Probabilistic Reasoning

Unit # 17

Stochastic Process

- A stochastic process is defined to be an indexed collection of random variables $\{X_t\}$, where the index t runs through a given set T.
- For example, X_t might represent the inventory level of a particular product at the end of week t.

Markovian Property

- A stochastic process {Xt} is said to have the Markovian property if $P(X_{i+1} = j \mid X_0 = k_0, X_1 = k_1, \ldots\ldots, X_{t-1} = k_{t-1}, X_i = i\} = P(X_{i+1} = j \mid X_i=i\}$, for t=0, 1, … and every sequence I, j, k_0, k_1, ….., k_{t-1}.
- This Markovian property says that the conditional probability of any future "event" given any past "events" and the present state $X_t = i$, is independent of the past events and depends only upon the present state.

Markov Chain

- A stochastic process $\{X_t\}$ (t = 0, 1, ….) is a Markov chain if it has the Markovian property.
- The conditional probabilities $P\{X_{t+1} =j \mid X_t =i\}$ for a Markov chain are called (one-step) transition probabilities.
- If for each i and j

$$P\{X_{t+1} = j \mid X_t =i\} = P(X_1 =j \mid X_0 =i\}$$

- Then the (one-step) transition probabilities are said to be stationary. Thus, having stationary transition probabilities implies that the transition probabilities do not change over time.

Weather Example

- Xt = 0 if day t is dry
- = 1 if day t has rain
- P{Xt+1 = 0 | Xt = 0} = 0.8
- P{Xt+1 = 0 | Xt = 1} = 0.6
- P = $\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix}$

Stock Example

- At the end of a given day, the price of a stock is recorded. If the stock has gone up, the probability that it will go up tomorrow is 0.7. If the stock has gone down, the probability that it will go up tomorrow is 0.5.
- This is a Markov chain, where the possible states for each day are as follows:
  - State 0: the stock increased on this day
  - State 1: the stock decreased on this day
- P = $\begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$

Gambling Example

- Suppose that a player has $1 and with each play of the game wins $1 with probability p > 0 or loses $1 with probability 1 – p. The game ends when the player either accumulates $3 or goes broke.
- This game is a Markov chain with the states representing the player's current holding of money, that is, 0, $1, $2, or $3.
- P = ?

Chapman Kolmogorov Equation

- The n-step transition probability matrix $P^{(n)}$ can be obtained by computing the nth power of the one-step transition matrix P.
- For the weather example,
- $P^{(2)} = P \cdot P = \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{pmatrix}$
- If the weather is in state 0 (dry) on a particular day, the probability of being in state 0 two days later is 0.76. Similarly, if the weather is in state 1 now, the probability of being in state 0 two days later is 0.72.

Steady-State Probabilities

- Steady-state equations:
  - $\pi_j = \sum_{i=0}^{M} \pi_i p_{ij}$ for j = 0, 1, ….., M
  - $\sum_{j=0}^{M} \pi_i = 1$
- For the weather example:
- $\pi_0 = \pi_0 p_{00} + \pi_1 p_{10}$
- $\pi_1 = \pi_0 p_{01} + \pi_1 p_{11}$
- $1 = \pi_0 + \pi_1$

i=0

Application of DBN in Anti-Money Laundering (Saleha and Sajjad)

- Suspicious activity reporting (SAR) has been a crucial part of anti-money laundering (AML) systems.
- Financial transactions are considered suspicious when they deviate from the regular behavior of their customers.
- Money launderers pay special attention to keep their transactions as normal as possible to disguise their illicit nature.

Application (Cont'd)

- This study presents an approach, called SARDBN (Suspicious Activity Reporting using Dynamic Bayesian Network), that employs a combination of clustering and dynamic Bayesian network (DBN) to identify anomalies in sequence of transactions.
- SARDBN applies DBN to capture patterns in a customer's monthly transactional sequences as well as to compute an anomaly index

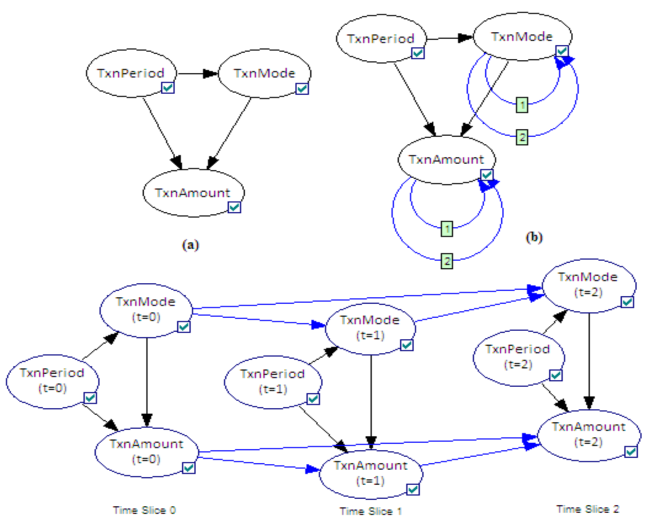
Application (Cont'd)

- The first step is to form clusters of customers that exhibit similar transactional pattern.
- The similarity in the transactional behavior is assessed by a customer's average monthly credit and debit amounts, frequency of credit and debit transactions, and delay in two consecutive transactions.

Application (Cont'd)

- Once clustering is performed, the next step is to learn the parameters of a DBN for each cluster.
- The structure of these DBNs can vary in different scenarios and is devised with the aid of subject matter experts.
- For this study, the three variables under consideration are transaction amount (TxnAmount), mode of transaction (TxnMode) (e.g. cheque withdrawal/deposit, ATM withdrawal, salary transfer, POS payment, bank draft), and period of transaction (TxnPeriod) (i.e. start/middle/end of month).

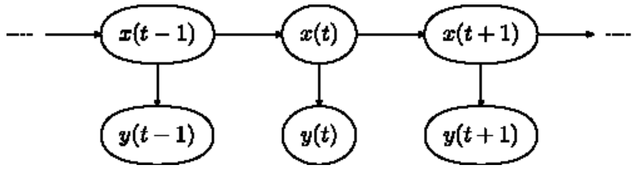
Application (Cont'd)



(c)

Hidden Markov Model

- A hidden Markov model (HMM) is a Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states.
- An HMM can be considered as the simplest dynamic Bayesian network.



Hidden Markov Model (Cont'd)

- In simpler Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters.
- In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible.
- Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states.
- Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; even if the model parameters are known exactly, the model is still 'hidden'.