12/12

[redacted] – PART 2

1) Contribution to the likelihood function of the $i-th$ experiment result.

2

If $|y_i| > 1$ then y_i represents x_i and its contribution to the likelihood function is

$$\frac{1}{\sigma} \cdot \phi\left(\frac{y_i - \mu}{\sigma}\right)$$

If $y_i = 0$ then the only information that gives y_i is that $|x_i| \geq 1$, thus

$$\Phi\left(\frac{1-\mu}{\sigma}\right) - \Phi\left(\frac{-1-\mu}{\sigma}\right)$$

2) Log-likelihood function for the complete data

1

The complete data means that the small quantities of x have also been observed and then only x_i have to be taken into account

$$l = \prod_{i=1,..,n} \frac{1}{\sigma} \cdot \phi\left(\frac{x_i - \mu}{\sigma}\right)$$

$$\mathcal{L} = \log(l) = \sum_{i=1,\dots,n} \log\left(\frac{1}{\sigma} \cdot \phi\left(\frac{x_i - \mu}{\sigma}\right)\right)$$

3) Log-likelihood function for the observed data

1

$|y_i| > 1$

$$l = \prod_{i|y_i>1} \frac{1}{\sigma} \cdot \phi\left(\frac{x_i - \mu}{\sigma}\right) \cdot \prod_{i|y_i=0} \left(\Phi\left(\frac{1-\mu}{\sigma}\right) - \Phi\left(\frac{-1-\mu}{\sigma}\right)\right)$$

$$\mathcal{L} = \log(l) = \sum_{i|y_i>1} \frac{1}{\sigma} \cdot \phi\left(\frac{x_i - \mu}{\sigma}\right) + \sum_{i|y_i=0} \left(\Phi\left(\frac{1-\mu}{\sigma}\right) - \Phi\left(\frac{-1-\mu}{\sigma}\right)\right)$$

log

Después lo hace bien

4) Expectation step in the EM algorithm. Give the expression of $Q(\mu, \sigma | \mu_m, \sigma_m)$

2

Given a random variable X with density function $f(x)$ and distribution function$F(x)$, the density function of X conditional to $a < X < b$ is (with our nomenclature)

$$f_{X|y=0}(x) = \begin{cases} \dfrac{\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{1-\mu}{\sigma}\right) - \Phi\left(\frac{-1-\mu}{\sigma}\right)} & if\ -1 \leq x \leq 1 \\ 0 & otherwise \end{cases}$$

And the expectation of the variable is

$$E_{\mu,\sigma}(X|a < X < b) = \mu + \frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{1-\mu}{\sigma}\right) - \Phi\left(\frac{-1-\mu}{\sigma}\right)} \cdot \sigma$$

So,

$$Q(\mu,\sigma|\mu_{it},\sigma_{it}) = E_{\mu_{it},\sigma_{it}}(\mathcal{L}(\mu,\sigma)|X = x)$$

$$\mathcal{L}(\mu,\sigma) = \log(l) = \sum_{i=1,\dots,n} \log\left(\frac{1}{\sigma} \cdot \phi\left(\frac{x_i-\mu}{\sigma}\right)\right)$$

We know that, for a given μ_{it}, σ_{it}

$$E(x_i|y_i = 0) = \mu_{it} + \frac{\phi\left(\frac{a-\mu_{it}}{\sigma_{it}}\right) - \phi\left(\frac{b-\mu_{it}}{\sigma_{it}}\right)}{\Phi\left(\frac{1-\mu_{it}}{\sigma_{it}}\right) - \Phi\left(\frac{-1-\mu_{it}}{\sigma_{it}}\right)} \cdot \sigma_{it}$$

$$E(x_i||y_i| > 1) = y_i$$

$$\phi(x) = \frac{e^{-\frac{1}{2}\cdot x^2}}{\sqrt{2\cdot\pi}}$$

$$\mathcal{L}(\mu,\sigma) = \sum_{i||y_i|>1} \log\left(\frac{1}{\sigma} \cdot \frac{e^{-\frac{1}{2}\cdot\left(\frac{x_i-\mu}{\sigma}\right)^2}}{\sqrt{2\cdot\pi}}\right) + \sum_{i|y_i=0} \log\left(\frac{1}{\sigma} \cdot \frac{e^{-\frac{1}{2}\cdot\left(\frac{x_i-\mu}{\sigma}\right)^2}}{\sqrt{2\cdot\pi}}\right)$$

$$\mathcal{L}(\mu,\sigma) = \sum_{i||y_i|>1} -\frac{1}{2}\cdot\left(\frac{x_i-\mu}{\sigma}\right)^2 - \log\left(\sqrt{2\cdot\pi\cdot\sigma}\right) + \sum_{i|y_i=0} -\frac{1}{2}\cdot\left(\frac{x_i-\mu}{\sigma}\right)^2 - \log\left(\sqrt{2\cdot\pi\cdot\sigma}\right)$$

$$Q(\mu,\sigma|\mu_{it},\sigma_{it}) = \frac{\mathrm{n}\cdot\log(2\cdot\pi\cdot\sigma^2)}{2} + \frac{1}{2\cdot\sigma^2}\left(\sum_{i||y_i|>1} E[(x_i-\mu)^2] + \sum_{i|y_i=0} E[(x_i-\mu)^2]\right)$$

It is known that

$$E_{\mu_{it},\sigma_{it}}((X-\tau)^2|\, a < X < b) = var(X|\, a < X < b) + (E(X|\, a < X < b) - \tau)^2$$

For the case where $|y_i| > 1$

$$E_{\mu_{it},\sigma_{it}}[(X-\tau)^2] = var(X) + (E(X) - \tau)^2 = \sigma^2 + (y_i - \mu)^2$$

For the case where $y_i = 0$

$$\hat{x} = E_{\mu_{it},\sigma_{it}}((X_i-\tau)^2|\, a < X < b) = var(X|\, a < X < b) + (E(X|y_i = 0) - \tau)^2$$

Notice that for a given μ_{it} and σ_{it} the variance and expectation are fixed values.

Casi perfecto.

$$Q(\mu,\sigma|\mu_{it},\sigma_{it}) = \frac{\mathrm{n}\cdot\log(2\cdot\pi\cdot\sigma^2)}{2} + \frac{1}{2\cdot\sigma^2}\cdot\left(\sum_{i||y_i|>1}\sigma^2_{it} + (y_i-\mu)^2 + \sum_{i|y_i=0}\hat{x}_i\right)$$

2

5) Maximization step. Prove that maximizing $Q(\mu,\sigma|\mu_{it},\sigma_{it})$ in (μ,σ) is equivalent to maximizing the complete log-likelihood calculated from a sample $\widetilde{y_1},\dots,\widetilde{y_n}$ with

$$\tilde{y_i} = \begin{cases} y_i & if\ |y_i|>1 \\ E_{\mu_{it},\sigma_{it}}[X|-1<X<1] & otherwise \end{cases}$$

In this case the complete log-likelihood function is

$$\mathcal{L}(\mu,\sigma) = \frac{\mathrm{n}\cdot\log(2\cdot\pi\cdot\sigma^2)}{2} + \frac{1}{2\cdot\sigma^2}\cdot\left(\sum_{i||y_i|>1}(y_i-\mu)^2 + \sum_{i|y_i=0}(E(x_i|y_i=0)-\mu)^2\right)$$

Remind that the Q was defined as

$$Q(\mu,\sigma|\mu_{it},\sigma_{it}) = \frac{\mathrm{n}\cdot\log(2\pi\sigma^2)}{2} + \frac{1}{2\cdot\sigma^2}\cdot\left(\sum_{i||y_i|>1}\sigma^2 + (y_i-\mu)^2 + \sum_{i|y_i=0}C_1 + (C_2-\mu)^2\right)$$

We now derive the gradient of the function

$$\frac{\partial}{\partial\mu}\mathcal{L}(\mu,\sigma) = \frac{1}{2\cdot\sigma^2}\cdot\left(\sum_{i||y_i|>1}-2\cdot(y_i-\mu) + \sum_{i|y_i=0}-2\cdot(C_2-\mu)\right)$$

$$\frac{\partial}{\partial\mu}Q(\mu,\sigma) = \frac{1}{2\cdot\sigma^2}\cdot\left(\sum_{i||y_i|>1}-2\cdot(y_i-\mu) + \sum_{i|y_i=0}-2\cdot(C_2-\mu)\right)$$

Where C_2 stands for $E(x_i|y_i=0)$ in both cases (they are equivalents when maximizing μ).

When it comes to σ^2

$$\frac{\partial}{\partial\sigma}\mathcal{L}(\mu,\sigma) = \frac{1}{2\cdot\sigma^2}$$

$$\frac{\partial}{\partial\sigma}Q(\mu,\sigma) = \frac{1}{2\cdot\sigma^2}$$

The MLE for the normal distribution are the following, which are found from the analytical derivatives

$$\hat{\mu} = \bar{X}$$

$$\sigma = \frac{1}{n} \sum_{i=1,n} (x - \bar{X})^2$$

$$\frac{1}{2 \cdot \sigma^2} \cdot \left(\sum_{i=1,n} -2 \cdot (y_i - \mu) \right) = 0 \rightarrow n \cdot \mu = \sum_{i=1,n} y_i$$

casi perfecto

6. Script $\rightarrow$ 2

+ 2