ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

Q1 Manual Calcul^n of one round of EM for a GMM

M step

1] write down the likelihood function you are trying to optimize:

$$Q(\theta, \theta^{(t-1)}) \triangleq \sum_i \sum_k r_{ik} \log \pi_k + \sum_i \sum_k r_{ik} \log p(x_i | \theta_k)$$

2] $\pi_k = \frac{1}{N} \sum_i r_{ic}$ — formula from Murphy book

$\therefore \pi_1 = \frac{1}{3} \left(\sum_{i=1}^{3} r_{ik} \right)$

$\pi_1 = \frac{1}{3} \left(1 + 0.3 + 0 \right) = \frac{1.3}{3} = 0.43333$

$\pi_2 = \frac{1}{3} \left(0 + 0.7 + 1 \right) = \frac{1.7}{3} = 0.56667$

3] $\mu_c = \frac{\sum_{i=1}^{3} r_{ic} x_i}{r_c}$ — from Murphy's book

$\mu_1 = \frac{\sum_{i=1}^{3} r_{i1} x_i}{r_1} = \frac{1 \times 1 + 10 \times 0.3 + 0 \times 0.7}{1.3}$

$\mu_1 = \frac{4}{1.3} = 3.0769$

$\mu_2 = \frac{\sum_{i=1}^{3} r_{i2} x_i}{r_2} = \frac{0 \times 1 + 0.7 \times 10 + 1 \times 20}{1.7} = \frac{27}{1.7} = 15.8823$

$$\boxed{\mu_1 = 3.0769 \quad \mu_2 = 15.8823}$$

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

(4) $\Sigma_c = \dfrac{\sum_{i=1}^{N} r_{ic} x_i x_i^T}{r_c} - \mu_c \mu_c^T$ — from Murphy's book

$\therefore \Sigma_1 = \dfrac{\sum_{i=1}^{3} r_{i1} x_i x_i^T}{r_1} - \mu_1 \mu_1^T$

$\Sigma_1 = \dfrac{1\times1\times1 + 0.3\times10\times10}{1.3} - (3.0769)^2 = \dfrac{31}{1.3} - (3.0769)^2$

$\Sigma_1 = 14.378840 \quad \therefore \quad \sigma_1 = \sqrt{\Sigma_1}$

$\sigma_1 = \sqrt{14.378840}$

$\boxed{\sigma_1 = 3.7919}$

$\Sigma_2 = \dfrac{\sum_{i=1}^{3} r_{i2} x_i x_i^T}{r_2} - \mu_2 \mu_2^T$

$\Sigma_2 = \dfrac{0 + 0.7\times100\times100 + 1\times400}{1.7} - (15.8823)^2$

$\Sigma_2 = \dfrac{70 + 400}{1.7} - (15.8823)^2 = \dfrac{470}{1.7} - (15.8823)^2$

$\Sigma_2 = 24.223 \quad \therefore \quad \boxed{\sigma_2 = 4.92170}$

Summary

$\pi = \begin{bmatrix} 0.4333 \\ 0.5667 \end{bmatrix}$ (2) $\quad \mu = \begin{bmatrix} 3.0769 \\ 15.8823 \end{bmatrix}$ (3) $\quad \sigma = \begin{bmatrix} 3.7919 \\ 4.92170 \end{bmatrix}$ (4)

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

Q1 E Step

1. formula for the probability of observation x_i belonging to cluster c

$$r_{ic} = \frac{\pi_c \, P(x_i \mid \theta_c^{(t-1)})}{\sum_{c'} \pi_{c'} \, P(x_i \mid \theta_{c'}^{(t-1)})}$$

As, given is mixture of gaussian.

$$\therefore P(x_i \mid \theta_c^{(t-1)}) = \frac{1}{\sqrt{2\pi}\,\sigma_c} \exp\left(\frac{-(x_i - \mu_c)^2}{2\sigma_c^2}\right)$$

$$P(x_1 \mid \theta_1) = \frac{1}{\sqrt{2\pi}\, 3.7919} \exp\left(\frac{-(1 - 3.0769)^2}{2(3.7919)^2}\right)$$

$$= \frac{1}{9.5047} \times \exp(-0.1499) = 0.09056$$

$$\boxed{P(x_1 \mid \theta_1) = 0.09056}$$

$$P(x_2 \mid \theta_1) = \frac{1}{\sqrt{2\pi}\, 3.7919} \exp\left(\frac{-(10 - 3.0769)^2}{2(3.7919)^2}\right)$$

$$= \frac{1}{9.5047} \exp(-1.6667) = 0.01987$$

$$\boxed{P(x_2 \mid \theta_1) = 0.01987}$$

$$P(x_3 \mid \theta_1) = \frac{1}{\sqrt{2\pi}\, 3.7919} \exp\left(\frac{-(20 - 3.0769)^2}{2(3.7919)^2}\right)$$

$$= \frac{1}{9.5047} \exp(-9.9590) = 4.97 \times 10^{-6}$$

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

$$P(x_1|\theta_2) = \frac{1}{\sqrt{2\pi}\, 4.92170} \exp\left(\frac{-(1-15.8823)^2}{2 \times 4.92170^2}\right)$$

$$= 0.08105 \times \exp(-4.5715)$$

$$\boxed{P(x_1|\theta_2) = 8.3825 \times 10^{-4}}$$

$$P(x_2|\theta_2) = \frac{1}{\sqrt{2\pi}\, 4.92170} \exp\left(\frac{-(10-15.8823)^2}{2 \times 4.92170^2}\right)$$

$$= 0.039684$$

$$\boxed{P(x_2|\theta_2) = 0.039684}$$

$$P(x_3|\theta_2) = \frac{1}{\sqrt{2\pi}\, 4.92170} \exp\left(-\frac{(20-15.8823)^2}{2 \times 4.92170^2}\right)$$

$$= 0.08105 \times \exp(-0.350035)$$

$$= 0.057114$$

$$\boxed{P(x_3|\theta_2) = 0.057114}$$

Probability table for Reference,

$x \downarrow$ \ $\theta \rightarrow$	1	2
1	0.09056	8.3825×10^{-4}
2	0.01987	0.039684
3	4.97×10^{-6}	0.057114

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

$$r_{ic} = \frac{\pi_c \, p(x_i \mid \theta_c^{(t-1)})}{\sum_{c'} \pi_{c'} \, p(x_i \mid \theta_{c'}^{(t-1)})}$$

$$r_{11} = \frac{\pi_1 \, p(x_1 \mid \theta_1^{(t-1)})}{\pi_1 \, p(x_1 \mid \theta_1^{(t-1)}) + \pi_2 \, p(x_1 \mid \theta_2^{(t-1)})}$$

$$r_{11} = \frac{0.4333 \times 0.09056}{0.4333 \times 0.09056 + 8.3825 \times 10^{-4} \times 0.5667}$$

$\boxed{r_{1,1} = 0.9880}$ $\therefore$ $\boxed{r_{1,2} = 0.012}$

$$r_{2,1} = \frac{0.4333 \times 0.01987}{0.4333 \times 0.01987 + 0.5667 \times 0.03968}$$

$\boxed{r_{2,1} = 0.2768}$ $\boxed{r_{2,2} = 0.7232}$

$$r_{3,1} = \frac{0.4333 \times 4.97 \times 10^{-6}}{0.4333 \times 4.97 \times 10^{-6} + 0.5667 \times 0.057124}$$

$\boxed{r_{3,1} = 6.65 \times 10^{-5}}$ $\boxed{r_{3,2} = 0.999}$

$$r = \begin{bmatrix} 0.9880 & 0.012 \\ 0.2768 & 0.7232 \\ 6.65 \times 10^{-5} & 0.999 \end{bmatrix}$$

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

Q2 PCA via Successive Deflation

1. Show that the covariance of the deflated matrix,

$$\tilde{C} = \frac{1}{n} \tilde{X} \tilde{X}^T \text{ is given by}$$

$$\tilde{C} = \frac{1}{n} XX^T - \lambda_1 v_1 v_1^T$$

$\rightarrow$

$$\tilde{C} = \frac{1}{n} \tilde{X} \tilde{X}^T \quad \text{— given} \quad \text{— (1)}$$

$$\tilde{X} = (I - v_1 v_1^T) X \quad \text{— given} \quad \text{— (2)}$$

putting eq^n (2) in (1) we get,

$$\tilde{C} = \frac{1}{n} (I - v_1 v_1^T) X ((I - v_1 v_1^T) X)^T$$

$$= \frac{1}{n} (I - v_1 v_1^T) X (X^T (I - v_1 v_1^T)^T)$$

$$= \frac{1}{n} (I - v_1 v_1^T) X \; X^T (I - v_1 v_1^T)$$

— $(I - v_1 v_1^T)$ is symmetric and using $(u.v)^T = v^T.u^T$ rule

$$= \frac{1}{n} (IX - v_1 v_1^T X)(X^T I - X^T v_1 v_1^T)$$

$$= \frac{1}{n} \left[(X - v_1 v_1^T X)(X^T - X^T v_1 v_1^T) \right] \quad \text{— } IX = X \text{ property.}$$

$$= \frac{1}{n} \left[XX^T - XX^T v_1 v_1^T - v_1 v_1^T XX^T - v_1 v_1^T XX^T v_1 v_1^T \right]$$

we know that, $XX^T v_1 = n \lambda_1 v_1$

hence $= \frac{1}{n} \left[XX^T - n\lambda_1 v_1 v_1^T - v_1 v_1^T XX^T - v_1 v_1^T XX^T v_1 v_1^T \right]$

$$\tilde{c} = \frac{1}{n}\left[XX^T - n\lambda_1 v_1 v_1^T\right] + \frac{1}{n}\left[-v_1 v_1^T X X^T + v_1 v_1^T X X^T v_1 v_1^T\right] \quad \text{—③}$$

we know that,

$$XX^T v_1 = n\lambda_1 v_1$$

taking transpose on both sides,

$$(XX^T v_1)^T = (n\lambda_1 v_1)^T$$

$$v_1^T X X^T = n\lambda_1 v_1^T \quad \text{— } n, \lambda_1 \text{ constant.}$$

—④

using eq^n ④ in eq^n ③

$$\tilde{c} = \frac{1}{n}\left[XX^T - n\lambda_1 v_1 v_1^T\right] + \frac{1}{n}\left[-v_1 n\lambda_1 v_1^T + v_1 n\lambda_1 v_1^T v_1 v_1^T\right]$$

this can be rearranged as follows, as n & λ_1 are constant

$$\tilde{c} = \frac{1}{n}\left[XX^T - n\lambda_1 v_1 v_1^T\right] + \frac{1}{n}\left[-n\lambda_1 v_1 v_1^T + n\lambda_1 v_1 \underline{v_1^T v_1} v_1^T\right]$$

we are given, $v_1^T v_1 = 1$

$$\therefore \tilde{c} = \frac{1}{n}\left[XX^T - n\lambda_1 v_1 v_1^T\right] + \frac{1}{n}\left[-n\lambda_1 v_1 v_1^T + n\lambda_1 v_1 v_1^T\right]$$

$$\tilde{c} = \frac{1}{n}\left[XX^T - n\lambda_1 v_1 v_1^T\right] + \frac{1}{n}[0]$$

$$\tilde{c} = \frac{1}{n}\left[XX^T - n\lambda_1 v_1 v_1^T\right]$$

$$\therefore \boxed{\tilde{c} = \frac{1}{n} XX^T - \lambda_1 v_1 v_1^T}$$

Hence proved.

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

$\boxed{2.2}$ ST for $j \neq 1$, if v_j is principal eigenvector of C with corresponding eigenvalue λ_j

→ from Q2.1, we know that,

$$\tilde{C} = \frac{1}{n} XX^T - \lambda_1 v_1 v_1^T$$

$$\tilde{C} = C - \lambda_1 v_1 v_1^T$$

multiplying both sides by v_j we get

$$\tilde{C} \cdot v_j = C \cdot v_j - \lambda_1 v_1 v_1^T v_j \quad \forall j \text{ — } (1)$$

we know that $C \cdot v_j = \lambda_j v_j$

$$\therefore \tilde{C} \cdot v_j = \lambda_j v_j - \lambda_1 v_1 v_1^T v_j \quad \forall j$$

$$\tilde{C} \cdot v_j = \begin{cases} \lambda_1 v_1 - \lambda_1 v_1 v_1^T v_1 & j = 1 \\ \lambda_j v_j - \lambda_1 v_1 v_1^T v_j & j \neq 1 \end{cases}$$

$\boxed{\text{case 1}}$ if $j = 1$

$$\tilde{C} \cdot v_j = \lambda_1 v_1 - \lambda_1 v_1 v_1^T v_1$$

$$= \lambda_1 v_1 - \lambda_1 v_1 \quad \text{as } v_1^T v_1 = 1 \text{ - given}$$

$\boxed{\tilde{C} \cdot v_j = 0}$ no eigenvector info can be inferred.

$\boxed{\text{Case 2}}$ $j \neq 1$

$$\tilde{C} \cdot v_j = \lambda_j v_j - \lambda_1 v_1 v_1^T v_j \quad j \neq 1$$

we know that, $v_i^T v_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

$$\therefore \boxed{v_1^T v_j = 0 \text{ if } j \neq 1}$$

PTO

$\therefore \tilde{C} \cdot v_j = \lambda_j \cdot v_j - \lambda_1 v_1 \times 0$

$$\boxed{\tilde{C} \cdot v_j = \lambda_j v_j}$$

$\therefore v_j$ is also a principal eigenvector of $\tilde{C}$.

$\boxed{2.3}$ ~~Let u be first~~

from Q 2.2 we know that,

$\tilde{C} \cdot v_j = \lambda_j v_j - \lambda_1 v_1 v_1^T v_j$ — ①

if $j = 1$

$\tilde{C} \cdot v_1 = \lambda_1 v_1 - \lambda_1 v_1 v_1^T v_1$

$\tilde{C} \cdot v_1 = \lambda_1 v_1 - \lambda_1 v_1$ — as $v_1^T v_1 = 1$

$\therefore \boxed{\tilde{C} \cdot v_1 = 0}$ hence, v_1 is not principal eigen vector.

~~we~~ $v_1, v_2, v_3, \ldots, v_n$ are principal eigenvector in decreasing order.

~~but v_1 does not exist~~

but v_1 is not ~~a principal~~ eigenvector. ~~Hence~~.

let's verify if v_2 is eigenvector or not.

putting $j=2$ in eqn ①

$\tilde{C} \cdot v_2 = \lambda_2 v_2 - \lambda_1 v_1 v_1^T v_2$

we know that, $v_1^T v_2 = 0$ — given.

$\therefore \tilde{C} \cdot v_2 = \lambda_2 v_2 - 0 = \lambda_2 v_2$

$\therefore v_2$ is the eignen vector.

Hence $u = v_2$ is the first eigenvector of $\tilde{C}$.

ML ASSIGNMENT 4: BHUSHAN SONAWANE (111511679)

$\boxed{2.4}$

function ReturnLeadingVectors (C, k):

Input: C : (positive definite matrix)

k :: number of leading principal basis vectors.

steps:

1. vectors[K] Initialize to empty/Null
2. for i from 1 to K
3. $\lambda_i, v_i = f(C)$
4. $C = C - \lambda_i v v^T$
5. vectors[i] = v
6. end of for loop
7. Return vectors
8. end of function