Tema HMM, Programare și modelare probabilistică

Adia-Ioana Romanescu (3A3)

November 5, 2024

Enunț exercițiul 1

Pe baza definiţiei unui model Markov ascuns, demonstraţi formula recurenţei pentru variabila forward de la pagina 23 a cursului 4.

Demonstrație:

Formula pe care dorim să o demonstrăm este:

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

unde:

- $\alpha_t(j) = P(o_1, o_2, \ldots, o_t, S_t = q_j|\lambda)$, adică probabilitatea de a observa secvența $o_1, o_2, \ldots, o_t$ și de a fi în starea q_j la momentul t,
- $a_{ij} = P(S_t = q_j|S_{t-1} = q_i)$ reprezintă probabilitatea de tranziție între stările q_i și q_j,
- $b_j(o_t) = P(o_t|S_t = q_j)$ este probabilitatea de observare a simbolului o_t în starea q_j.

Vom demonstra folosind metoda inducției matematice.

Pasul de bază: Verificăm formula pentru $t = 0$

Pentru $t = 1$, formula forward $\alpha_0(j)$ ar trebui să reprezinte probabilitatea ca secvența observată să conțină doar observația inițială o_1 și să fie în starea q_j la timpul $t = 1$. Conform definiției, avem:

$$\alpha_1(j) = P(o_1, S_1 = q_j|\lambda) = P(S_1 = q_j|\lambda) \cdot P(o_1|S_1 = q_j)$$

Întrucât probabilitatea inițială a stării q_j este dată de $\pi_j = P(S_1 = q_j)$ și probabilitatea observației o_1 în starea q_j este $b_j(o_1)$, rezultă că:

$$\alpha_1(j) = \pi_j \cdot b_j(o_1)$$

Aceasta este formula corectă pentru $t = 0$, deci pasul de bază este verificat.

Pasul inductiv:

Presupunem că formula este valabilă pentru $t - 1$ și vom demonstra că are loc și pentru timpul t.
Cunoscând că formula are loc pentru pasul $t - 1$, adică:

$$\alpha_{t-1}(i) = \left[\sum_{k=1}^{n} \alpha_{t-2}(k) \cdot a_{ki}\right] b_i(o_{t-1})$$

Ar trebui să rezulte că această relație este valabilă și pentru timpul t, adică:

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

Demonstrăm formula pentru timpul t

Pentru a calcula $\alpha_t(j)$, vom folosi definiția probabilităților condiționate și structura de tranziție a modelului Markov ascuns. Astfel, $\alpha_t(j)$ poate fi descompusă ca suma probabilităților de a ajunge în q_j la momentul t din fiecare stare q_i la momentul $t-1$, urmată de observarea lui o_t:

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

Aplicăm regula sumei pentru toate stările posibile q_i la timpul $t-1$:

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

Folosind independența condițională a HMM, obținem:

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

Observăm că $P(o_1, o_2, \ldots, o_{t-1}, S_{t-1} = q_i|\lambda)$ este tocmai $\alpha_{t-1}(i)$. În plus, $P(S_t = q_j|S_{t-1} = q_i) = a_{ij}$ și $P(o_t|S_t = q_j) = b_j(o_t)$.

Astfel, formula devine:

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

sau, echivalent:

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

Concluzie

Am demonstrat formula de recurență pentru variabila *forward* folosind inducția matematică, arătând că relația este valabilă pentru $t = 0$ și că, dacă este valabilă pentru $t-1$, atunci este valabilă și pentru t.