Hidden Markov Models

Overview

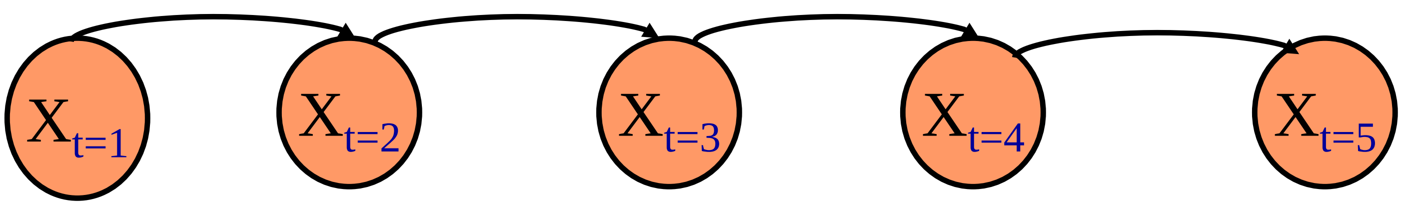
- Markov models
- Hidden Markov models(HMM)
- Issues Regarding HMM
- Algorithmic approach to Issues of HMM

Markov Models

- A Markov model is a finite state machine with N distinct states begins at (Time t = 1) in initial state .
- It moves from current state to Next state according to the transition probabilities associated with the Current state
- This kind of system is called Finite or Discrete Markov model.

Markov Property

- Markov Property : The Current state of the system depends only on the previous state of the system (Memory Less)
- The State of the system at Time [T+1] depends on the state of the system at time T.
-



Discrete Markov Model : Example

- A Discrete Markov Model with 5 states.
- Each aij represents the probability of moving from state ‘ i’ to state ’j’.

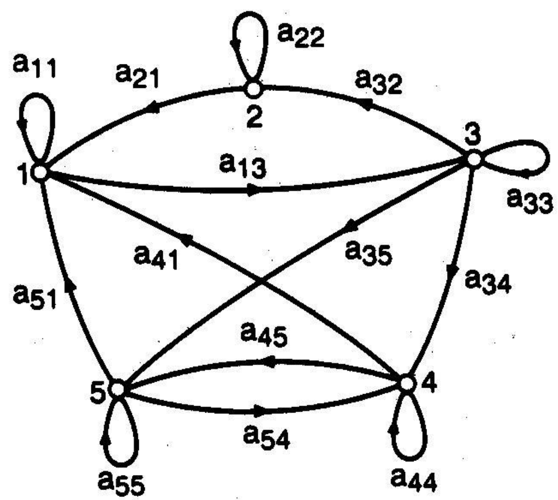


Figure 6.1 A Markov chain with five states (labeled 1 to 5) with selected state transitions.

Example

- The probability to start in a given state I is π_i.
- The Vector π represents the start probabilities.
- To define Markov model, the following probabilities have to be specified: transition probabilities $a_{ij} = P(S_i \mid S_j)$ and

  initial probabilities

$$\pi_i = P(S_i)$$

Hidden Markov Models

- A Hidden Markov model is a statistical model in which the system being modelled is assumed to be markov process with unobserved hidden states.

- In Regular Markov models the state is clearly visible to others in which the state transition probabilities are the parameters only where as in HMM the state is not visible but the output is visible.

Description

- It consists of set of states : S1,S2,S3…….sn.
- Process moves from One state to another state generating a sequence of states Si1,Si2,….Sik…
- Markov chain property: probability of each subsequent state depends only on what was the previous state

  P(Sik | Sk1,Si2,……..Sik-1) = P (Sik | Sik-1)
- States are not visible, but each state randomly generates one of M observations (or visible states)

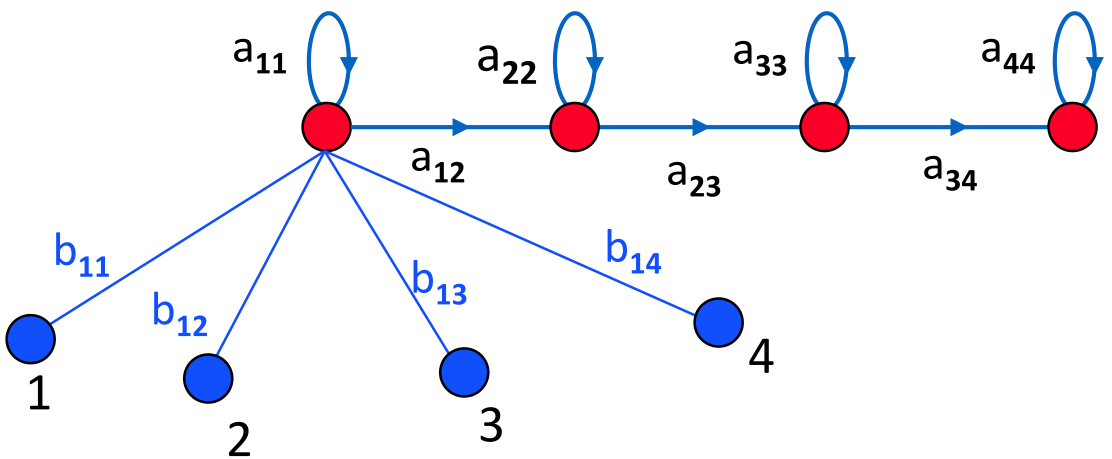
  V = { v1,v2,v3…..vk…..}

Essentials

- To define hidden Markov model, the following probabilities have to be specified: matrix of transition probabilities $A=(a_{ij})$, $a_{ij}= P(s_i \mid s_j)$, matrix of observation probabilities $B=(b_i(v_m))$, $b_i(v_m) = P(v_m \mid s_i)$ and a vector of initial probabilities $\pi=(\pi_i)$, $\pi_i = P(s_i)$. Model is represented by $M=(A, B, \pi)$.

Hidden Markov model
(Probabilistic finite state automata)

- The Scenarios where states cannot be directly observed.
- We need an extension i.e, Hidden Markov Models

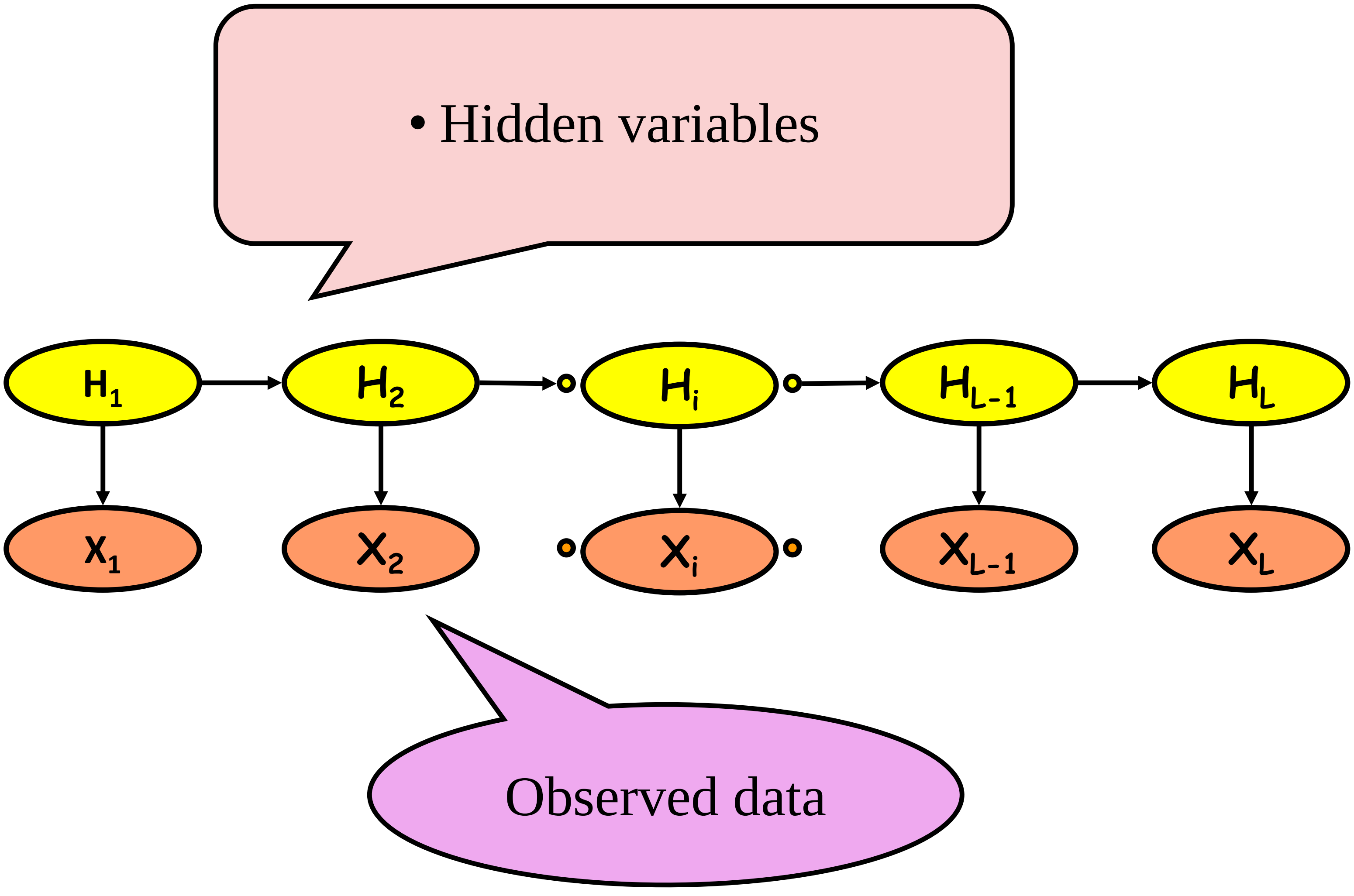


Hidden Markov model

- a_{ij} are state transition probabilities.
- b_{ik} are observation (output) probabilities.

- $b_{11} + b_{12} + b_{13} + b_{14} = 1$,
- $b_{21} + b_{22} + b_{23} + b_{24} = 1$.

Hidden Markov Models - HMM



Main issues ?

- **Evaluation problem:** Given the HMM

  M = { A, B, π } and observation sequence

  O = o1,o2 ……ok, Caluculate the probability that model m has generated sequence O.

- **Decoding problem : :** Given the HMM

  M = { A, B, π } and observation sequence

  O = o1,o2 ……ok, Caluculate the most likely sequence of hidden states Si that generated sequence O.

Problems ?

- **Learning Problem :** Given some training observation sequences O = o1,o2 ……ok, and general structure of HMM(visible and hidden states) Determine HMM parameters that best fit the training data.

Solutions to evaluation problem ?

- Evaluation problem: For this problem We use an Forward- Backward algorithm
- This algorithm mainly consists of defining a forward or backward variable as the joint probability of partial state sequence such as

  O = o1, o2, …..ok and the hidden state Si at time k is $\alpha k(i) = p(o1\ o2\ o3 \ldots ok, Qk = Si)$.
- The three states in this algorithm are initilaisation, forward recursion and termination.

Solutions to learning problem

- The solution to this problem is to estimate parameters.
- The parameters that need to be estimated are Transmission probabilities and emission probabilities. Since they sum upto 1, only 2 transmission and 2 estimation parameters are to be found.
- More parameter estimation be done using

  Baun-Welch algorithm

Solution to decoding problem ?

- Decoding problem: Viterbi Algorithm
- In this algorithm we go through the observations from start to end referring a state of hidden machine for each observation.
- We also record the values of Overall Probability,

  Viterbi path (sequence of states) and the viterbi probability(Probability of observed state sequences in viterbi path)
- The probability of possible step given its corresponding observation is probability of transmission times the emission probability.

Viterbi algorithm

- Overall Probability : Multiply each new probability with the old one.
- Viterbi probability : Take the highest next step probability and multiply with the next step viterbi probability.
- Viterbi path : Add the next step path to viterbi path.

Viterbi algorithm with example

- A person basically does 3 activities walk, clean and shop depending on the weather conditions?

- Possibility of weather conditions are ‘Rainy’ and ‘sunny’.

- In this example weather condition states are **hidden** and we will know the weather condition by her activities.

Viterbi algorithm with example

- As we discussed in earlier slides for every hidden markov model (HMM) we need an Transition probabilities and Emission probabilities.
- The transition probabilities are :

P(R ---> R) (Rainy stays rainy) = 0.7

P(R ---> S) (Rainy turns into Sunny) = 0.3

P (S ---> S) (Sunny stays into sunny) = 0.6

P (S----> R) (Sunny turns into rainy) = 0.4

Viterbi algorithm with example

- The Observations of her activities is

  If it is Rainy the behaviour is

  Walk = 0.1

  Clean = 0.5

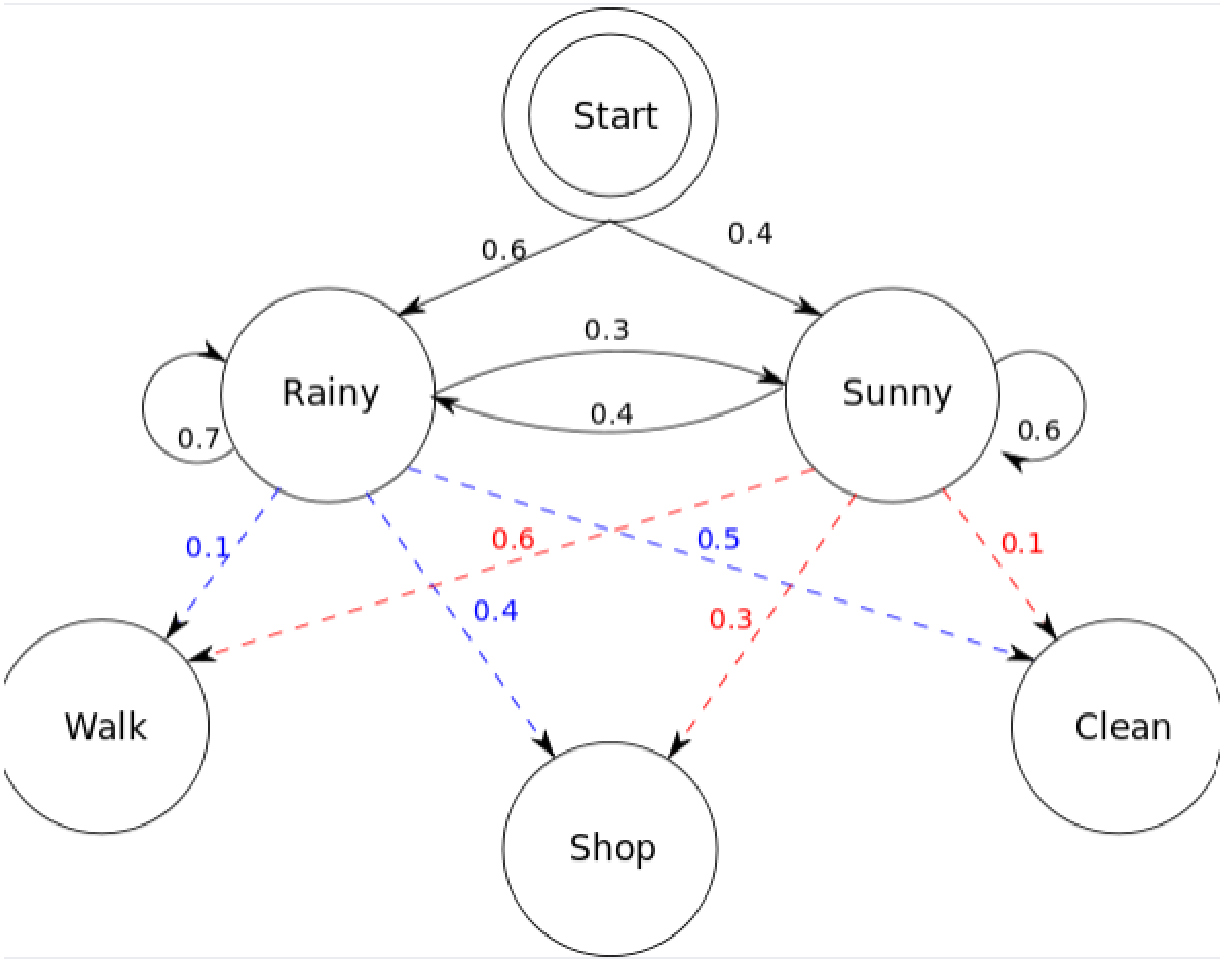
  Shop = 0.4
- If it is Sunny the behaviour is

  Walk = 0.6

  Clean = 0.3

  Shop = 0.1

Viterbi algorithm with example



Viterbi algorithm with example

- If the observations are WCSW
- Then according to algorithm find the overall prob, vit Prob, vit_path.
- In vi_path you get the sequence of states which need to compare with the original states in order to know the accuracy
- Through many examples the accuracy varies between 80-90%

Applications of HMM

- Cryptanalysis
- Speech Recognition
- Pattern Recognition
- Activity Recognition
- Machine Translation

References

- http://en.wikipedia.org/wiki/Hidden_Markov_model
- www.evl.uic.edu/shalini/coursework/hmm
- www.cedar.buffalo.edu/~govind/CS661/Lec12.**ppt**
- www.bios.niu.edu/johns/bioinf.../**Hidden**%20**Markov**%20**Models.ppt**
- www.ece.drexel.edu/gailr/ECE-S690-503/**markov_models.ppt**.pdf

Thank you