①

* Learning Method: Baum Welch algorithm *

/ /

* when The label of data set is available ⇝ maximum liklihood
* but if The label isn't available How we learn The parameters of The system ?? using a type of Expectation Maximization Called «Baum welch algorithm»

We will use Baum welch algorithm to find maximum liklihood estimate of The parameters of hidden markov chain

---

HMM parameters: $\lambda(P, E, \pi)$

$P = \{P_{ij}\} = P(X_t = j \mid X_{t-1} = i)$ is «Transition matrix»

$\pi = \{\pi_i\} = P(X_1 = i)$ is «The initial state distribution»

$E = \{e_j(O_t)\} = P(O_t = o_t \mid X_t = j)$ «is The emission matrix»

X_t ⇝ hidden state O_t ⇝ observable var.

«I have To learn These parameters»

Problem: if we have The corpus not labeled but The word sequence are given

given observations يعني كدة إيه :

The algorithm Tries to find The parameter λ That maximize The probability of The «observation»

optimal

---

aim: To find out what are The optimal parameter λ That maximize The prob. of liklihood of The observation

عشان كدة بنستخدم EM عشان نلاقي optimal λ

الي تزود فرصة حدوث ال sequence بتاعي

---

②

The algorithm idea:-

is to start with some random initial conditions on the parameter λ, estimate best values of state paths X_t to compute state path probabilities, then re-estimate the parameters λ using the just computed values of X_t, iteratively (until it convergens)

use some par. λ to get some liklihood on the hidden states once you have these liklihood or probabilities use that to compute my λ

intuition :- ① Choose some random initial values for λ

② Determine probable (state) paths $X_{t-1}=i$, $X_t=j$

③ Count the expected number of transitions b_{ij} as well as the expected number of times various emissions $b_j(O_t)$

Count the number of paths : because we are not finding the actual paths

④ Re-esimate $\lambda=(P, \Pi, E)$ using b_{ij} and $b_j(O_t)$s

Repeat ②, ③, ④ until convergence

طب هنحسب رقم ③ ازاي هنستخدم

A forward-backward algo. are used for finding probable paths

③

Forward Backward algo.

① Forward procedure:-

② Backward procedure

Forward $\downarrow$ $\qquad\qquad$ backward $\downarrow$

$X_t = i \qquad X_{t+1} = j$

$y_1 \quad y_2 \quad \cdots \quad y_t \qquad y_{t+1} \quad \cdots \quad y_{T-1} \quad y_T$

The prob. of observing y_1 to y_t & $X_t = i$ given λ

α^I_t

• $P(Y_1 = y_1, Y_2 = y_2 \cdots Y_t = y_t, X_t = i \mid \lambda)$

we want to compute α^i_t for all possible i s & all possible values of t

$\alpha^i_1 = P(Y_1 = y_1, X_1 = i \mid \lambda)$

$\pi_i b_i(y_1)$

For all possible values

$\alpha^j_{t+1} = \sum_{i=1}^{S} \alpha^i_t \, b_{ij} \, e_j(o_t)$

The probability of observing This given $X_t = i$ "STATE"

B_t

$P(Y_{t+1} = y_{t+1}, \cdots, Y_T = y_T \mid X_t = i, \Theta)$

$B^i_T = 1$

end of sequence

لأنك ليس لك الا قبلها حصل

• probability of ending The Sequence With Y_{t+1} To y_T given The State i and λ •

(This is my back ward)

$B^i_T = 1$ (end of The Sequence)

$B_i(t) = \sum_{j=1}^{S} B_j(t+1) \, b_{ij} \, e_j(o_{t+1})$

---

* We Want To Calculate These probabilities of Various paths:

$P(X_t = i \mid Y, \lambda)$ $\qquad$ $P(X_t = i, X_{t+1} = j \mid Y, \lambda)$

هقدر احسب π من هنا

هقدر احسب

Transition Probability

* Finding probabilities of paths:-

① $\delta^{i}_{t} = P(X_t = i \mid Y, \lambda) = \dfrac{P(X_t = i, Y \mid \lambda)}{P(Y \mid \lambda)} = \dfrac{P(X_t = i, Y \mid \lambda)}{P(\sum_{j} X_t = j, Y \mid \lambda)}$

(all possible كانها states at time)

$$\therefore \delta^{i}_{t} = \frac{\alpha^{i}_{t} \, \beta^{i}_{t}}{\sum_{j=1}^{S} \alpha^{j}_{t} \, \beta^{j}_{t}}$$

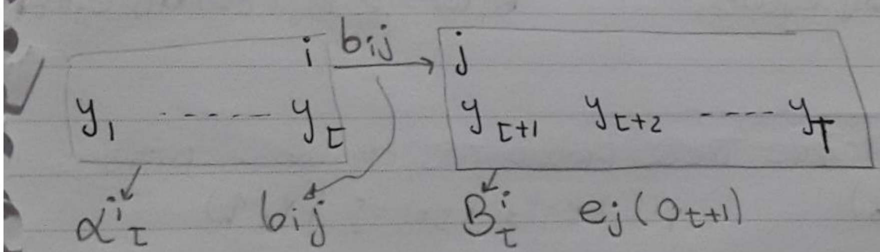
② Probability of being in state i & j at time t and $t+1$ respectively given Y and λ

$\tau_{ij}(t) = P(X_t = i, X_{t+1} = j \mid Y, \lambda)$

$= \dfrac{P(X_t = i, X_{t+1} = j, Y \mid \lambda)}{\sum\sum P(X_t = i, X_{t+1} = j, Y \mid \lambda)}$

all possible i & j

j و i فيها :-
شكلنا كدة هنستخدمها
عشان نجيب $b_{i,j}$
☺



$$\therefore \tau_{ij}(t) = \frac{\alpha^{i}_{t} \, b_{ij} \, \beta^{t+1}_{j} \, e_j(o_{t+1})}{\sum_{j=1}^{S}\sum_{i=1}^{S} \alpha^{i}_{t} \, a_{ij} \, \beta^{t+1}_{j} \, e_j(o_{t+1})}$$

* I found some of the probabilities of possible state paths

/ /

* How To re-estimate The Parameters again:-

$\pi_i = \delta_i(1)$ expected no of Times state i was seen at Time 1

* Transition probability:

$$b_{ij} = \frac{\sum_t \tau_{ij}(t)}{\sum_t \delta_i(t)}$$

بنقسمها على overall Time لما شفت T وكنت في State i

$e_j(o_k)$

How many Times we are in state j by The observation we can divide by # of Times you are in state J

$$e_j(o_k) = \frac{\sum_{t=1}^{T} 1_{y_t = o_k} \, \delta_j(t)}{\sum_{t=1}^{T} \delta_j(t)}$$

$$1_{y_t = o_k} = \begin{cases} 1 & \text{if } o_k = o_t \\ 0 & \text{otherwise} \end{cases}$$

identicator
