Expectation- Maximisation.

22/6/24

$x = x_1, x_2, \dots$ → set of observed variables

$Z = z_1, z_2, \dots$ → latent variables.

$z_i \in Z$ → discrete

θ → parameter we wish to estimate

To find the MLE of θ, maximize the log marginal likelihood of the data

$$\ell(\theta) = \log p(x|\theta)$$

$$= \log \left[\sum_{z \in Z} p(x, z|\theta) \right]$$

$$= \log \left[\sum_{z \in Z} p(x|z, \theta)\, p(z|\theta) \right]$$

↓

has no closed-form solution

- sum inside log makes partial diff w.r.t θ difficult
- space of all possible Z is immense

So, we can construct a lower bound that can be easily optimise

q → a valid distr. function of z.

$$\ell(\theta) = \log \sum_{z \in Z} \frac{p(x, z|\theta)}{q(z)}\, q(z)$$

$$= \log \left[E_{q(z)} \left(\frac{p(x, z|\theta)}{q(z)} \right) \right]$$

$$\geq E_{q(z)} \log(\cdot)$$ → Jensen for concave function.

$= E_q(\log p - \log q)$

$= E_q[\log p(x, z|\theta)] + H(q)$

$= ELBO(\theta, q) \hookrightarrow Q(\theta)$.

Now, finding θ that maximises ELBO is easier

$$Q(\theta) = \sum_{z \in Z} q(z) \log(p(x, z|\theta))$$

$\hookrightarrow$ log is inside the sum

$H(q) \rightarrow$ doesn't depend on θ, it is sufficient to maximise $Q(\theta)$

$$ELBO(q, \theta) = E_{q(z)}\left[\log \frac{p(x, z|\theta)}{q(z)}\right]$$

$$= E_{q(z)}\left[\log \frac{p(x, z|\theta)\, p(z|x, \theta)}{p(z|x, \theta)\, q(z)}\right]$$

$$= E_{q(z)}\left[\log \frac{p(z|x, \theta)\, p(x|\theta)\, \cancel{p(z|x, \theta)}}{\cancel{p(z|x, \theta)}\, q(z)}\right]$$

$$= E_q[\log p(x|\theta)] + E_q\left[\frac{p(z|x, \theta)}{q}\right]$$

= maximum when KL $[q \| p(z|x, \theta)]$

= 0

$\rightarrow$ iterate b/w optimising ELBO w.r.t θ keeping q fixed.

then " $q(z)$ " " " θ fixed