Brief Overview of Hidden Markov Model

Sagar Gaikwad

Big Data Lab

Agenda

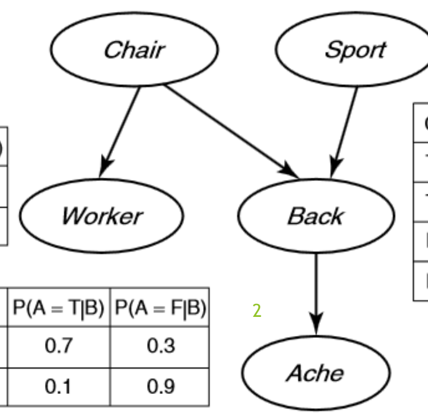
- Bayesian Network
- Discrete Markov Model
- Hidden Markov Model
- Limitations of HMM
- Extensions of HMM
- Application of HMM in Predictive Analysis
- References

Bayesian Networks

- Probabilistic Directed Acyclic Graph Model
- Each node in the network represents random variable
- Edges between the nodes represent the probabilistic dependencies between the nodes
- Conditional probability matrices are stochastic in nature
- Parameters of the model consistent with the Markov property

P(C = T)	P(C = F)
0.8	0.2

P(S = T)	P(S = F)
0.02	0.98

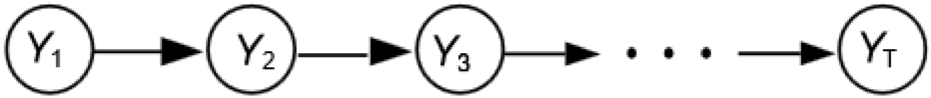


C	P(W = T|C)	P(W = F|C)
T	0.9	0.1
F	0.01	0.99

C	S	P(B = T|C,S)	P(B = F|C,S)
T	T	0.9	0.1
T	F	0.2	0.8
F	T	0.9	0.1
F	F	0.01	0.99

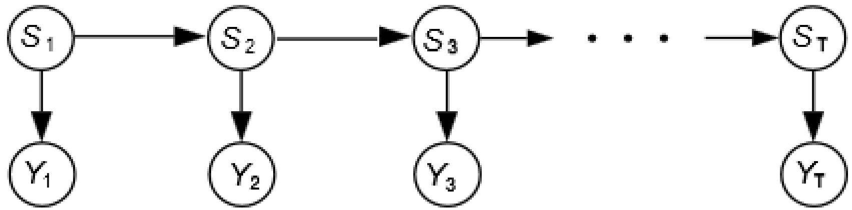
B	P(A = T|B)	P(A = F|B)
T	0.7	0.3
F	0.1	0.9

Discrete Markov Model



- Discrete **Markov model** is a stochastic model used to model randomly changing systems
- It is assumed that future states depend only on the present state and not on the sequence of events that preceded it (that is, it assumes the Markov property).
- Generally, this assumption enables reasoning and computation with the model that would otherwise be intractable.
- An example use of a Markov chain is Markov Chain Monte Carlo, which uses the Markov property to prove that a particular method for performing a random walk will sample from the joint distribution of a system.

Hidden Markov Model



- A hidden Markov model is a Markov chain for which the state is only partially observable.
- An HMM can be presented as the simplest dynamic Bayesian network.
- Observations are related to the state of the system, but they are typically insufficient to precisely determine the state.
- In a *hidden* Markov model, the state is not directly visible, but output, dependent on the state, is visible.
- Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics and time-series data.

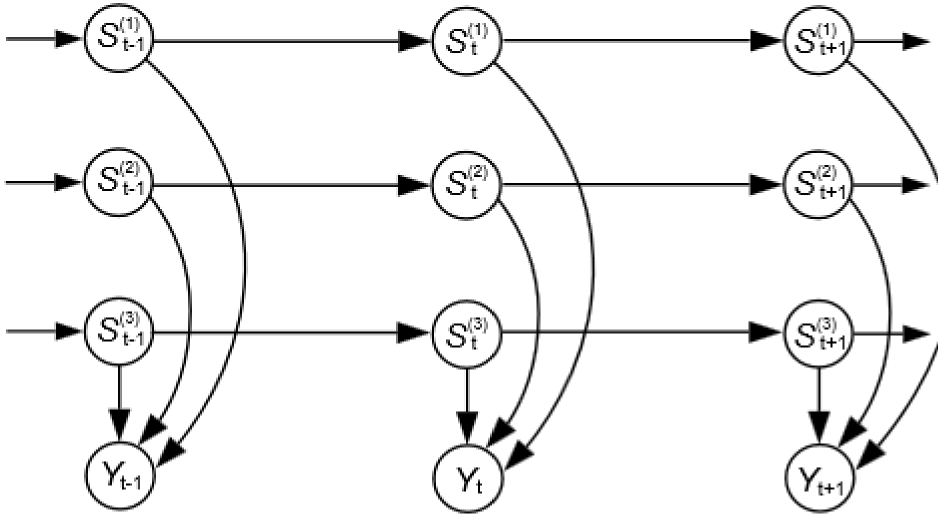
Limitations of HMM

- In image processing,
  - M objects
  - K positions
- K^M possible distinct states to model the system
- Inefficient and difficult to interpret
- This unconstrained HMM would require K^2M parameters in the transition matrix
- This encourages random interactions of all variables, which will not help to discover the structure of system.

Extensions of HMM

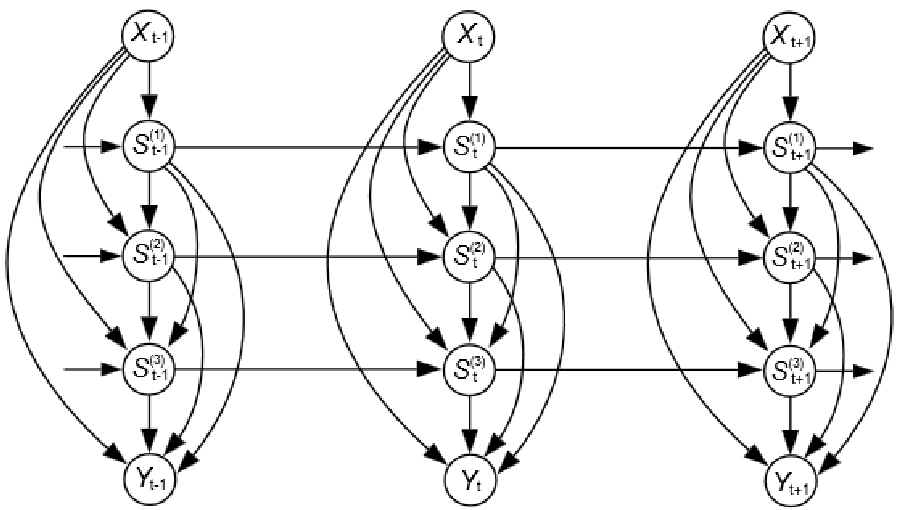
- Factorial HMMs
- Underlying state transitions are constrained (i.e. each state is independent)
- Each state variable evolves according its own dynamics and uncoupled from other state variables

$$P(S_t|S_{t-1}) = \prod_{m=1}^{M} P(S_t^{(m)}|S_{t-1}^{(m)}).$$



Extensions of HMM

- Tree structured HMM
- This architecture can be interpreted as a probabilistic decision tree with Markovian dynamics linking the decision variable.
- Note that, the interaction between state variables is constrained
- Provide useful starting point for modeling time series with both temporal and spatial structure of multiple resolutions



Application of HMM in predictive analysis

- As per requirement, both extensions can be used in predictive analysis.
- Factorial HMMs can be used in case of mutually independent state variables.
- Factorial HMMs can also provide a very efficient approach to construct a model with state variables completely decoupled.
- Tree structured HMMs provide a more open and controlled interaction between state variables which maintains the efficiency of model while allowing the state variables to interact with each other.

References

- Ghaharamani, Z (2001) “An Introduction to Hidden Markov Models and Bayesian Networks”, International Journal of Pattern Recognition and Artificial Intelligence.
- http://en.wikipedia.org/wiki/Hidden_Markov_model
- http://en.wikipedia.org/wiki/Markov_chain

Questions ?