Homework 4 Part A: HMM, Viterbi

Aditya Shastry

October 20, 2016

1 HMM

1. As the state S_1 is not fixed for the RHS, the output $\square$ can be emitted from either of the states A or B. As the state S_1 is fixed as B for the LHS, both sides as not equal. So, this equation is FALSE

2. As the next state of an HMM depends on the previous state and not the emission of the previous state, both sides are equal. So, this equation is TRUE.

3. Given that the state S_1 is A and S_3 is also A, as there are different transmission probabilities from A to itself and another state B and vice versa,the second state can be any of A or B. As the emission values for both $\square$ and $\triangle$ from states A and B are different, both sides of the equation are not equal. So, this equation is FALSE.

4. $P(o_1 = \square) = P(o_1 = \square | S_1 = A) \times P(S_1 = A|Start) + P(o_1 = \square | S_1 = B) \times P(S_1 = B|Start)$
$= 0.5 \times 0.5 + 0.7 \times 0.5 = 0.6$

5. To compute $P(S_1 = A|o_2 = \square, S_3 = END)$. This can be computed using the total probability rule as below:
$\frac{P(S_1=A,S_2=A,o_2=\triangle,S_3=END)+P(S_1=A,S_2=B,o_2=\triangle,S_3=END)}{P(S_1=A,S_2=A,o_2=\triangle,S_3=END)+P(S_1=A,S_2=B,o_2=\triangle,S_3=END)+P(S_1=B,S_2=A,o_2=\triangle,S_3=END)+P(S_1=B,S_2=B,o_2=\triangle,S_3=END)}$
$= \frac{(0.2)(0.5)(0.5)+(0.3)(0.3)(0.2)}{(0.2)(0.5)(0.5)+(0.3)(0.3)(0.2)+(0.4)(0.5)(0.5)+(0.4)(0.3)(0.5)} = \frac{0.068}{0.228} = 0.298$

2 Viterbi (log-additive form)

1. Under HMM, $P(\vec{y}, \vec{w}) = \prod_{t=1}^{T} P(y_t|y_{t-1})P(w_t|y_t) = \prod_{t=1}^{T} P(y_t|y_{t-1}) \prod_{t=1}^{T} P(w_t|y_t)$
Taking log on both sides,
$log(P(\vec{y}, \vec{w})) = log(\prod_{t=1}^{T} P(y_t|y_{t-1}) \prod_{t=1}^{T} P(w_t|y_t)) = log(\prod_{t=1}^{T} P(y_t|y_{t-1})) + log(\prod_{t=1}^{T} P(w_t|y_t))$
$= \sum_{t=1}^{T} log(P(y_t|y_{t-1})) + \sum_{t=1}^{T} log(P(w_t|y_t))$
So, for the notation $G(\vec{y}) = \sum_{t=1}^{T} B_t(y_t) + \sum_{t=1}^{T} A(y_{t-1}, y_t)$,
$B_t(y_t) = log(P(w_t|y_t))$, which is the emission probability of each output given a state, and
$A(y_{t-1}, y_t) = log(P(y_t|y_{t-1}))$, which is the transmission probability between consecutive states.

2. The viterbi function has been completed. The function returns the sequence of states with the highest goodness score. This has been tested with 4 iterations of the randomized_test function and the viterbi sequence is same as the one returned by exhaustive function for all of them.