A Toy Example for EM Algorithm

Setup:

$x \in \{0,1\}$ $\quad P_X(0) \triangleq \theta_0, \quad P_X(1) \triangleq \theta_1$

$y \in \{0,1\}$ $\quad P_{Y|X}(i|j) \triangleq \theta_{i|j}$ $\quad i \in \{0,1\}$, $j \in \{0,1\}$

1°. Single observation $\boxed{y=1}$ x hidden.

Goal: find $\theta = (\theta_0, \theta_1, \theta_{0|0}, \theta_{0|1}, \theta_{1|0}, \theta_{1|1})$ s.t. $P_Y(y;\theta)$ is maximized.

Solution: EM algo.

Initialization: $\theta_0^{(0)} = \theta_1^{(0)} = 1/2. \quad \theta_{0|0}^{(0)} = 0.9 = \theta_{1|1}^{(0)} \quad \theta_{1|0}^{(0)} = 0.1 = \theta_{0|1}^{(0)}$

These numbers are just made up for the toy example, can be anything.

E-step: $q^{(1)}(\cdot) = P_{X|Y}(x|y)$

$$\therefore q^{(1)}(0) = P_{X|Y}(0|1) = \frac{P_{X,Y}(0,1)}{P_Y(1)} = \frac{P_X(0)\,P_{Y|X}(1|0)}{P_X(0)\,P_{Y|X}(1|0) + P_X(1)\,P_{Y|X}(1|1)}$$

$$\underline{\text{using } \theta^{(0)}} \quad \frac{0.1 \times \frac{1}{2}}{0.1 \times \frac{1}{2} + 0.9 \times \frac{1}{2}} = 0.1$$

similarly $q^{(1)}(1) = 0.9$.

M-step: $\theta^{(1)} = \arg\max_{\theta} \underbrace{E_{q^{(1)}(\cdot)}\left[\log P_{X,Y}(x, y; \theta)\right]}_{(*)}$

$$(*) = \sum_{x=0}^{1} q^{(1)}(x) \log P_{X,Y}(x, 1; \theta)$$

$$= q^{(1)}(0)\left[\log \theta_0 + \log \theta_{1|0}\right] + q^{(1)}(1)\left[\log \theta_1 + \log \theta_{1|1}\right] \quad (**)$$

want to maximize $(**)$ subject to:

$\theta_0 + \theta_1 = 1. \quad \theta_0, \theta_1, \theta_{1|0}, \theta_{1|1} \in [0,1]$.

$$\Rightarrow \underbrace{\theta_0^{(1)} = 0.1 \quad \theta_0^{(1)} = 0.9}_{\text{use Lagrangian multipliers.}} \quad \cdot \quad \theta_{0|0}^{(1)} = 1 \quad \theta_{1|0}^{(1)} = 0 \quad \theta_{0|1}^{(1)} = 0 \quad \theta_{1|1}^{(1)} = 1$$

~~take partial derivative → 0.~~

2°. Multiple (IID) Observations $\boxed{y^{(1)}=1,\ y^{(2)}=0}$

$x^{(1)}$ $x^{(2)}$ hidden.

$$\log P_Y(y^{(1)}, y^{(2)};\theta) = \log \sum_{x^{(1)}}\sum_{x^{(2)}} P_{X,Y}(x^{(1)}, x^{(2)}, y^{(1)}, y^{(2)};\theta)$$

$$\geq \mathbb{E}_{q(x^{(1)},x^{(2)}|y^{(1)},y^{(2)})}\left[\log P_{X,Y}(x^{(1)}, x^{(2)}, y^{(1)}, y^{(2)};\theta)\right]$$

E-step

$$q(x^{(1)},x^{(2)}|y^{(1)},y^{(2)}) = P(x^{(1)},x^{(2)}|y^{(1)},y^{(2)})$$

$$= P(x^{(1)}|y^{(1)})\cdot P(x^{(2)}|y^{(2)})$$

due to indepeme btw $(x^{(1)}, y^{(1)})$ & $(x^{(2)}, y^{(2)})$

$$\Rightarrow q^{(1)}(0,0|1,0) = 0.1\times 0.9 = 0.09$$

$$q^{(1)}(0,1|1,0) = 0.1\times 0.1 = 0.01$$

$$q^{(1)}(1,0|1,0) = 0.9\times 0.9 = 0.81$$

$$q^{(1)}(1,1|1,0) = 0.9\times 0.1 = 0.09$$

(Computation of $P(X^{(i)}|y^{(i)})$: see 1°)

M-step: $\mathbb{E}_{q(\cdot,\cdot|y^{(1)},y^{(2)})}\left[\log P_{X,Y}(x^{(1)}, x^{(2)}, y^{(1)}, y^{(2)};\theta)\right]$

$$= \sum_{x^{(1)}=0}^{1}\sum_{x^{(2)}=0}^{1} q^{(1)}(x^{(1)},x^{(2)})\cdot \log P_{X,Y}(x^{(1)}, x^{(2)}, 1, 0;\theta)$$

$= 0.09(\log\theta_0 + \log\theta_0 + \log\theta_{1|0} + \log\theta_{0|0})$ ← $x^{(1)}=0$ $x^{(2)}=0$

$+0.01(\log\theta_0 + \log\theta_1 + \log\theta_{1|0} + \log\theta_{0|1})$ 0, 1

$+0.81(\log\theta_1 + \log\theta_0 + \log\theta_{1|1} + \log\theta_{0|0})$ 1, 0

$+0.09(\log\theta_1 + \log\theta_1 + \log\theta_{1|1} + \log\theta_{0|1})$ 1, 1

$$= \log\theta_0 + \log\theta_1 + 0.9\log\theta_{0|0} + 0.1\log\theta_{1|0} + 0.1\log\theta_{0|1} + 0.9\log\theta_{1|1}$$

Subject to $\theta_0+\theta_1=1$ $\theta_{0|0}+\theta_{1|0}=1$ $\theta_{0|1}+\theta_{1|1}=1$

Use Lagrangian multiplier $\Rightarrow \theta_0^{(1)} = \frac{1}{2} = \theta_1^{(1)}$

$$\theta_{0|0}^{(1)} = 0.9 = \theta_{1|1}^{(1)} \qquad \theta_{1|0}^{(1)} = 0.1 = \theta_{0|1}^{(1)}$$