Kumar Ayush
16MT30009

Assignment 2

Problem 1

$x_1 = (-1, 1)$
$x_2 = (5, -6)$
$x_3 = (0, -1)$
$x_4 = (-4, 4)$
$x_5 = (1, 0)$
$x_6 = (0, -2)$
$x_7 = (6, 4)$
$x_8 = (-6, 5)$
$x_9 = (-2, -1)$
$x_{10} = (-5, -7)$

Given K = 2

a) Taking x_1 and x_5 as initial means $\implies$ $\mu_1 = (-1, 1),\ \mu_2 = (1, 0)$

Using

$$\boxed{C_k = \{n : k = arg\ min\ ||x_n - \mu_k||^2\}}$$

* Stopping criteria: number of iterations = 3

Iteration 1:

$C_1 = arg\ min\ (0^2 + 0^2 = 0,\ 2^2 + 1^2 = 5) = 1$
$C_2 = arg\ min\ (6^2 + 7^2 = 85,\ 4^2 + 6^2 = 52) = 2$
$C_3 = arg\ min\ (1^2 + 2^2 = 5,\ 1^2 + 1^2 = 2) = 2$
$C_4 = arg\ min\ (3^2 + 3^2 = 18,\ 5^2 + 4^2 = 41) = 1$
$C_5 = arg\ min\ (2^1 + 1^2 = 5,\ 0^2 + 0^2 = 0) = 2$
$C_6 = arg\ min\ (1^2 + 3^2 = 10,\ 1^2 + 2^2 = 5) = 2$
$C_7 = arg\ min\ (7^2 + 3^2 = 58,\ 5^2 + 4^2 = 41) = 2$
$C_8 = arg\ min\ (5^2 + 4^2 = 41,\ 7^2 + 5^2 = 74) = 1$
$C_9 = arg\ min\ (1^2 + 2^2 = 5,\ 3^2 + 1^2 = 10) = 1$
$C_{10} = arg\ min\ (4^2 + 8^2 = 80,\ 6^2 + 7^2 = 85) = 1$
$\mu_1 = \left(\dfrac{-1-4-6-2-5}{5}, \dfrac{1+4+5-1-7}{5}\right) = \left(\dfrac{-18}{5}, \dfrac{2}{5}\right)$
$\mu_2 = \left(\dfrac{5+0+1+0+6}{5}, \dfrac{-6-1+0-2+4}{5}\right) = \left(\dfrac{12}{5}, -1\right)$

Iteration 2:

$C_1 = arg\ min\ (7.12,\ 15, 56) = 1$
$C_2 = arg\ min\ (114.92,\ 31.76) = 2$
$C_3 = arg\ min\ (14.92,\ 5, 76) = 2$
$C_4 = arg\ min\ (13.12,\ 65.96) = 1$
$C_5 = arg\ min\ (21.32,\ 2.96) = 2$

$C_6 = arg\ min\ (18.72,\ 6,76) = 2$
$C_7 = arg\ min\ (105.12,\ 37.96) = 2$
$C_8 = arg\ min\ (26.92,\ 106.56) = 1$
$C_9 = arg\ min\ (4.52,\ 19.36) = 1$
$C_{10} = arg\ min\ (56.72,\ 90.76) = 1$

$$\mu_1 = \Big(\frac{-1-4-6-2-5}{5},\ \frac{1+4+5-1-7}{5}\Big) = \Big(\frac{-18}{5}, \frac{2}{5}\Big)$$
$$\mu_2 = \Big(\frac{5+0+1+0+6}{5},\ \frac{-6-1+0-2+4}{5}\Big) = \Big(\frac{12}{5}, -1\Big)$$

Iteration 3:
$C_1 = arg\ min\ (7.12,\ 15.56) = 1$
$C_2 = arg\ min\ (114.92,\ 31.76) = 2$
$C_3 = arg\ min\ (14.92,\ 5.76) = 2$
$C_4 = arg\ min\ (13.12,\ 65.96) = 1$
$C_5 = arg\ min\ (21.32,\ 2.96) = 2$
$C_6 = arg\ min\ (18.72,\ 6.76) = 2$
$C_7 = arg\ min\ (105.12,\ 37.96) = 2$
$C_8 = arg\ min\ (26.92,\ 106.56) = 1$
$C_9 = arg\ min\ (4.52,\ 19.36) = 1$
$C_{10} = arg\ min\ (56.72,\ 90.76) = 1$

$$\mu_1 = \Big(\frac{-1-4-6-2-5}{5},\ \frac{1+4+5-1-7}{5}\Big) = \Big(\frac{-18}{5}, \frac{2}{5}\Big)$$
$$\mu_2 = \Big(\frac{5+0+1+0+6}{5},\ \frac{-6-1+0-2+4}{5}\Big) = \Big(\frac{12}{5}, -1\Big)$$

Clusters = [(-1, 1), (-4, 4), (-6, 5), (-2, -1), (-5, -7)] and [(5, -6), (0, -1), (1, 0), (0, -2), (6, 4)]

b) Kernalized K-means

$$\boxed{||\phi(x_n) - \phi(\mu_k)||^2 = K(x_n,\ x_n) + K(\mu_k,\ \mu_k) - 2\cdot k(X_n,\ \mu_k)}$$

For Gaussian Kernel,

$$\boxed{\phi(x) = K(x, x') = exp\Big(\frac{-||x-x'||^2}{\sigma^2}\Big)}$$

Taking $\sigma = 1$,
$\implies\ ||\phi(x_n) - \phi(\mu_k)||^2 = e^0 + e^0 - 2\cdot K(x_n, \mu_k) = 2 - 2\cdot K(x_n, \mu_k)$
Taking same initial means as in part a, i.e., $x_1,\ x_5$
$\mu_1 = (-1,\ 1),\ \mu_2 = (1,\ 0)$

Using

$$\boxed{C_k = \{n : k = arg\ min\ ||\phi(x_n) - \phi(\mu_k)||^2\}}$$

Iteration 1:

$C_1 = arg\ min\ (0,\ 1.98652) = 1$
$C_2 = arg\ min\ (2,\ 2) = 1$ Random tie break
$C_3 = arg\ min\ (1.9865,\ 1.729) = 2$
$C_4 = arg\ min\ (1.9999, 2.0) = 1$
$C_5 = arg\ min\ (1.98652,\ 2.0) = 1$
$C_6 = arg\ min\ (1.9990, 1.98652) = 2$
$C_7 = arg\ min\ (2.0, 2.0) = 2$
$C_8 = arg\ min\ (2.0, 2.0) = 1$
$C_9 = arg\ min\ (1.98652,\ 1.9999) = 1$
$C_{10} = arg\ min\ (2.0, 2.0) = 2$

$$\mu_1 = \Big(\frac{1+5-4+1-6-2}{6},\ \frac{1-6+4+0+5-1}{6}\Big) = \Big(\frac{-7}{6}, \frac{3}{6}\Big)$$
$$\mu_2 = \Big(\frac{0+0+6-5}{4},\ \frac{-1-2+4-7}{4}\Big) = \Big(\frac{1}{4}, \frac{-6}{4}\Big)$$

Iteration 2:

$C_1 = arg\ min\ (0.4850,\ 1.991) = 1$
$C_2 = arg\ min\ (2,\ 2) = 1$
$C_3 = arg\ min\ (1.9459,\ 0.5367) = 2$
$C_4 = arg\ min\ (1.9999, 2.0) = 1$
$C_5 = arg\ min\ (1.9857,\ 1.87989) = 2$
$C_6 = arg\ min\ (1.9990, 0.5367) = 2$
$C_7 = arg\ min\ (2.0, 2.0) = 2$
$C_8 = arg\ min\ (2.0, 2.0) = 1$
$C_9 = arg\ min\ (1.894,\ 1.99014) = 1$
$C_{10} = arg\ min\ (2.0, 2.0) = 1$

$$\mu_1 = \Big(\frac{-1+5-4-6-2-5}{6},\ \frac{1-6+4+5-1-7}{6}\Big) = \Big(\frac{-13}{6}, \frac{-4}{6}\Big)$$
$$\mu_2 = \Big(\frac{0+1+0+6}{4},\ \frac{-1+0-2+4}{4}\Big) = \Big(\frac{7}{4}, \frac{1}{4}\Big)$$

Iteration 3:

$C_1 = arg\ min\ (1.99999,\ 1.9994) = 2$
$C_2 = arg\ min\ (2,\ 2) = 1$
$C_3 = arg\ min\ (1.9999,\ 1.98) = 2$
$C_4 = arg\ min\ (2.0, 2.0) = 2$
$C_5 = arg\ min\ (1.9999,\ 0.929) = 2$

$C_6 = arg\ min\ (1.9996, 1.9994) = 2$
$C_7 = arg\ min\ (2.0, 1.999) = 2$
$C_8 = arg\ min\ (2.0, 2.0) = 1$
$C_9 = arg\ min\ (1.99975,\ 1.9999) = 1$
$C_{10} = arg\ min\ (1.9999, 2.0) = 1$

$$\mu_1 = \Big(\frac{5-6-2-5}{6},\ \frac{-6+5-1-7}{6}\Big) = \Big(-2, \frac{-9}{4}\Big)$$
$$\mu_2 = \Big(\frac{-1+0-4+1+0+6}{6},\ \frac{1-1+4+0-2+4}{4}\Big) = \Big(\frac{2}{6}, 1\Big)$$

Clusters: [(5, -6), (-6, 5), (-2, -1), (-5, -7)] and [(-1, 1), (0, -1), (-4, 4), (1, 0), (0, -2), (6, 4)]

Problem 2

a. Geometric

parameter: θ , $\theta(1-\theta)^{x-1}$

PMF = $\theta(1-\theta)^{x-1}$

Parameter → p

p : success probability

$$PMF = p(1-p)^{x-1} = f_X(x,\theta)$$

$$E(X) = \frac{1}{p}$$

$$V(X) = \frac{1-p}{p^2}$$

$$L_N(p; x_1, \dots x_N) = \prod_{i=1}^{N} f_X(x_i; p) = p^N (1-p)^{\sum_{i=1}^{N}(x_i-1)}$$

$$l_N(p;x) = \log(L_N(p; x_1, \dots x_N)) = N \ln(p) + (\ln(1-p)) \sum_{i=1}^{N} x_i - 1$$

$$\frac{\partial l_N(p;x)}{\partial p} = \frac{N}{p} - \frac{\sum_{i=1}^{N}(x_i-1)}{1-p} = 0$$

$$\frac{\partial^2 l_N(p;x)}{\partial p^2} = -\frac{N}{p^2} - \left(\frac{1}{1-\theta}\right)^2 \sum_{i=1}^{N}(x_i-1)$$

$$(1-p)\frac{N}{p} = \frac{\left(\sum_{i=1}^{N} x_i\right) - N}{N}$$

$$\Rightarrow \left(\frac{1}{p} - \cancel{1}\right) = \frac{\sum_{i=1}^{N} x_i}{N} - \cancel{1}$$

$$\Rightarrow \boxed{p = \frac{1}{\bar{X}}}$$

$\bar{X}$ → mean

Page No.

$$-\left[\frac{\cancel{\sum_{i=1}^{m} x_i}}{\cancel{\sum x_i}}\,\frac{m^2n^2}{\sum_{i=1}^{m} x_i} - \left(\frac{mn - \sum x_i m}{(mn - \sum x_i)^2}\right.\right.$$

$$-\left[mn\left[\frac{mn}{\sum x_i} - \frac{1}{mn - \sum x_i}\right)\right]$$

$$-mn\left[\frac{m^2n^2 - mn\sum x_i - \sum x_i}{(\sum x_i)(mn - \sum x_i)}\right] < 0$$

Binomial distribution

$$\text{PMF} = \binom{N}{x} p^x (1-p)^{N-x}$$

Parameter : (N, p)

$$L(x_1, x_2, \ldots, x_m, p) = \prod_{i=1}^{m} \binom{n}{x_i} p^{x_i} (1-p)^{n-x_i}$$

$$l(x_1, x_2, \ldots x_m, p) = \ln L(x_1, x_2 \ldots x_m, p)$$

$$= \sum_{i=1}^{m} \left(\ln \binom{n}{x_i} + x_i \ln p + (n - x_i) \ln(1-p) \right)$$

$$= \sum_{i=1}^{m} \ln \binom{n}{x_i} + \left(\sum_{i=1}^{m} x_i \right) \ln p + \left(mn - \sum_{i=1}^{m} x_i \right) \ln(1-p)$$

$$\frac{\partial l(x_1, x_2, \ldots x_m, p)}{\partial p} = 0 + \left(\frac{\sum_{i=1}^{m} x_i}{p} \right) - \left(\frac{mn - \sum_{i=1}^{m} x_i}{1-p} \right) = 0$$

$$\Rightarrow \frac{\sum x_i (1-p)}{p} = mn - \sum_{i=1}^{m} x_i$$

$$\Rightarrow \frac{1}{p} \sum_{i=1}^{m} x_i - \cancel{\sum_{i=1}^{m} x_i} = mn - \cancel{\sum x_i}$$

$$\Rightarrow p = \frac{\sum_{i=1}^{m} x_i}{mn} =$$

$$\frac{\partial^2 l(x_1, \ldots x_m, p)}{\partial p^2} = \frac{-\sum_{i=1}^{n} x_i}{p^2} - \left(\frac{mn - \sum_{i=1}^{n} x_i}{(1-p)^2} \right)$$

Page No.

Date ___ / ___ / ______

$$\Rightarrow \sum_i \frac{x_i}{\lambda} = 1$$

$$\Rightarrow \lambda = \sum_i x_i = N$$

$$\Rightarrow \theta_i = \frac{x_i}{N}$$

Multinomial distribution

$$PMF = \frac{N!}{x_1! x_2! \cdots x_k!} \prod_{k=1}^{K} p_k^{x_k}$$

$$l(\theta, x_i, \ldots x_k) = \log N! - \sum_{i=1}^{N} \log(x_i!) + \sum_i x_i \log \theta_i$$

Using Lagrange Multiplier

$$l(\theta, x_i, \ldots x_k) = \log N! - \log(x_i!) + \sum x_i \log \theta + \lambda \left(1 - \sum p_i\right)$$

$$\frac{dl}{d\theta_i} = \frac{x_i}{\theta_i} - \lambda$$

$$\frac{dl}{d\lambda} = 1 - \sum_i \theta_i$$

We have system of eqn's

$$\theta_1 = \frac{x_1}{\lambda}$$

$$\vdots$$

$$\theta_k = \frac{x_k}{\lambda}$$

$$\sum_i \theta_i = 1$$

Page No.

Date ___ / ___ / ___

$$\Rightarrow \sum_i \frac{x_i}{\lambda} = 1$$

$$\Rightarrow \lambda = \sum_i x_i = N$$

$$\Rightarrow \theta_i = \frac{x_i}{N}$$

Problem 3

3. MAP estimation for parameters of:

a) Poisson (Gamma as prior)

$$G(\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \cdot x^{\alpha-1} e^{-x/\beta} \quad ; \text{ for } x, \alpha, \beta > 0$$

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$$

$\lambda \sim \text{Gamma}(\alpha,\beta)$; $X \sim \text{Exp}(\lambda)$

n: number of variables

$$P(\lambda|x) \propto P(x|\lambda) \cdot P(\lambda)$$

$$\propto \lambda^n e^{-\lambda \Sigma x_i} \cdot \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$\propto e^{-\lambda(\sum_i x_i - \beta)} \cdot \lambda^{n+\alpha-1}$$

Taking log,

$$\log(P(\lambda|x)) \propto -\lambda\left(\sum_i x_i + \beta\right) + (n+\alpha-1)\log$$

differentiating & setting derivative $= 0$

$$\Rightarrow -\sum_i x_i - \beta + \frac{(n+\alpha-1)}{\lambda} = 0$$

$$\Rightarrow \lambda = \frac{n+\alpha-1}{\sum_i x_i + \beta}$$

b) Dirichlet distribution as prior for categorical distribution

$$f(x_1, \dots x_K; \alpha_1, \dots \alpha_K) = \frac{1}{B(\alpha)} \cdot \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

$$\therefore \; B(\alpha) = \frac{\prod_{i=1}^{n} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)}$$

for $P(\theta | x) = \dfrac{\mathcal{L}(\theta)\, \pi(\theta)}{P(x)}$

$$\Rightarrow \; \hat{\theta}_{max} = \arg\max \left(\frac{\Gamma\left(\sum_{i=1}^{K} x_i + 1\right)}{\prod_i \Gamma(x_i + 1)} \cdot \prod_{i=1}^{K} \theta_i^{x_i} \cdot \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)} \cdot \prod_{i=1}^{K} \theta_i^{\alpha_i - 1} \right)$$

terms that does not depend on θ will not affect result, hence dropping those terms.

$$\hat{\theta}_{max} = \arg\max \left(\prod_{i=1}^{K} \theta_i^{x_i} \cdot \prod_{i=1}^{K} \theta_i^{\alpha_i - 1} \right)$$

$$= \arg\max \left(\prod_{i=1}^{K} \theta_i^{(\alpha_i - 1 + x_i)} \right)$$

taking log ~~on both sides~~

$$\log \hat{\theta}_{max} = \arg\max \left(\sum_{i=1}^{\alpha} \log \left(\theta^{(x_i + \alpha_i - 1)} \right) \right)$$

$$= \arg\max \left(\sum_i (x_i + \alpha_i - 1) \log(\theta_i) \right)$$

Above function is monotonically increasing & we have constraint $\sum \theta_i = 1$

Using Langrange's multiplier

$$\mathcal{L}(\theta, \lambda) = \sum_{i=1}^{K} (x_i + \alpha_i - 1) \log(\theta_i) - \lambda \left[\sum_{i=1}^{K} \theta_i \right.$$

taking derivative w.r.t. θ_i

$$\Rightarrow \frac{\partial \mathcal{L}(\theta, \lambda)}{\partial \theta_i} = \frac{(x_i + \alpha_i - 1)}{\theta_i} - \frac{(x_k + \alpha_k - 1)}{1 - \sum_{i=1}^{K-1} \theta_i}$$

Substituting θ_K for $1 - \sum_{i=1}^{K-1} \theta_i$

$$\Rightarrow \frac{\partial (L(\theta))}{\partial \theta_i} = \frac{x_i + \alpha_i - 1}{\theta_i} - \frac{x_K + \alpha_K - 1}{\theta_K}$$

$$\Rightarrow \hat{\theta}_i = \frac{x_i + \alpha_i - 1}{\sum_j (x_j + \alpha_j - 1)}$$

$$\text{or } \lambda = \frac{x_y + \alpha_y - 1}{\theta_y} \text{ for each } y, \sum_i \theta_i = 1$$

(c) Multivariate Gaussian (with multivariate normal as prior)

$$\mathcal{N}(x|\mu,\Sigma) = \frac{1}{2\pi^{D/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$

$\therefore\ x \sim \mathcal{N}(\mu, \Sigma),\ \mu \in R^d$

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

$$p\left(\begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \middle| \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix}\right) = (2\pi)^{-p/2} \cdot \begin{vmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{vmatrix}^{-1/2} \times e^{\Omega}$$

$$\Omega = -\frac{1}{2}\left[(x_1-\mu_1) \cdots (x_p-\mu_p)\right] \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix}^{-1} \begin{bmatrix} x_1-\mu_1 \\ \vdots \\ x_p-\mu_p \end{bmatrix}$$

Taking log

$$\log(p(D|\mu,\Sigma)) = -\frac{Nd}{2}\log(x_i) - \frac{N}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{N}(x_i-\mu)^T \cdot \Sigma^{-1}(x_i-\mu)$$

differentiating)

$$\frac{d\log(p(D|\mu,\Sigma))}{d\mu} = \sum_{i=1}^{N}(x_i-\mu)^T\Sigma^{-1}$$

Forming likelihood prior

$\prod_{i=1}^{N} p_i \cdot p(\mu)$: p_i form of bivariate Gaussian

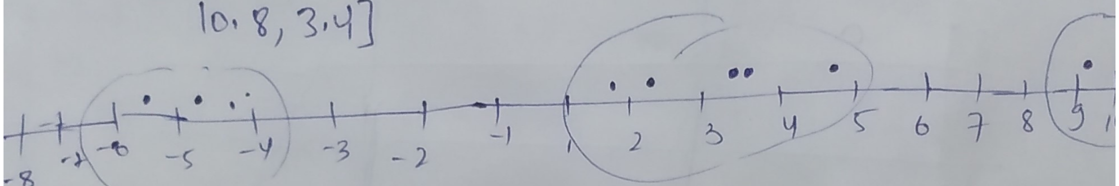
Taking log of expression & derivative w.r.t. μ & equating to 0.

$$\Rightarrow \mu_{MAP} = \left(\Sigma_0^{-1} + m\Sigma^{-1}\right)^{-1}\left(\left\{ \mu_0^T \Sigma_0^{-1} + \sum_{i=1}^{N} x^i \Sigma^{-1}\right)\right.$$

$$\mu_0 = [\mu_{10}, \mu_{20}]^T, \quad 4\ x_i = [x_1^{(i)}, x_2^{(i)}]^T$$

4. observations.

[2.3, 4.7, −5.5, −4.8, 9.1, 3.5, 10.4, −4.3, 11.2, 1.9, 10.8, 3.4]



looking at data & clusters, 3 clusters is needed.

$$\text{Choosing } (\mu^0, \sigma^0, \pi^0) = (-4.5, 0.2, 0.2)$$
$$(\mu^1, \sigma^1, \pi^1) = (3.2, 0.4, 0.4)$$
$$(\mu^2, \sigma^2, \pi^2) = (9.8, 0.6, 0.4)$$

$$\mu_j^K = \frac{\sum_{i=1}^{N} \gamma_{ij} x_i}{\sum_{i=1}^{N} \gamma_{ij}}, \quad \pi_j = \frac{\sum \gamma_{ij}}{1}, \quad \sigma_j = \frac{\sum_{i=1}^{N} \gamma_{ij}}{\sum_{i=1}^{N} \gamma_{ij}}$$

$\cancel{\sigma_j^K = \sum_{i=1}^{N} \gamma_{ij}}$

Problem 4

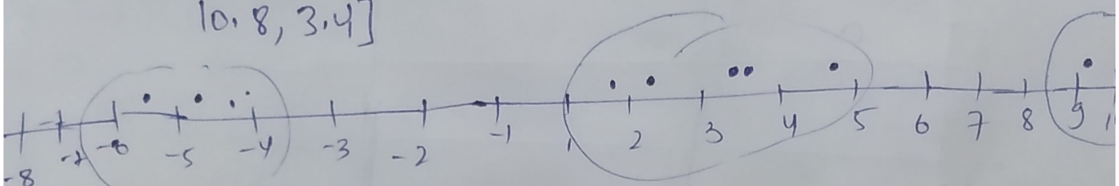
Taking log of expression & derivative w.r.t. μ & equating to 0.

$$\Rightarrow \hat{\mu}_{MAP} = \left(\Sigma_0^{-1} + m\Sigma^{-1}\right)^{-1}\left(\left(\mu_0^T \Sigma_0^{-1} + \sum_{i=1}^{N} x^i \Sigma^{-1}\right)\right.$$

$$\mu_0 = [\mu_{10}, \mu_{20}]^T, \quad \& \quad x_i = [x_1^{(i)}, x_2^{(i)}]^T$$

4. observations.

[2.3, 4.7, -5.5, -4.8, 9.1, 3.5, 10.4, -4.3, 11.2, 1.9, 10.8, 3.4]



looking at data & clusters, 3 clusters is needed.

$$\text{Choosing } (\mu^0, \sigma^0, \pi^0) = (-4.5, 0.2, 0.2)$$
$$(\mu^1, \sigma^1, \pi^1) = (3.2, 0.4, 0.4)$$
$$(\mu^2, \sigma^2, \pi^2) = (9.8, 0.6, 0.4)$$

$$\mu_j^K = \frac{\sum_{i=1}^{N} \gamma_{ij} x_i}{\sum_{i=1}^{N} \gamma_{ij}}, \quad \pi_j^K = \frac{\sum \gamma_{ij}}{1}, \quad \sigma_j^1 = \frac{\sum \gamma_{ij}}{\sum_{i=1}^{N} \gamma_{ij}}$$

$$\cancel{\sigma_j^K = \sum_{i=1}^{N} \gamma_{ij}}$$

γ_{ij} table

Clusters

Datapoints	1	2	3
1	2.3885×10^{-250}	1	6.57×10[illegible]
2	0	1	1.0293×10[illegible]
3	1	2.531×10^{-98}	$3.7625 \times$[illegible]
4	1	2.131×10^{-87}	1.82×10^{-1}[illegible]
5	0	2.5389×10^{-47}	1
6	0	1	$6.7527 \times$[illegible]
7	0	1.635×10^{-70}	1
8	1	3.76×10^{-77}	4.4093×10[illegible]
9	0	4.737×10^{-86}	1
10	1.721×10^{-220}	1	1.979×10[illegible]
11	0	3.67×10^{-78}	1
12	0	1	5.89×10[illegible]

μ	σ	π
-4.667	0.37667	0.2500
3.16	0.97	0.41667
10.375	0.95250	0.3333

2rd Iteration

	μ	σ	π
1	-4.8867	0.28209	0.25
2	3.1600	0.97440	0.41667
3	10.375	0.62188	0.33333

Parameters θ

$\because (\mu^1, \sigma^1, \pi^1) \approx (\mu^2, \sigma^2, \pi^2)$

$\Rightarrow$ 2 iteration is enough