Modelos Ocultos de Markov II

Jorge Luis Guevara Díaz
Universidad Nacional de Trujillo
Escuela de Informática

Solucion al problema II

- Basado en programación dinámica, existe un algoritmo llamado algoritmo de viterbi, que permite encontrar la secuencia de estados Q para una secuencia de observaciones O, básicamente es igual que el algoritmo forward-backward pero utilizando la parte fordward

Algoritmo de Viterbi

- Definir una variable

$$\delta_t(i) = \max_{q_1, q_2, \cdots, q_{t-1}} P[q_1 \, q_2 \cdots q_t = i, O_1 \, O_2 \cdots O_t | \lambda]$$

- Y un array donde guardaremos el argumento que maximiza la ecuación anterior

$$\psi_t(j)$$

Algoritmo de Viterbi

- Inicialización

$$\delta_1(i) = \pi_i b_i(O_1), \qquad 1 \le i \le N$$

$$\psi_1(i) = 0.$$

- Recursión

$$\delta_t(j) = \max_{1 \le i \le N} [\delta_{t-1}(i) a_{ij}] b_j(O_t), \qquad 2 \le t \le T$$

$$1 \le j \le N$$

$$\psi_t(j) = \operatorname*{argmax}_{1 \le i \le N} [\delta_{t-1}(i) a_{ij}], \qquad 2 \le t \le T$$

$$1 \le j \le N.$$

Algoritmo de Viterbi

- Terminacion

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)]$$

$$q_T^* = \operatorname*{argmax}_{1 \leq i \leq N} [\delta_T(i)].$$

- Path backtraking

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \qquad t = T-1, T-2, \cdots, 1.$$

Solucion al problema III

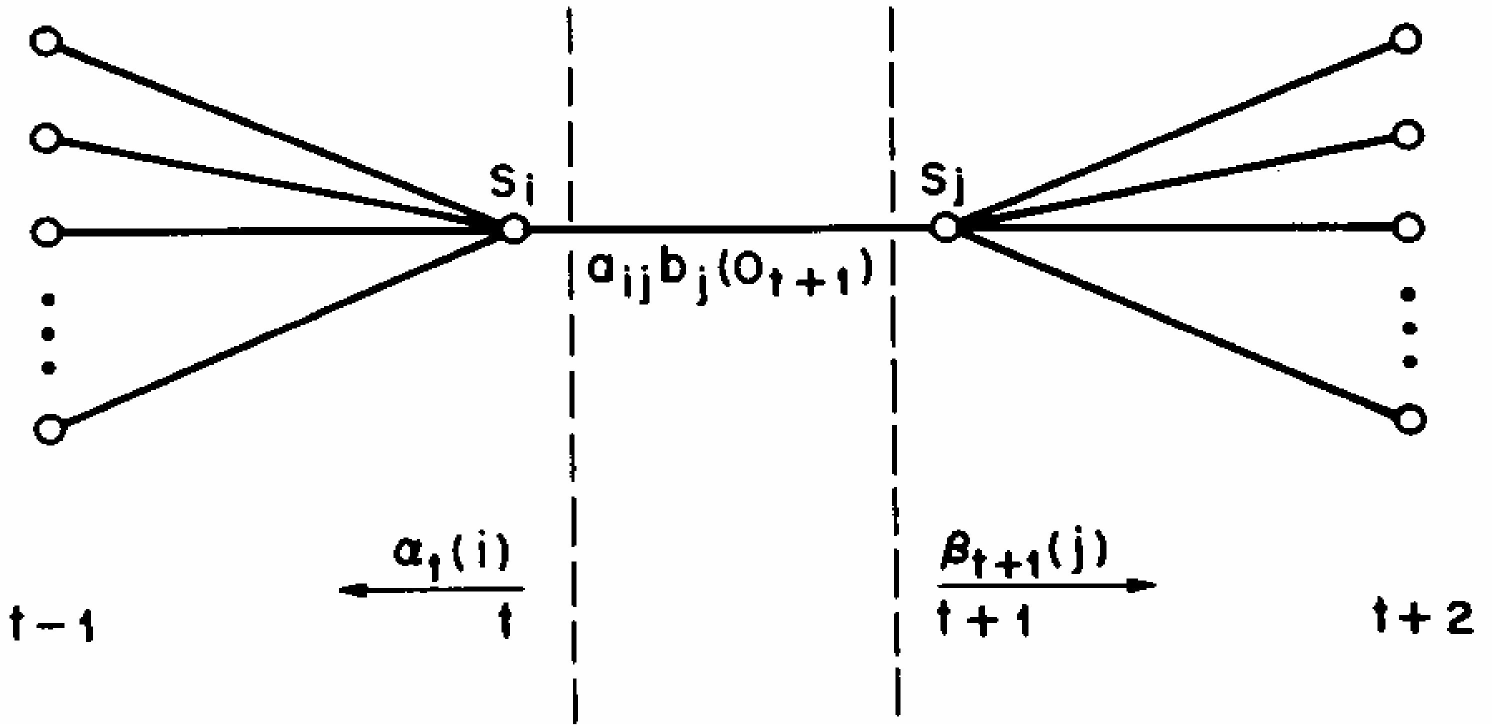
- En este problema debemos encontrar las probabilidades del modelo, para esto haremos uso del algorimo Baum-Welch

- Primeramente definiremos una variable para indicar la probabilidad de iniciar en el estado S_i en el tiempo t e ir al estado S_j en el tiempo $t + 1$ de la siguiente manera

Algoritmo Baum-Welch

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda).$$

- Para calcular esta variable gráficamente lo podemos ver como sigue:



Algoritmo Baum-Welch

- Según el gráfico puede ser escrita de la siguiente manera

$$\xi_t(i, j) = \frac{\alpha_t(i)\, a_{ij} b_j(O_{t+1})\, \beta_{t+1}(j)}{P(O|\lambda)}$$

$$= \frac{\alpha_t(i)\, a_{ij} b_j(O_{t+1})\, \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i)\, a_{ij} b_j(O_{t+1})\, \beta_{t+1}(j)}$$

Algoritmo Baum-Welch

- Definamos una variable para indicar la probabilidad de iniciar en el estado S_j en el tiempo t dada la secuencia de observaciones y el modelo

$$\gamma_t(i) = P(q_t = S_i | O, \lambda)$$

$$\gamma_t(i) = \frac{\alpha_t(i)\,\beta_t(i)}{P(O|\lambda)} = \frac{\alpha_t(i)\,\beta_t(i)}{\sum_{i=1}^{N} \alpha_t(i)\,\beta_t(i)}$$

Algoritmo Baum-Welch

- Se puede relacionar con la variable antes mencionada de la siguiente manera

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i, j).$$

Algoritmo Baum-Welch

- Estas variables podemos interpretarlas de la siguiente manera

$$\sum_{t=1}^{T-1} \gamma_t(i)$$

- Número esperado de transiciones del estado $\boldsymbol{S_i}$

Algoritmo Baum-Welch

$$\sum_{t=1}^{T-1} \xi_t(i, j)$$

- Número esperado de transiciones de $\boldsymbol{S_i}$ a $\boldsymbol{S_j}$

- Usando esas fórmulas podemos estimar los parámetros de un HMM de la siguiente manera

Algoritmo Baum-Welch

- Frecuencia esperada en el estado S_i en el tiempo $t = 1$

$$\bar{\pi}_i = \gamma_1(i)$$

- Número esperado de transiciones del estado del estado S_i al estado S_j / número esperado de transiciones en el estado S_i

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

Algoritmo Baum-Welch

- Número esperado de veces de estar en el estado ***j*** y observar el símbolo $\boldsymbol{V_k}$ / número esperado de veces de estar en el estado ***j***

$$\bar{b}_j(k) = \frac{\sum_{\substack{t=1 \\ \text{s.t. } O_t = v_k}}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}.$$

Algoritmo Baum-Welch

- Despues de calcular esos valores, se tendran las probabilidades del HMM, luego se iterará de nuevo el algoritmo con los valores ya calculados, hasta que el algoritmo converga, esta etapa es la etapa de entrenamiento del HMM

Algoritmo Baum-Welch

1. Inicializar $\lambda=(A,B,\pi)$
2. Calcular α, β, ξ
3. Estimar nuevo $\lambda'=(A,B,\pi)$
4. Remplazar λ con λ'
5. Si no converge ir a etapa 2
6. Fin

Topicos en HMM

- HMM continuos
- HMM semicontinuos
- HMM regresivos

Topicos en RAH

- Reconocimiento del habla continua usando HMM, GMM ANN, SVM, etc

Fin de la primera parte del curso