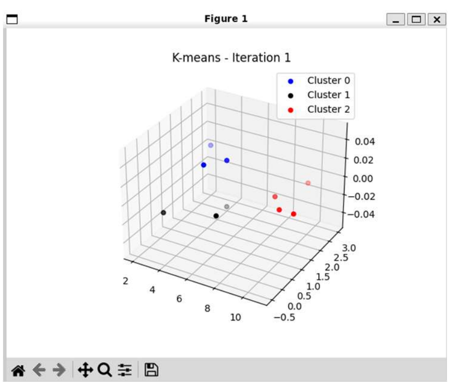
Center

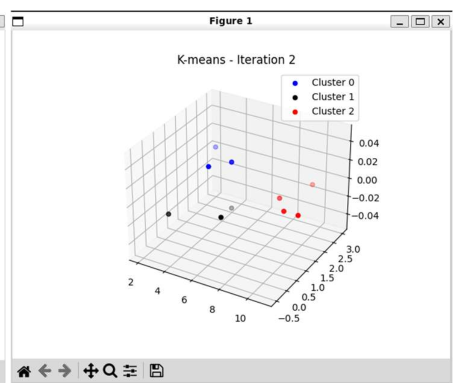
$C1 =\ A1 = (2,\ 3)$

$C2 = A4 = (4,\ \ -0.5)$

$C3 = A5 = (10,\ 2.5)$

a)





This is the the result for the process for the first epoch:

Datapoint	Ditance to:			Cluster
	(2, 3)	(4, -0.5)	(10, 2.5)	
$A1 = (2,\ 3)$	0.0	4.03	8.02	C1
$A2\ = (7,\ 0.5)$	5.59	3.16	3.61	C2
$A3 = (11,\ 1)$	9.22	7.16	1.8	C3
$A4 = (4,\ \ -0.5)$	4.03	0.0	6.71	C2
$A5 = (10,\ 2.5)$	8.02	6.71	0.0	C3
A6= $(3,\ 2)$	1.41	2.69	7.02	C1
$A7 = (7,\ 0)$	5.83	3.04	3.91	C2
$A8 = (9,\ 1.5)$	7.16	5.39	1.41	C3
$A9 = (4,\ 2.5)$	2.06	3.0	6.0	C1

$A10 = (10,\ 1)$		8.25	6.18	1.5		C3

With the formula :

$$d(p1, p2) = \sqrt{(x_2 + x_1)^2 + (y_2 - y_1)^2}$$

Exemple use :

$d(A1, C1) = \sqrt{(2+2)^2 + (3-3)^2} = 0$

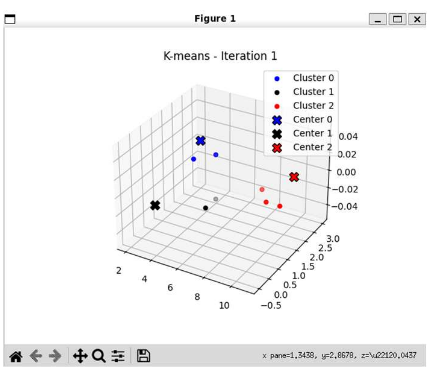
Epoch 1 :

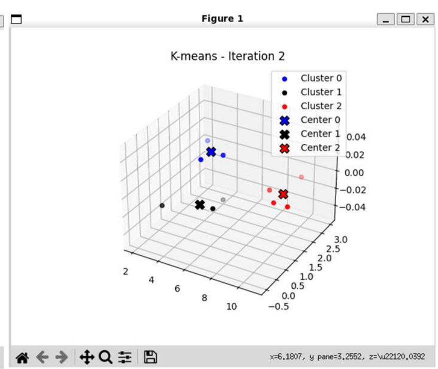
- **Cluster 1** : A1, A6, A9
- **Cluster 2** : A2, A4, A7
- **Cluster 3** : A3, A5, A8, A10

Epoch 2 :

- **Cluster 1** : A1, A6, A9
- **Cluster 2** : A2, A4, A7
- **Cluster 3** : A3, A5, A8, A10

b)





how to calculate the cluster center:

$$\frac{C1 + C1 + C1}{3}$$

$$\frac{C2 + C2 + C2}{3}$$

$$\frac{C3 + C3 + C3}{3}$$

Epoch 1 :

- **Cluster center 1**: (2, 3)
- **Cluster center 2** : (4, -0.5)
- **Cluster center 3** : (10, 2.5)

Epoch 2 :

- **Cluster center 1** : (3.0, 2.5)
- **Cluster center 2** : (6.0, 0.0)
- **Cluster center 3** : (10.0, 1.5)

c)

it converge after 1 epoch.

Problem 2

a)

$$L(p \mid \mathbf{x}) = \prod_{i=1} p^{x_i}(1-p)^{1-x_i}$$

$x = (x1, x2, \dots, xn), xi \in \{0,1\}$

$logL(\, p \mid x\,) = \sum \log\,(p^{x_1}(1-p)^{1-x_1}))$

Rules :

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

$\log(a^b) = b * log(a)$

$\frac{d}{dp}\log(p) = \frac{1}{p}$

$logL(\, p \mid x\,) = \sum[x_i log(p) + (1-x_i)log(1-p)]$

$x_i \log(p) \to x_i * \frac{1}{p}$

$(1-x_i)log(1-p) \to (1-x_i) * -\frac{1}{1-p}$

$(1-x_i) * -\frac{1}{1-p} \to -\frac{1-x_i}{1-p}$

$logL(\, p \mid x\,) = \frac{1}{p}\sum(x_i) - \frac{1}{1-p}\sum(1-x_i)$

$S = \sum x_i \to success$

$n - S = \sum(1-x_i) \to fails$

Results:

Derivative of log-vraisemblance:

$\frac{dp}{d} logL(\, p \mid x\,) = \frac{S}{p} - \frac{n-S}{1-p}$

b)

Derivative of log-vraisemblance

$$\frac{dp}{d} logL(p \mid x) = \frac{S}{p} - \frac{n-S}{1-p}$$

$$\frac{dp}{d} logL(p \mid x) = 0$$

$$\frac{\mathrm{S}}{\mathrm{p}} - \frac{\mathrm{n}-\mathrm{S}}{1-\mathrm{p}} = 0$$

Make a cross product:

$$\mathrm{S}(1-\mathrm{p}) = (\mathrm{n}-\mathrm{S})\mathrm{p}$$

$$S - Sp = np - Sp$$

$$S - \cancel{Sp} = np - \cancel{Sp}$$

$$S = np$$

$$p = \frac{S}{n}$$

Lettres translate:

$$S = \sum x_i \rightarrow success$$

$$n - S = \sum (1 - x_i) \rightarrow fails$$

Results:

$$p = \frac{S}{n} \rightarrow \frac{n}{1} \sum x_i$$

$$\hat{p}_{\mathrm{ML}} = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Problem 3

a)

$y_i = a + bx_i + \epsilon_i \; for \; i = 1,2, \dots, n, and \; \epsilon i \sim N(0, \sigma 2)$

Joint density of independent variables:

$$J(\epsilon 1, \dots, \epsilon n) = i = \prod \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{\epsilon_i^2}{2\sigma^2})$$

Rules:

$$i = \prod exp(ai) \; = exp(a1) \cdot exp(a2) \rightarrow \; exp(a_n) = exp(\sum a_i)$$

$$J = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot exp(-\frac{1}{2\sigma^2} \sum \epsilon_i^2)$$

We want to find ϵ_i:

$y_i = a + bx_i + \epsilon_i \rightarrow \; \epsilon_i = y_i - a - bx_i$

So, the result is:

$$J(a, b, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^n \cdot exp(-\frac{1}{2\sigma^2} \sum (y_i - a - bx_i)^2)$$

b)

<u>Criteria of the smallest squares</u>

$$\sum (yi - \hat{y}^i)^2 = \sum \left(y_i - (a + bx_i)\right)^2$$

<u>maximizing the natural logarithm of J</u>

$$J(a, b, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^n \cdot exp(-\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2)$$

<u>Rules:</u>

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

$$\log\left(a^b\right) = b * log(a)$$

$$\frac{d}{dp}\log(p) = \frac{1}{p}$$

$$\log(x^n) = nlogx$$

$$log(e^x) = x$$

$$\sqrt{A} \Rightarrow A^{-\frac{1}{2}}$$

$$\frac{1}{\sqrt{A}} = A^{-\frac{1}{2}}$$

$$\log\left(\exp\left(f(x)\right)\right) = f(x)$$

$$log\ \mathrm{J(a, b, \sigma^2)} = \log\left(\left(\frac{1}{2\pi\sigma^2}\right)^n \cdot \exp(-\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2)\right)$$

$$\log \mathrm{J(a, b, \sigma^2)} = \log\left(\left(\frac{1}{2\pi\sigma^2}\right)^n\right) + \log\left(\exp(-\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2)\right)$$

$$log\ \mathrm{J(a, b, \sigma^2)} = \log\left((2\pi\sigma^2)^{\frac{-n}{2}}\right) + -\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2$$

$$\log \mathrm{J(a, b, \sigma^2)} = -\frac{n}{2} log(2\pi\sigma^2) + -\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2$$

Don’t need $-\frac{n}{2}(2\pi\sigma^2)$.

So, the result is:

$$\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2$$

$$\frac{1}{2\sigma^2}\sum (y_i - a - bx_i)^2 = \sum \left(y_i - (a + bx_i)\right)^2$$

c)

maximizing the natural logarithm of J:

$$\log \mathrm{J}(\mathrm{a}, \mathrm{b}, \sigma^2) = \log \left(\left(\frac{1}{2\pi\sigma^2}\right)^{\mathrm{n}} \cdot \exp(-\frac{1}{2\sigma^2}\sum(\mathrm{y_i} - \mathrm{a} - \mathrm{bx_i})^2)\right)$$

$$\log \mathrm{J}(\mathrm{a}, \mathrm{b}, \sigma^2) = -\frac{n}{2} log(2\pi\sigma^2) + -\frac{1}{2\sigma^2}\sum(\mathrm{y_i} - \mathrm{a} - \mathrm{bx_i})^2$$

Rules:

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

$\log(a^b) = b * log(a)$

$\frac{d}{dp}\log(p) = \frac{1}{p}$

$\log(x^n) = nlogx$

$log(e^x) = x$

$\sqrt{A} \Rightarrow A^{-\frac{1}{2}}$

$\frac{1}{\sqrt{A}} = A^{-\frac{1}{2}}$

$\log(\exp(f(x))) = f(x)$

$log(x) = \frac{1}{x}$

$(\frac{x}{1}) = -\frac{x}{x^2}$

$-\frac{2}{n} log(2\pi) = 0$

$$\log \mathrm{J}(\mathrm{a}, \mathrm{b}, \sigma^2) = -\frac{n}{2} log(2\pi\sigma^2) + -\frac{1}{2\sigma^2}\sum(\mathrm{y_i} - \mathrm{a} - \mathrm{bx_i})^2$$

$\sum(\mathrm{y_i} - \mathrm{a} - \mathrm{bx_i})^2 = S$ (we know a and b)

$$\log \mathrm{J}(\mathrm{a}, \mathrm{b}, \sigma^2) = -\frac{n}{2} log(2\pi\sigma^2) + \frac{S}{2\sigma^2}$$

$$-\frac{n}{2} log(2\pi\sigma^2) \to -\frac{n}{2} \cdot \frac{1}{\sigma^2}$$

$$\frac{S}{2\sigma^2} + \frac{1}{\sigma^2} \to \frac{S}{2(\sigma^2)^2}$$

$$\log \mathrm{J}(\mathrm{a}, \mathrm{b}, \sigma^2) = -\frac{n}{2} + \frac{S}{2(\sigma^2)^2}$$

$$\log \mathrm{J}(\mathrm{a}, \mathrm{b}, \sigma^2) = \frac{-\frac{n}{2} + \frac{S}{2(\sigma^2)^2}}{2(\sigma^2)^2}$$

$$-n\sigma^2 + S = 0$$

$$S = -n\sigma^2$$

$$\sigma^2 = \frac{S}{n}$$

$$\sigma^2 = \frac{S}{n} = \frac{1}{\mathrm{n}} \sum (\mathrm{y_i} - \mathrm{a} - \mathrm{bx_i})^2$$

The result is:

$$\sigma^2 = \frac{1}{\mathrm{n}} \sum (\mathrm{y_i} - \mathrm{a} - \mathrm{bx_i})^2$$