Modele Markov ascunse - TEMA

① Dem. formulei de recurență pt. variabila forward

Un model Markov ascuns este definit prin următorii parametrii:
- Setul de stări ascunse $Q = \{q_1, q_2, \ldots, q_N\}$ - stările prin care trece modelul
- Setul de observații $O = \{o_1, o_2, \ldots, o_M\}$ - posibilele valori vizibile generate de stările ascunse
- Vectorul de probabilități inițiale $\pi = \{\pi_i\}$ unde $\pi_i = P(S_1 = q_i)$ - probabilitatea ca modelul să înceapă în starea q_i
- Matricea de tranziție între stări $A = \{a_{ij}\}$ unde $a_{ij} = P(S_{t+1} = q_j | S_t = q_i)$ descrie probabilitatea de a trece de la starea q_i la q_j
- Matricea de probabilități de observație $B = \{b_j(k)\}$ unde $b_j(k) = P(O_t = o_k | S_t = q_j)$ - probabilitatea de a observa o_k atunci când modelul este în stare q_j

$$\alpha_t(j) = \left[\sum_i \alpha_{t-1}(i)\, a_{ij}\right] b_j(o_t)$$

Variabila forward ($\alpha_t(j)$) reprezintă probabilitatea ca secvența parțială de observații ($o_1, o_2, \ldots, o_t$) să fie generată și ca sistemul să fie în starea q_j la momentul t:

$$\alpha_t(j) = P(o_1, o_2, \ldots, o_t, S_t = q_j | \lambda)$$

S_t = starea sistemului la momentul t

$\lambda = (A, B, \pi)$ parametrii modelului HMM

Pentru a exprima $\alpha_t(j)$ în funcție de $\alpha_{t-1}(i)$, luăm în considerare că sistemul poate ajunge în starea q_j la momentul t din oricare dintre stările posibile q_i la momentul $t-1$.

Astfel avem:

$$\alpha_t(j) = P(o_1, o_2, \ldots, o_t, S_t = q_j | \lambda)$$

Exprimăm această probabilitate în funcție de starea anterioară S_{t-1}, folosind teorema totală a probabilităților:

$$\alpha_t(j) = \sum_{i=1}^{N} P(o_1, o_2, \ldots, o_t, S_{t-1} = q_i, S_t = q_j | \lambda)$$

Folosim proprietatea lanțului Markov, conform căreia probabilitatea de a fi în starea $S_t = q_j$ depinde de starea anterioară $S_{t-1} = q_i$.

Deci, descompunem astfel:

$$\alpha_t(j) = \sum_{i=1}^{N} P(o_t | S_t = q_j, o_1, o_2, \dots, o_{t-1}, S_{t-1} = q_i) \cdot P(S_t = q_j | S_{t-1} = q_i) \cdot$$

$$\cdot P(o_1, o_2, \dots, o_{t-1}, S_{t-1} = q_i | \lambda)$$

Într-un HMM, observațiile sunt independente de stările anterioare, starea curentă fiind dată. Deci:

$$P(o_t | S_t = q_j, o_1, o_2, \dots, o_{t-1}, S_{t-1} = q_i) = P(o_t | S_t = q_j)$$

iar, probabilitatea de tranziție $P(S_t = q_j | S_{t-1} = q_i)$ este dată de matricea de tranziție A:

$$\alpha_t(j) = \sum_{i=1}^{N} P(o_t | S_t = q_j) \cdot P(S_t = q_j | S_{t-1} = q_i) \cdot P(o_1, o_2, \dots, o_{t-1}, S_{t-1} = q_i | \lambda)$$

Vedem că $P(o_1, o_2, \dots, o_{t-1}, S_{t-1} = q_i | \lambda) = \alpha_{t-1}(i)$

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) \cdot a_{ij} \cdot b_j(o_t)$$

$a_{ij} = P(S_t = q_j | S_{t-1} = q_i)$ - probabilitatea de tranziție de la starea q_i la starea q_j

$b_j(o_t) = P(o_t | S_t = q_j)$ - probabilitatea de emisie a observației o_t dată de starea q_j

Astfel, avem formula de recurență:

$$\alpha_t(j) = \left(\sum_{i=1}^{N} \alpha_{t-1}(i) \cdot a_{ij}\right) b_j(o_t)$$