Maximum likelihood for Bernoulli (1)

$$\boxed{P(X,\mu) = \prod_{i=1}^{n} P(x_i, \mu) = \prod_{i=1}^{n} \mu^{x_i} (1-\mu)^{1-x_i}}$$

log likelihood

$$\log P(X,\mu) = \log \prod_{i=1}^{n} P(x_i,\mu)$$

$$= \sum_{i=1}^{n} \log \mu^{x_i} (1-\mu)^{1-x_i}$$

$$= \sum_{i=1}^{n} \left\{ x_i \log \mu + (1-x_i) \log (1-\mu) \right\}$$

log likelihood of mixture Model

$$= \log \left[\prod_{i=1}^{N} \sum_{j=1}^{K} P_j \, P(\vec{X_i} \mid \theta_j) \right]$$

$$= \sum_{i=1}^{N} \log \sum_{j=1}^{K} P_j \, P(\vec{X_i} \mid \theta_j)$$

copy Page (2) from Binomial till equation (3)

$$\sum_{i=1}^{N} P(j|x_i) \frac{\partial}{\partial \mu_j} \log P(\vec{x_i}|\mu_j) = 0$$ (2)

$$\log P(\vec{x_i}|\mu_j) = \log \left[\prod_{i=1}^{N} \mu^{x_i} (1-\mu)^{1-x_i} \right]$$

$$= \sum_{i=1}^{N} \log \mu^{x_i} (1-\mu)^{1-x_i}$$

$$\frac{\partial}{\partial \mu_j} \log P(\vec{x_i}|\mu_j) = \frac{\partial}{\partial \mu_j} \sum_{i=1}^{N} \log \mu^{x_i} (1-\mu)^{1-x_i}$$

$$= \frac{\partial}{\partial \mu_j} \sum_{i=1}^{N} \left[\log \mu^{x_i} + \log (1-\mu)^{1-x_i} \right]$$

$$= \sum_{i=1}^{N} \left[\frac{x_i}{\mu} + \frac{(1-x_i)}{(1-\mu)} (-1) \right]$$

$$= \sum_{i=1}^{N} \left[\frac{x_i}{\mu} - \frac{(1-x_i)}{(1-\mu)} \right]$$

$$= \sum_{i=1}^{N} \left[\frac{x_i - \mu}{\mu(1-\mu)} \right]$$

$$\sum_{i=1}^{N} P(j|x_i) \sum_{i=1}^{N} \left[\frac{x_i - \mu}{\mu(1-\mu)} \right] = 0$$ (3)

$$\sum_{i=1}^{N} P(j|x_i) \left[\sum_{i=1}^{N} x_i - \mu \sum_{i=1}^{N} \right] = 0$$

$$\sum_{i=1}^{N} P(j|x_i) \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} P(j|x_i) \mu N = 0$$

$$\sum_{i=1}^{N} P(j|x_i) \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} P(j|x_i) \mu N$$

$$\boxed{\hat{\mu} = \frac{\sum_{i=1}^{N} P(j|x_i) \sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} P(j|x_i) N}}$$

Scanned by CamScanner