

ECE59500NL Lecture 8: Hidden Markov Models

Jeffrey Mark Siskind

School of Electrical and Computer Engineering

Spring 2021



© 2021 Jeffrey Mark Siskind. All rights reserved.

Size Hyperparameters and Indices

- ▶ L : number of sentences
- ▶ l : sentence index
- ▶ $I(l)$: number of words in sentence l
- ▶ i : word position in sentence l
- ▶ K : number of words in dictionary
- ▶ k : word index in dictionary
- ▶ J : number of states
- ▶ j : state index

Random Variables

- ▶ $T_{li} \in \{1, \dots, J\}$: state index at position i of sentence l
- ▶ $W_{li} \in \{1, \dots, K\}$: word index at position i of sentence l

Model Parameters

- ▶ $b_j = \Pr(T_{l1} = j) \ (\forall l)$
- ▶ $a_{jj'} = \Pr(T_{li+1} = j' | T_{li} = j) \ (\forall l)(\forall i)$
- ▶ $c_{jk} = \Pr(W_{li} = k | T_{li} = j) \ (\forall l)(\forall i)$

Sampling

- ① $i := 1$
- ② Sample T_i from b .
- ③ Sample W_i from c_{T_i} .
- ④ $i := i + 1$
- ⑤ Sample T_i from $a_{T_{i-1}}$.
- ⑥ Go to 3.

Forward Probabilities

► $\alpha_{lij} = \Pr(W_{l1} = k_{l1}, \dots, W_{li} = k_{li}, T_{li} = j | a, b, c)$

Backward Probabilities

► $\beta_{lij} = \Pr(W_{li+1} = k_{li+1}, \dots, W_{U(l)} = k_{U(l)}, T_{li} = j | a, b, c)$

Forward Algorithm

- ▶ $\alpha_{l1j} = b_j c_{jW_{l1}} \quad (\forall l)(\forall j)$
- ▶ $\alpha_{lij} = \left(\sum_{j'} \alpha_{li-1j'} a_{j'j} \right) c_{jW_{li}} \quad (\forall l)(\forall i > 1)(\forall j)$

Backward Algorithm

- ▶ $\beta_{ll(l)j} = 1 \ (\forall l)(\forall j)$
- ▶ $\beta_{lij} = \sum_{j'} a_{jj'} c_{j'} w_{li+1} \beta_{li+1j'} \ (\forall l)(\forall i < I(n))(\forall j)$

Likelihood Estimation

- ▶ $\Pr(W_{l1} = k_{l1}, \dots, W_{l(l)} = k_{l(l)} | a, b, c)$
- ▶ $= \sum_j \alpha_{l(l)j}$
- ▶ $= \sum_j b_j c_{lW_{l1}} \beta_{l1j}$

Forward-Backward Algorithm

► $\gamma_{lij} = \Pr(W_{l1} = k_{l1}, \dots, W_{lI(l)} = k_{lI(l)}, T_{li} = j | a, b, c) \propto \alpha_{lij} \beta_{lij}$

Baum-Welch Reestimation Procedure

- ▶ $b_j \propto \sum_l \gamma_{lij} \quad (\forall j)$
- ▶ $a_{jj'} \propto \sum_{l,i} \gamma_{lij} \gamma_{li+1j'} \quad (\forall j)(\forall j')$
- ▶ $c_{jk} \propto \sum_{l,i, W_{li}=k} \gamma_{lij} \quad (\forall j)(\forall k)$