Sistemas Inteligentes

Escuela Técnica Superior de Informática

Universitat Politècnica de València

Tema B2T7:
Estimación de modelos de Markov.

DSIC – UPV

índice

Estimación de probabilidades de un modelo de Markov

Estimar las probabilidades de un modelo de Markov, M, mediante una secuencia de cadenas de entrenamiento $Y = \{y_1, \ldots, y_n\}$.

Idea básica:

1. Inicializar modelo de Markov inicial M_0 y $i = 0$ (ver traspa 10)

2. Analizar en M_i cada cadena de Y por Viterbi para obtener la secuencia de estados

3. A partir de esta secuencia de estados, contabilizar las frecuencias de uso de transiciones y símbolos.

4. Normalizar frecuencias para obtener nuevas probabilidades del modelo M_{i+1}

5. Repetir pasos 2-4 hasta convergencia ($M_i = M_{i+1}$).

Estimación mediante el algoritmo de Viterbi: ejemplo

Se dispone de tres cadenas para re-estimar las probabilidades de un modelo de Markov. Utilizando el algoritmo de Viterbi se han obtenido las siguientes *secuencias óptimas de estados* para cada *cadena*:

Cadena: **`aaaaaddcdcdcdcdccbabaabababccccb`**
Secuencia óptima de estados: **`11111222222222222333333333344444F`**

Cadena: **`aaaaaddcdcdcdccdcbababababaabcccdcbb`**
Secuencia óptima de estados: **`1111122222222222233333333334444444F`**

Cadena: **`aaaadcdcdcdcdcdcbababababccdccbaab`**
Secuencia óptima de estados: **`1111222222222222333333333444444444F`**

Estimación mediante el algoritmo de Viterbi: ejemplo (cont.)

$$\pi_1 = 3/3 = 1 \quad \pi_2 = \pi_3 = \pi_4 = 0$$

A	1	2	3	4	F
1	$4+4+3$	$1+1+1$	0	0	0
2	0	$11+11+11$	$1+1+1$	0	0
3	0	0	$9+9+8$	$1+1+1$	0
4	0	0	0	$4+6+8$	$1+1+1$

$\Rightarrow$

A	1	2	3	4	F
1	$\frac{11}{14}$	$\frac{3}{14}$	0	0	0
2	0	$\frac{33}{36}$	$\frac{3}{36}$	0	0
3	0	0	$\frac{26}{29}$	$\frac{3}{29}$	0
4	0	0	0	$\frac{18}{21}$	$\frac{3}{21}$

B	a	b	c	d
1	$5+5+4$	0	0	0
2	0	0	$6+6+6$	$6+6+6$
3	$5+5+4$	$5+5+5$	0	0
4	$0+0+2$	$1+2+2$	$4+4+4$	$0+1+1$

$\Rightarrow$

B	a	b	c	d
1	$\frac{14}{14}$	0	0	0
2	0	0	$\frac{18}{36}$	$\frac{18}{36}$
3	$\frac{14}{29}$	$\frac{15}{29}$	0	0
4	$\frac{2}{21}$	$\frac{5}{21}$	$\frac{12}{21}$	$\frac{2}{21}$

Algoritmo de reestimación por Viterbi

Input: $M_0 = (Q_0, \Sigma_0, \pi_0, A_0, B_0)$ /* Modelo inicial */
$Y = \{y_1, \ldots, y_n\}$ /* cadenas de entrenamiento */
Output: $M = (Q, \Sigma, \pi, A, B)$ /* Modelo optimizado */

$M = M_0$
repeat $M' = M$; $\pi = 0$; $A = 0$; $B = 0$
  for $k = 1$ **to** n **do**
    $m = |y_k|$ /* secuencia de estados más probable para y_k, */
    $\tilde{q}_1, \ldots, \tilde{q}_m =$ argmax$_{q_1,\ldots,q_m} P(y_k\,, q_1, \ldots, q_m \mid M')$ /* por Viterbi */
    $\pi_{\tilde{q}_1}$++; $B_{\tilde{q}_1, y_{k,1}}$++ /* actualización de contadores */
    for $t = 2$ **to** m **do** $A_{\tilde{q}_{t-1}, \tilde{q}_t}$++; $B_{\tilde{q}_t, y_{k,t}}$++ **done**; $A_{\tilde{q}_m, F}$ ++
  done
  $s = \sum_{q \in Q} \pi_q$
  forall $q \in Q$ **do** /* normalización de contadores */
    $\pi_q = \pi_q / s$
    $a = \sum_{q' \in Q} A_{q,q'}$; **forall** $q' \in Q$ **do** $A_{q,q'} = A_{q,q'}/a$
    $b = \sum_{\sigma \in \Sigma} B_{q,\sigma}$; **forall** $\sigma \in \Sigma$ **do** $B_{q,\sigma} = B_{q,\sigma}/b$
  done
until $M = M'$

Algoritmo mediante el algoritmo de Viterbi: ejercicio

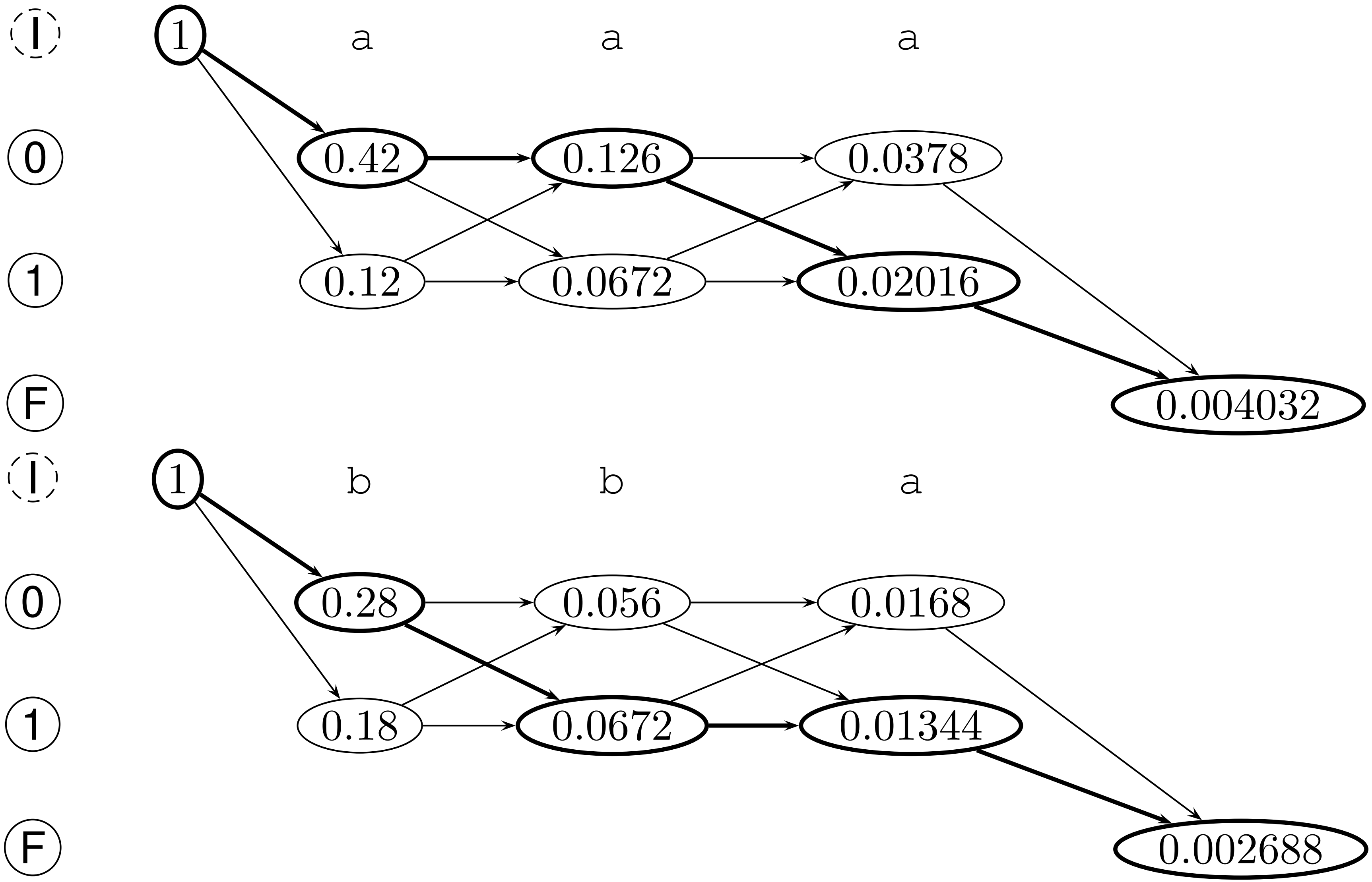
Sea M un modelo de Markov de conjunto de estados $Q = \{0, 1, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(0) = 0.7, \pi_0(1) = 0.3$; y probabilidades de transición entre estados y de emisión de símbolos:

A	0	1	F
0	0.5	0.4	0.1
1	0.3	0.5	0.2

B	a	b
0	0.6	0.4
1	0.4	0.6

Reestima los parámetros de M mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento “`a a a`” y “`b b a`”.

Ejercicio: secuencias de estados mas probables



Los pares *cadena-secuencia óptima de estados* obtenidos son:

a a a	b b a
0 0 1 F	0 1 1 F

Ejercicio: parámetros reestimados

$$\hat{\pi}_0(0) = \frac{2}{2} = 1$$
$$\hat{\pi}_0(1) = \frac{0}{2} = 0$$

A	0	1	F
0	$\frac{1}{3}$	$\frac{2}{3}$	0
1	0	$\frac{1}{3}$	$\frac{2}{3}$

B	a	b
0	$\frac{2}{3}$	$\frac{1}{3}$
1	$\frac{2}{3}$	$\frac{1}{3}$

índice

Inicialización de la re-estimación por Viterbi

Una idea elemental: Inicializar todas las probabilidades según distribuciones equiprobables.

Problema: Suele producir problemas de convergencia o convergencia a máximos locales inadecuados.

Una idea útil para modelos lineales:

- Segmentar cada cadena de Y en tantos segmentos de (aproximadamente) la misma longitud como estados tenga el modelo de Markov.
- Asignar los símbolos de cada segmento a su correspondiente estado
- Contabilizar las frecuencias de generación y transición
- Normalizar las frecuencias para obtener probabilidades iniciales requeridas

Inicialización por segmentación lineal: Ejemplo

Obtener un modelo de Markov de $N = 3$ estados mediante segmentación lineal a partir de las cadenas

$y_1 =$ aabbbcc $\qquad$ $y_2 =$ aaabbccc

$$Q = \{1, 2, 3, F\} \qquad \Sigma = \{a, b, c\}$$

$$q = \left\lfloor \frac{t \cdot N}{\mid y \mid +1} \right\rfloor + 1 \ :$$

aabbbcc	aaabbccc
1122233	11222333

$$\pi_1 = \tfrac{2}{2}, \quad \pi_2 = \pi_3 = 0$$

A	1	2	3	F
1	$\frac{2}{4}$	$\frac{2}{4}$	0	0
2	0	$\frac{4}{6}$	$\frac{2}{6}$	0
3	0	0	$\frac{3}{5}$	$\frac{2}{5}$

B	a	b	c
1	$\frac{4}{4}$	0	0
2	$\frac{1}{6}$	$\frac{5}{6}$	0
3	0	0	$\frac{5}{5}$

Inicialización por segmentación lineal

Input: $Y = \{y_1, \ldots, y_n\}$, N /* cadenas de entrenamiento, número de estados */
Output: $M = (Q, \Sigma, \pi, A, B)$ /* modelo */

$Q = \{1, 2, \ldots, N, F\}$; $\Sigma = \{y \in y_k \in Y\}$ /* estados y símbolos */
$\pi = 0$; $A = 0$; $B = 0$ /* inicialización de contadores */

for $k = 1$ **to** n **do** /* actualización de contadores por */
$q = 1$; π_q++; $B_{q,y_{k,1}}$++ /* alineamiento lineal de y_k con los estados */
for $t = 2$ **to** $|y_k|$ **do** $q' = q$; $q = \left\lfloor \frac{t}{|y_k|+1} N \right\rfloor + 1$; $A_{q',q}$++; $B_{q,y_{k,t}}$++ **done**
$A_{q,F}$ ++
done

$s = \sum_{q \in Q} \pi_q$
forall $q \in Q$ **do** /* normalización de contadores */
$\pi_q = \pi_q \,/\, s$
$a = \sum_{q' \in Q} A_{q,q'}$; **forall** $q' \in Q$ **do** $A_{q,q'} = A_{q,q'} \,/\, a$
$b = \sum_{\sigma \in \Sigma} B_{q,\sigma}$; **forall** $\sigma \in \Sigma$ **do** $B_{q,\sigma} = B_{q,\sigma} \,/\, b$
done