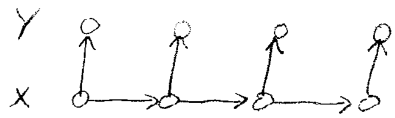
Extension



$$P\{X_0 = i\} = \pi_i \qquad i = 1, \ldots, M$$

$$p(y_s \mid x_s) \sim N(\mu_{x_s}, \sigma^2_{x_s})$$

$$p(y \mid x) = \prod_{s=1}^{N} p(y_s \mid x_s)$$

$$P_{ij} = P\{X_s = j \mid X_{s-1} = i\}$$

Each element is a vector

$$\theta = [\pi, P, \mu, \gamma]$$

matrix

$$\hat{\theta} = \underset{\theta \in \Omega}{\arg\max}\ P\{Y = y \mid \theta\}$$

$$\Omega = \left\{ \theta : \sum_{i=1}^{M} \pi_i = 1,\ \sum_{j=1}^{M} P_{ij} = 1 \quad \forall i \right\}$$

Lecture 31

$$\log p(y\, x|\theta) = \log p(y|x,\theta) + \log p(x|\theta)$$

$$Q(\theta',\theta) = E[\log p(Y,X|\theta)\,|\,Y=y,\theta]$$

$$= E[\log p(Y|X\theta')|Y=y,\theta] \qquad 1^{st} \text{ term}$$

$$+ E[\log p(X|\theta')\,|\,Y=y,\theta] \qquad 2^{nd} \text{ term}$$

1st term same as previous exampl

$$= \sum_{j=1}^{M} \sum_{s=1}^{N} \left\{ \frac{1}{2\gamma_j'}(y_s - \mu_j')^2 + \frac{1}{2}\log(2\pi\gamma_j')\right\} P\{X_s=j|Y=y,\theta\}$$

$$f_s(j|\theta) \triangleq P\{X_s=j\,|\,Y=y,\theta\}$$

2nd term = ?

$$\log p(X|\theta') = \log\left\{\pi'_{X_1} \prod_{s=1}^{N-1} P'_{X_s X_{s+1}}\right\}$$

Define

$$\tau_i = \delta(X_1 - i)$$
$$= \#\text{ of } X_1\text{'s} = i$$
$$K_{ij} = \#\text{ of } (X_s=i, X_{s+1}=j) \text{ pairs}$$

$$\log p(X|\theta') = \log\left(\left(\prod_{i=1}^{M} \pi_i'^{\tau_i}\right)\left(\prod_{\substack{i=1\\j=1}}^{M} P_{ij}'^{K_{ij}}\right)\right)$$

$$= \sum_{i=1}^{M} \tau_i \log \pi_i' + \sum_{i=1}^{M}\sum_{j=1}^{M} K_{ij} \log P_{ij}'$$

$$E[\log p(X|\theta) \mid Y=y, \theta]$$

$$= \sum_{i=1}^{M} \bar{\tau}_i \log \pi_i' + \sum_{i=1}^{M}\sum_{j=1}^{M} \bar{K}_{ij} \log P_{ij}''$$

$$\bar{\tau}_i \triangleq P\{X_1 = i \mid Y=y, \theta\} = f_1(i|\theta)$$

$$\bar{K}_{i,j} = E[K(i,j) \mid Y=y, \theta]$$

$$= \sum_{s=1}^{N-1} \underbrace{P\{X_s = i, X_{s+1} = j \mid Y=y, \theta\}}_{g_s(i,j|\theta)}$$

M-step

$$Q(\theta', \theta) =$$

$$\sum_{s=1}^{N} \sum_{j=1}^{M} \left\{ \frac{1}{2\gamma_j'} (y_s - \mu_j')^2 + \frac{1}{2} \log (2\pi \gamma_j') \right\} f_s(j|\theta)$$

$$+ \sum_{i=1}^{M} \bar{\tau}_i \log \pi_i' \qquad + \sum_{i,j=1}^{M} \bar{K}_{i,j} \log P_{ij}'$$

maximize with respect to π, P, μ, γ

1) From previous example

for $j = 1, .., M$

$$\hat{N}_j = \sum_{s=1}^{N} f_s(j|\theta)$$

$$\hat{\mu}_j' = \frac{1}{\hat{N}_j'} \sum_{s=1}^{N} y_s \, f_s(j'|\theta)$$

$$\hat{\gamma}_j' = \frac{1}{\hat{N}_j'} \sum_{s=1}^{N} (y_s - \hat{\mu}_j')^2 f_s(j'|\theta)$$

2) Maximize P_{ij} subject to $\sum_{j=1}^{M} P_{ij} = 1$ constraint

Use Lagrange Multipliers λ_i

$$\frac{\partial}{\partial P_{m,n}} \left\{ \sum_{i,j} \overline{K}_{i,j} \log P_{ij} + \sum_{i=1}^{M} \lambda_i \sum_{j=1}^{M} P_{ij} \right\} \Bigg|_{P=\hat{P}} = 0$$

$$\frac{\overline{K}_{m,n}}{\hat{P}_{m,n}} + \lambda_m = 0$$

$$\hat{P}_{m,n} = \frac{\overline{K}_{m,n}}{-\lambda_m}$$

$$\sum_{n=1}^{N} \hat{P}_{m n} = 1 \Rightarrow \boxed{\hat{P}_{m n} = \frac{\overline{K}_{m,n}}{\sum_{n=1}^{M} \overline{K}_{m,n}}}$$

3) maximize π_i subject to $\sum_{i=1}^{M} \pi_i = 1$

$$\boxed{\hat{\pi}_i = \frac{\overline{\tau}_i}{\sum_{i=1}^{M} \overline{\tau}_i}}$$

E step

(Details in HMM paper by Rabiner + Juang)

How do we compute

$$f_s(i|\theta),\ \hat{R}(i,j),\ \hat{\tau}(i)$$

note $\hat{\tau}(i) = f_1(i|\theta)$

First we must compute the following constants

$$\alpha_s(i) = P\{Y_r = y_r\ r \le s,\ X_s = i \mid \theta\}$$

$$\beta_s(i) = P\{Y_r = y_r\ r > s \mid X_s = i, \theta\}$$

Exercise: show that $\alpha_s(i)$ and $\beta_s(i)$ can be computed using the following recursions

Define $p(y_s|j) \triangleq p(y_s \mid x_s = j)$
$\sim N(\mu_j, \sigma_j^2)$

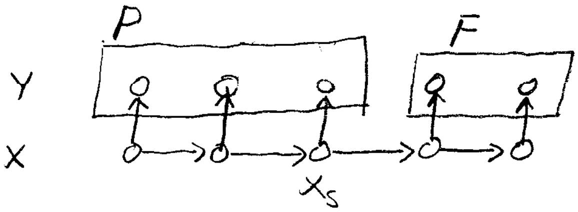
$$\boxed{\begin{aligned} \alpha_1(i) &= \pi_i\, p(y_1|i) \\ \alpha_{s+1}(i) &= \left[\sum_{i=1}^{M} \alpha_s(i) P_{ij}\right] p(y_{s+1}|j) \end{aligned}}$$

forward recursion

$$\boxed{\begin{aligned} \beta_N(i) &= 1 \\ \beta_s(i) &= \sum_{j=1}^{M} P_{ij}\, p(y_{s+1} | j)\, \beta_{s+1}(j) \end{aligned}}$$

backward recursion

Intuition



present +
P - past observations
F - future observations

$$\alpha_s(i) = P\{P \quad , x_s = i \,/\, \theta\}$$

$$\beta_s(i) = P\{F \,/\, x_s = i, \theta\}$$

$$f_s(i/\theta) = p(x_s = i \,/\, y, \theta)$$

$$= \frac{p(x_s, F, P/\theta)}{p(F, P/\theta)} \quad \overbrace{p(F,P/\theta) = p(y/\theta)}^{} \leftarrow \text{constant}$$

$$\boxed{f_s(i/\theta) = \frac{\beta_s(i)\,\alpha_s(i)}{\sum_{i=1}^{M} \beta_s(i)\,\alpha_s(i)}}$$

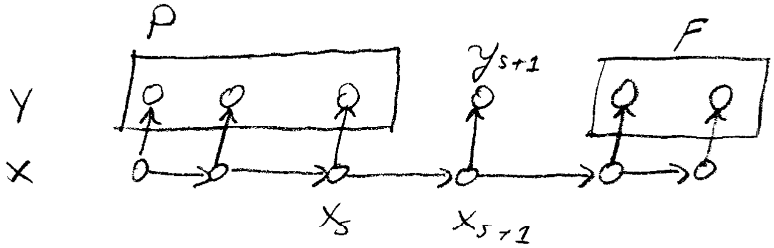
notice: $p(y/\theta) = \sum_{i=1}^{M} \beta_s(i)\,\alpha_s(i)$

$$\hat{R}(i,j) = \sum_{s=1}^{N} P\{X_s=i, X_{s+1}=j \mid Y=y, \theta\}$$

drop θ for convenience

$$= \sum_{s=1}^{N} \frac{P\{X_s=i, X_{s+1}=j, Y=y\}}{p(y) \leftarrow \text{constant}}$$

$$= \sum_{s=1}^{N} \frac{P\{F \mid X_{s+1}=j\}\, p(y_{s+1} \mid j)\, P_{X_s X_{s+1}}\, P\{X_s=i, P\}}{p(y)}$$



$$= \boxed{\sum_{s=1}^{N-1} \frac{\beta_{s+1}(j)\, p(y_{s+1} \mid j)\, P_{ij}\, \alpha_s(i)}{p(y)} = \hat{R}(i,j)}$$

Exercise: Write all update equations together

Summary of EM for HMM

M step

$$N_j^{(k+1)} = \sum_{s=1}^{N} f_s(j \mid \theta^{(k)})$$

$$\mu_j^{(k+1)} = \frac{1}{N_j^{(k+1)}} \sum_{s=1}^{N} y_s \, f_s(j \mid \theta^{(k)})$$

$$\gamma_j^{(k+1)} = \frac{1}{N_j^{(k+1)}} \sum_{s=1}^{N} (y_s - \mu_j^{(k+1)}) \, f_s(j \mid \theta^{(k)})$$

$$\pi_j^{(k+1)} = f_0(j \mid \theta^{(k)})$$

$$P_{ij}^{(k+1)} = \frac{g(i, j \mid \theta^{(k)})}{\sum_{l=1}^{M} g(i, l \mid \theta^{(k)})}$$

E step

$$\alpha_1^{(k)}(i) = p(y_1 \mid i, \theta^{(k)}) \, \pi_i^{(k)}$$

$$\alpha_{s+1}^{(k)} = \left[\sum_{i=1}^{M} \alpha_s^{(k)}(i) \, P_{ij}^{(k)} \right] p(y_{s+1} \mid j, \theta^{(k)})$$

$$\beta_N^{(k)}(i) = 1$$

$$\beta_{s-1}^{(k)}(i) = \sum_{j=1}^{M} P_{ij}^{(k)} \, p(y_s \mid j, \theta^{(k)}) \, \beta_s^{(k)}(j)$$

$$f_s(i \mid \theta^{(k)}) = \frac{\beta_s^{(k)}(i)\ \alpha_s^{(k)}(i)}{\sum_{l=1}^{M} \beta_s^{(k)}(l)\ \alpha_s^{(k)}(l)}$$

$$g(i,j \mid \theta^{(k)}) =$$

$$\sum_{s=1}^{N-1} \alpha_s^{(k)}(i)\ P_{ij}^{(k)}\ p(y_{s+1} \mid j, \theta^{(k)})\ \beta_{s+1}^{(k)}(j)$$