AML-
Test-2

DEEPAK MEWADA
20CS91P02

Soln 1 ~~Answer 1~~

$P(A) = 1/2$

$P(B) = \mu$

$P(C) = 2\mu$

$P(D) = \frac{1}{2} - 3\mu$

Grade	A	B	C	D
No of student	a	b	c	d

$\underbrace{a + b}_{} = h$

E-Step Evaluates the expected value of a and b

$$a' = \left(\frac{P(A)}{P(A)+P(B)}\right)h$$ by Bayes rule

$$a' = \boxed{\left(\frac{\frac{1}{2}}{\frac{1}{2}+\mu}\right)h}$$

$$b' = \left(\frac{P(B)}{P(A)+P(B)}\right)h = \boxed{\left(\frac{\mu}{\frac{1}{2}+\mu}\right)h}$$

M-Step Re-estimate the parameter using the current responsibilities

Assumption - Computing the MLE of μ, by assuming that unobserved variables are replaced by their expectations.

$$\mu' = \frac{b+c}{(b+c+d)}\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)$$

$$\mu' = \frac{b+c}{6(b+c+d)} = \boxed{\frac{h-a+c}{6(h-a+c+d)}}$$

P.T.O.

Deepak Memade

Conclusion

By iterating between the step E-step & M-step will always converges to a local optimum of μ (which may or may not be a global optimum value)

20CS91P02
Deepak Nemada

(2)

$P(X)$ — fixed distribution

$$q(x) = \mathcal{N}(x \mid \mu, \Sigma)$$

To show $\mu = \mathbb{E}_{p(x)}(X)$, $\Sigma = \text{cov}(X)$

$$D_{KL}(P \| q) = \mathbb{E}_{p(x)} \log \frac{P(x)}{q(x)}$$

$$= \int_x P(x) \log \frac{P(x)}{q(x)}$$

$$= \int P(x) \log \frac{P(x)}{\mathcal{N}(x \mid \mu, \Sigma)}$$

$$D_{KL}(P \| q) = \underbrace{\int_x P(x) \log P(x)}_{\text{constant}} - \underbrace{\int_x P(x) \log \mathcal{N}(x \mid \mu, \Sigma)}_{F(x)} \quad \text{--- eqn (1)}$$

$\int P(x) \log P(x)$ is constant hence to minimize D_{KL} we need to maximize $\int_x P(x) \log \mathcal{N}(x \mid \mu, \Sigma)$.

Let $F = \int_x P(x) \log \mathcal{N}(x \mid \mu, \Sigma)$

$$F = \int_x P(x) \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu) \right] dx$$

diff F wrt. μ using $\frac{\partial (x^T A x)}{\partial x} = x^T (A + A^T)$

$$\frac{dF}{d\mu} = \int_x P(x) \left[0 - 0 - \frac{1}{2} \left[(x-\mu)(\Sigma_k^{-1} + \Sigma_k) \right] \right]$$

For maximization equating it to zero

$$\int_x P(x) \left((x - \mu) \right) = 0$$

$$\int_x P(x) . x - \mu \int_x P(x) = 0$$

$$\mu = \int_x P(x) . x \quad \text{since } \int_x P(x) = 1$$

hence proved

$$\Rightarrow \boxed{\mu = \mathbb{E}_{p(x)}(X)}$$

$\mu = \mathbb{E}_p(X)$

Since $f(\mu)$ is maximum when it is given by the expectation of x under $p(x)$

from eq^n (1)

$$D_{KL}(p\|q) = \int_x p(x)\log p(x) - \int_x f(\mu)$$

Hence $D_{KL}(p\|q)$ will be minimised when $\mu = \mathbb{E}_p(x)$ i.e μ is given by the expectation of x under $p(x)$.

Deepak Menander
~~DESS~~
20CS91P02

Answer 2 Part 2

To show: minimization of KL divergence with respect to Σ leads to the result that Σ is given by its covariance.

from eqn 1

i.e

~~$D_{KL} = \int p(x)$~~

$$D_{KL}(P\|q) = \underbrace{\int p(x)\log p(x)}_{const} - \underbrace{\int p(x)\log N(x|\mu,\Sigma)\,dx}_{F}$$

Now ~~differ~~ differentiating F. w.r.t "Σ" & equate to 0

$$\frac{dF}{d\Sigma} = \frac{d}{d\Sigma}\left[\int p(x)\log N(x|\mu,\Sigma)\,dx\right] = 0$$

$$\frac{dF}{d\Sigma} = \frac{d}{d\Sigma}\left[\int p(x)\left(\frac{1}{2}\log 2\pi - \frac{1}{2}\log\Sigma - \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)\right] = 0$$

$$= \int p(x)\left(0 - \frac{1}{2\Sigma} + \frac{(x-\mu)^T(x-\mu)}{2\Sigma^2}\right) = 0$$

$$\int p(x)\left(-\Sigma + (x-\mu)^T(x-\mu)\right) = 0$$

$$\boxed{\Sigma = \frac{\int p(x)(x-\mu)^T(x-\mu)}{\int p(x)}}$$

$$\therefore \Sigma = E_p[(x-\mu)^T(x-\mu)]$$

hence proved that Σ is given by its covariance.