EM Algorithm

① Method of Moments

② M.L.E Maximum Likelihood estimator

Poisson   Binomial

Normal   Gamma

Two Modes

$$f(x) = \pi_0 N(\mu_0, \sigma_0^2) + \pi_1 N(\mu_1, \sigma_1^2)$$

$$\pi_0 + \pi_1 = 1$$

$$\theta = (\pi_0, \mu_0, \sigma_0^2, \mu_1, \sigma_1^2)$$

① MOM   Five equations

$$\frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x_i-\mu_0)^2}{2\sigma_0^2}} = \frac{1}{\sigma_0}\varphi\left(\frac{x_i-\mu_0}{\sigma}\right)$$

Thursday, September 25, 2014   2:25 PM

② MLE

$$f(x_i) = \pi_0$$

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

$$f(x_i) = \pi_0 \frac{1}{\sigma_0}\varphi\left(\frac{x_i-\mu_0}{\sigma_0}\right) + \pi_1 \frac{1}{\sigma_1}\varphi\left(\frac{x_i-\mu_1}{\sigma_1}\right)$$

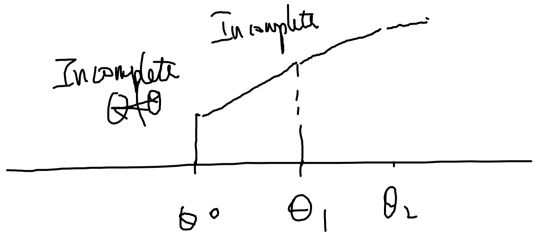
$$L(\theta) = \prod_{i=1}^{n}\left(\pi_0 \frac{1}{\sigma_0}\varphi\left(\frac{x_i-\mu_0}{\sigma_0}\right) + \pi_1 \frac{1}{\sigma_1}\varphi\left(\frac{x_i-\mu_1}{\sigma_1}\right)\right)$$

$$\Downarrow$$

$$l(\theta) = \sum_{i=1}^{n} \log(\qquad)$$

"MLE" fails

Expectation-Maximization Algorithm (EM)



Gaussian Mixture

$$\bar{Y} = (\tilde{X}, \tilde{Z}) \qquad \theta = (\pi_0, \mu_0, \sigma_0^2, \mu_1, \sigma_1^2)$$

$$L(\theta; \cancel{X}) = \prod_{i=1}^{n} \left(\pi_0 \frac{1}{\sigma_0} \varphi\left(\frac{x_i - \mu_0}{\sigma_0}\right) + \pi_1 \frac{1}{\sigma_1} \varphi\left(\frac{x - \mu_1}{\sigma_1}\right) \right)$$

$$L(\theta, \underset{\substack{X \\ Z}}{Y}) = \prod_{i=1}^{n} \sum_{j=0}^{1} 1(z_i = j)\, \pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right)$$

$$f(x_i, z_i) = \begin{cases} \pi_0 \frac{1}{\sigma_0} \varphi\left(\frac{x_i - \mu_0}{\sigma_0}\right) & \text{if } z_i = 0 \\ \pi_1 \frac{1}{\sigma_1} \varphi\left(\frac{x_i - \mu_1}{\sigma_1}\right) & \text{if } z_i = 1 \end{cases}$$

$$Q(\theta \mid \theta^t) = E\left(\log L(\theta; \tilde{X}, \tilde{Z}) \right)$$

$$= E\left\{ \sum_{i=1}^{n} \log\left(\sum_{j=0}^{1} 1(z_i = j)\, \pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right) \right) \right\}$$

$$= E\left\{ \sum_{i=1}^{n} \sum_{j=0}^{1} 1(z_i = j) \left(\log \pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right) \right) \right\}$$

$$= \sum_{i=1}^{n} \sum_{j=0}^{1} \underline{E\, 1(z_i = j)} \left(\log \underline{\pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right)} \right)$$

$$\sum_{i=1} \sum_{j=0}$$

$$= \sum_{i=1}^{n} \sum_{j=0}^{1} \log\left(\pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right) \right) P\left(Z_i = j \mid x_i, \theta \right)$$

$$P\left(Z_i = j \mid X_i, \theta \right) = \frac{P(x_i \mid Z_i = j)\, P(Z_i = j)}{\sum_{j=0}^{1} P(x_i \mid Z_i = j)\, P(Z_i = j)}$$

$$T_{j,i}^{t} \triangleq \frac{\pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right)}{\pi_0 \frac{1}{\sigma_0} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right) + \pi_1 \frac{1}{\sigma_1} \varphi\left(\frac{x_i - \mu_1}{\sigma_1} \right)}$$

$$Q(\theta \mid \theta^{T}) = \sum_{i=1}^{n} \sum_{j=0}^{1} T_{ji}^{t} \log\left(\pi_j \frac{1}{\sigma_j} \varphi\left(\frac{x_i - \mu_j}{\sigma_j} \right) \right)$$

$$= \sum_{i=1}^{n} \sum_{j=0}^{1} T_{ji}^{t} \left(\log \pi_j - \log \sigma_j - \frac{(x_i - \mu_j)^2}{2\sigma_j^2} - \ln \sqrt{2\pi} \right)$$

Maximization Step.

① π_0, π_1 $\quad \left(\sum_{i=1}^{n} T_{0i}^{t} \right) \log \pi_0 + \left(\sum_{i=1}^{n} T_{1i}^{t} \right) \log (1 - \pi_0)$

$$\pi_0^{t+1} = \frac{\sum_i T_{0i}^{t}}{\sum_i T_{0i}^{t} + \sum_i T_{1i}^{t}} = \frac{\sum T_{0i}^{t}}{n}$$

② (μ_0, σ_0^2) Set $j = 0$ $\triangleq \sum_{i=1}^{n} T_{0i}^{t} \left(-\frac{1}{2} \log \sigma_0^2 - \frac{(x_i - \mu_0)^2}{2\sigma_0^2} \right)$

$$\frac{\partial \ell}{\partial \mu_0} = \sum_{i=1}^{n} T_{0i}^{t} \left(+ \frac{\cancel{2}(x_i - \mu_0)}{\cancel{2}\sigma_0^2} \right) = \frac{1}{\cancel{2}\sigma_0^2} \sum_{i=1}^{n} T_{0i}^{t} (x_i - \mu_0)$$

$$= \sum_{i=1}^{n} T_{0i}^{t} x_i - \left(\sum_{i=1}^{n} T_{0i}^{t} \right) \mu_0$$

$$\mu_0^{t+1} = \frac{\sum T_{0i}^{t} x_i}{\sum T^{t}}$$

$$\mu_0^{t+1} = \frac{\sum T_{0i}^t x_i}{\sum T_{0i}^t}$$

$$\frac{\partial \Delta}{\partial \sigma_0^2} = \sum_{i=1}^{n} T_{0i}^t \left(-\frac{1}{2\sigma_0^2} + \frac{(x_i-\mu_0)^2}{2(\sigma_0^2)^2} \right) = 0$$

$$= \frac{1}{2\sigma_0^4} \sum_{i=1}^{n} T_{0i}^t \left(-\sigma_0^2 + (x_i-\mu_0)^2 \right) = 0$$

$$\sigma_0^{2,\,t+1} = \frac{\sum T_{0i}^t (x_i - \mu_0^{t+1})^2}{\sum T_{0i}^t}$$

Similarly (μ_1, σ_1^2)

① Calculate $$T_{ji}^t = \frac{\pi_j N(x_i, \mu_j, \sigma_j^2)}{\pi_0 N(x_i, \mu_0, \sigma_0^2) + \pi_1 N(x_i, \mu_1, \sigma_1^2)}$$

② $$\pi_j^{t+1} = \frac{\sum_{i=1}^{n} T_{ji}^t}{n}$$

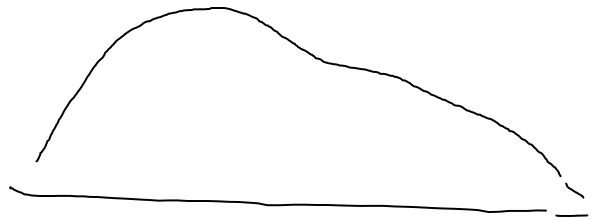
③ $$\mu_0^{t+1} = \frac{\sum T_{0i}^t x_i}{\sum T_{0i}^t} \qquad \sigma_0^{2,\,t+1} = \frac{\sum T_{0i}^t (x_i - \mu_0^{t+1})^2}{\sum T_{0i}^t}$$

④ $$\mu_1^{t+1} = \dots \qquad \sigma_1^{2,\,t+1} = \dots$$

If $\| \theta^{t+1} - \theta^t \| < \delta$ 0.0001

Based on EM

$$f(x_i) = 0.36\, N(54.6,\ 5.87) + 0.64\, N(80.1,\ 5.87)$$



1. HMM Baum-Welch
2. Multiple testing
3. Survival Analysis

KDE

- Gaussian
- Gamma
- Poisson
- Mixture of Gaussian

Parametric Model

Don't assume any parameter, Nonparametric estimation

$$X_1, X_2 \cdots X_n \sim f(x)$$

Kernel density Estimation, choose a kernel $K(\cdot)$

1. $\int K(x)\,dx = 1$
2. $K(-u) = K(u)$

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

K is choose as uniform, Epanechnikov, Normal

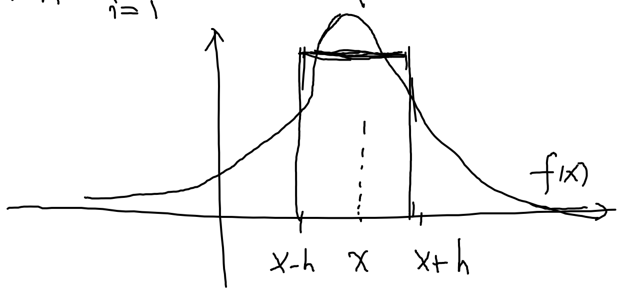
<u>Uniform</u>: $K(u) = \frac{1}{2} \mathbb{1}(|u| \le 1)$

<u>Normal</u>: $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$

<u>Uniform</u>: $K(\cancel{x}u) = \frac{1}{2} \mathbb{1}(|u| \le 1)$

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) = \frac{1}{nh} \sum_{i=1}^{n} \frac{1}{2} \mathbb{1}\left(\left|\frac{x - X_i}{h}\right| \le 1\right)$$

$$= \frac{1}{nh} \sum_{i=1}^{n} \frac{1}{2} \mathbb{1}\left(|X_i - x| \le h\right) = \frac{1}{2h} \frac{\#\{X_i \text{ falls in the interval}\}}{n}$$



$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}\underline{\varphi\left(\frac{x-X_i}{h}\right)}$$

$$= \text{dnorm}(x,\ x_i,\ h)$$

$$\underline{MISE} = E\int\left(\hat{f}(x) - f(x)\right)^2 dx$$

Pointwise. $E\left(\hat{f}(x) - f(x)\right)^2$ pointwise.

Bias- Variance decomposition

$$\left\{\begin{aligned} \text{Bias} &= \frac{h^2}{2} C_1 \\ \text{Variance} &= \frac{1}{nh} C_2 \end{aligned}\right. \qquad h = O\left(n^{-\frac{1}{5}}\right)$$

2. Silverman. $h \approx 1.06\ \hat{\sigma}\ n^{-\frac{1}{5}}$

default option of density ( ) in R

3. Cross- Validation (Leave - one - out- CV)

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right)$$

$$\hat{f}_h(X_j) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{X_j - X_i}{h}\right) \qquad f(x_i, \theta)$$

$$\prod^{n} \hat{f}(x_j) = \frac{1}{\ } \prod^{n} \sum^{n} K\left(\frac{X_j - X_i}{\ }\right)$$

$$\prod_{j=1}^{n} \hat{f}_h(x_j) = \frac{1}{nh} \prod_{j=1}^{n} \sum_{i=1}^{n} K\left(\frac{x_j - x_i}{h}\right)$$

$$\hat{f}_{h,j}(x_j) = \frac{1}{(n-1)h} \sum_{i \neq j} K\left(\frac{x_j - x_i}{h}\right)$$

$$MLCV = \frac{1}{n} \sum_{j} \log\left(\sum_{i \neq j} K\left(\frac{x_j - x_i}{h}\right) \frac{1}{(n-1)h} \right)$$

h maximize MLCV.

<u>Kodd</u> :: h, mlcv