Solution for Q2

Kavya Gupta

September 2023

Understanding the Gaussian Mixture Model (GMM) :

So interesting thing to note about GMM is that it is **not** a mere linear combination of random variables (r.v.). In latter, all the component r.v. X_is return a value whose linear combination is taken.
But in GMM, as clearly stated in the question, actually only one of the X_is are chosen with probability p_i and from that value is taken. Hence at a particular measurement, **only and exactly** one r.v. comes into action.

Now, for GMM, we will make use of **Law of Total Expectation/Law of Iterated Expectation/Adam's Rule** which states :-

$$X \sim \sum_i p_i X_i \Rightarrow E[g(X)] = \sum_i E[g(X)|(X = X_i)]P(X = X_i) = \sum_i E[g(X_i)]P(X = X_i) = \sum_i p_i E[g(X_i)]$$

Where g is a function on a r.v.

Solving the GMM :-

Before moving ahead, let's see what mean and variance of a Gaussian Variable $\mathcal{N}(\mu_i, \sigma_i^2)$

$$E[X_i] = E[\mathcal{N}(\mu_i, \sigma_i^2)] = \mu_i \tag{1}$$

$$E[(X_i - \mu_i)^2] = \sigma_i^2 \tag{2}$$

As $Var(Y) = E[Y^2] - E[Y]^2$ for any r.v. Y,

$$E[X_i^2] = Var(X_i) + E[X_i]^2 = \sigma_i^2 + \mu_i^2 \tag{3}$$

Mean E(X)

Using the above Law and equation (1)...

$$\mu = E[X] = \sum_i p_i E[X_i] = \sum_{i=1}^{K} p_i \mu_i \tag{4}$$

Variance Var(X)

$$Var(X) = E[(X-\mu)^2]$$

From the above Law,

$$E[(X-\mu)^2] = \sum_i p_i E[(X_i-\mu)^2]$$

Opening by the linearity of the expectation operator...

$$E[(X_i-\mu)^2] = E[X_i^2 - 2\mu X_i + \mu^2] = E[X_i^2] - 2\mu E[X_i] + E[\mu^2]$$

From equations (1), (3) and (4)...

$$E[X_i^2] - 2\mu E[X_i] + E[\mu^2] = \sigma_i^2 + \mu_i^2 - 2\mu\mu_i + \mu^2$$

$$\Rightarrow \sum_i p_i E[(X_i-\mu)^2] = \sum_i p_i(\sigma_i^2+\mu_i^2-2\mu\mu_i+\mu^2) = \sum_i [p_i(\sigma_i^2+\mu_i^2)-2\mu(\sum_i p_i\mu_i)+\mu^2(\sum_i p_i)]$$

Given $\sum_i p_i = 1$, so...

$$\Rightarrow \sum_i p_i E[(X_i-\mu)^2] = \sum_i [p_i(\sigma_i^2+\mu_i^2)] - 2\mu(\mu) + \mu^2(1) = \sum_i [p_i(\sigma_i^2+\mu_i^2)] - \mu^2$$

Hence...

$$Var(X) = \sum_{i=1}^{K} [p_i(\sigma_i^2+\mu_i^2)] - \mu^2$$

MGF(X) $\Phi_X(t)$

For a Gaussian r.v. $X_i = \mathcal{N}(\mu_i, \sigma_i^2)$, its MGF is...

$$\Phi_{X_i}(t) = E[e^{tX_i}] = e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2}$$

So using the above Law...

$$\Phi_X(t) = E[e^{tX}] = \sum_i p_i E[e^{tX_i}] = \sum_i p_i \Phi_{X_i}(t) = \sum_{i=1}^{K} p_i e^{\mu_i t + \frac{1}{2}\sigma_i^2 t^2}$$

Solving the Linear Combination System :

Now we have a r.v. Z such that $Z = \sum_i p_i X_i$ and as said in the very beginning, all the component r.v. X_is return a value and their linear combination is taken. We shall make extensive use of the **linearity property** of the expectation operator...

Mean E(Z)

$$E[Z] = E[\sum_i p_i X_i] = \sum_i E[p_i X_i] = \sum_i p_i E[X_i] = \sum_{i=1}^{K} p_i \mu_i = \mu$$

Peculiarly it's same as the GMM case..

Variance Var(Z)

$$Var(Z) = E[(Z-\mu)^2] = E[(\sum_i p_i X_i - \mu)^2]$$

$$\Rightarrow E[(\sum_i p_i X_i)^2 - 2\mu Z + \mu^2] = E[(\sum_i p_i X_i)^2] - 2\mu E[Z] + \mu^2 = E[(\sum_i p_i X_i)^2] - 2\mu(\mu) + \mu^2 = E[(\sum_i p_i X_i)^2] - \mu^2$$

Now, using this equation : $(\sum_i p_i X_i)^2 = \sum_{1 \le i \le K} (p_i X_i)^2 + 2 \sum_{1 \le i < j \le K} p_i p_j X_i X_j$...

$$E[(\sum_i p_i X_i)^2] = E[\sum_{1 \le i \le K} (p_i X_i)^2 + 2 \sum_{1 \le i < j \le K} p_i p_j X_i X_j]$$

$$\Rightarrow \sum_{1 \le i \le K} p_i^2 E[X_i^2] + 2 \sum_{1 \le i < j \le K} p_i p_j E[X_i X_j] = \sum_{1 \le i \le K} p_i^2 E[X_i^2] + 2 \sum_{1 \le i < j \le K} p_i p_j E[X_i] E[X_j]$$

$E[X_i X_j] = E[X_i]E[X_j]$ is possible because for $i \neq j, X_i, X_j$ are **independent random variables**.
Using equation (1) and (3)...

$$\Rightarrow E[(\sum_i p_i X_i)^2] = \sum_{1 \le i \le K} p_i^2 (\sigma_i^2 + \mu_i^2) + 2 \sum_{1 \le i < j \le K} p_i p_j \mu_i \mu_j = \sum_{1 \le i \le K} [p_i^2 \sigma_i^2] + (\sum_{1 \le i \le K} p_i^2 \mu_i^2 s + 2 \sum_{1 \le i < j \le K} p_i p_j \mu_i \mu_j)$$

$$\Rightarrow E[(\sum_i p_i X_i)^2] = \sum_{1 \le i \le K} [p_i^2 \sigma_i^2] + (\sum_i p_i \mu_i)^2 = \sum_{1 \le i \le K} [p_i^2 \sigma_i^2] + \mu^2$$

Finally...

$$Var(Z) = E[(\sum_i p_i X_i)^2] - \mu^2 = \sum_{1 \le i \le K} [p_i^2 \sigma_i^2] + \mu^2 - \mu^2 = \sum_{1 \le i \le K} p_i^2 \sigma_i^2$$

MGF(Z) $\Phi_Z(t)$:

Here we will make use of the fact that for r.v. $A = \sum_i A_i$ where A_is are independent r.v., then...

$$\Phi_A(t) = E[e^{tA}] = E[e^{t \sum_i A_i}] = E[\prod_i e^{tA_i}] = \prod_i E[e^{tA_i}] = \prod_i \Phi_{A_i}(t)$$

As A_is are independent r.v.s, then e^{tA_i}s are **also independent**... that's why they're expectation can be split into individual expectations.

Also, as A_is are independent, p_iA_is are also independent... So for $A = Z$ take $A_i = p_iX_i$, hence...

$$\Phi_Z(t) = \prod_i \Phi_{p_iX_i}(t)$$

Also, MGF(p_iX_i) is :

$$\Phi_{p_iX_i}(t) = E[e^{(p_it)X_i}] = e^{\mu_ip_it+\frac{1}{2}\sigma_i^2(p_it)^2}$$

$$\Rightarrow \Phi_Z(t) = \prod_i e^{\mu_ip_it+\frac{1}{2}\sigma_i^2(p_it)^2} = e^{\sum_i \mu_ip_it+\frac{1}{2}\sigma_i^2(p_it)^2} = e^{\sum_i(\mu_ip_i)t+\frac{1}{2}\sum_i(p_i^2\sigma_i^2)t^2}$$

And from results of mean and variance we calculated above...

$$\Rightarrow \Phi_Z(t) = e^{\mu t+\frac{1}{2}\sigma^2t^2}$$

This form of MGF is exactly same as of a Gaussian Variable with mean = μ and variance = $\sigma^2 = \sum_i p_i^2\sigma_i^2$, in other words, Z is also a Gaussian Variable : $Z = \mathcal{N}(\mu, \sigma^2)$.

PDF(Z):

Now that we know that Z is also a Gaussian Random Variable, we can calculate PDF for it easily.. it comes from the definition of the Gaussian itself...

$$PDF(Z) = f_Z(z) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Where $\mu = \sum_i p_i\mu_i$ and $\sigma^2 = \sum_i p_i^2\sigma_i^2$ as we had calculated above.