Graphical Models:

Hidden Markov Models: Formulation

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

Objective

Introduce Hidden Markov Models

Objective

Illustrate HMM with intuitive examples

Hidden Markov Models

- Hidden Markov Models (HMMs) are a type of dynamic Bayesian Network
  - Modeling a process indexed by time
- “Hidden”: the observations are due to some underlying (hidden) states not directly observable.
- “Markov”: the state transitions are governed by a Markov process.

Discrete Markov Process

- Consider a system which may be described at any time as being in one of a set of N distinct states, $S_1, \ldots, S_N$.

- At time instances $t=1,2,3, \ldots$, the system changes its state according to certain probability. The full description requires us to know $P(s^t=S_j \mid s^{t-1}=S_i, s^{t-2}=S_k, \ldots, s^1=S_m)$ for all $t, i, k, \ldots, m$, where s^t stands for the state of the system at time t.
  - For a first-order Markov chain, we need to consider only

    $$P(s^t=S_j \mid s^{t-1}=S_i)$$

  - Further assume Ps are "stationary":

    $a_{ij} = P(s^t=S_j \mid s^{t-1}=S_i)$, $1 \le i,j \le N$, for any t.

A Simple Example

- Assume one of the three states for each day:

  S_1-rainy, S_2-cloudy, S_3-sunny

- Assume the **transition probability matrix**

$$A = \{a_{ij}\} = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}$$

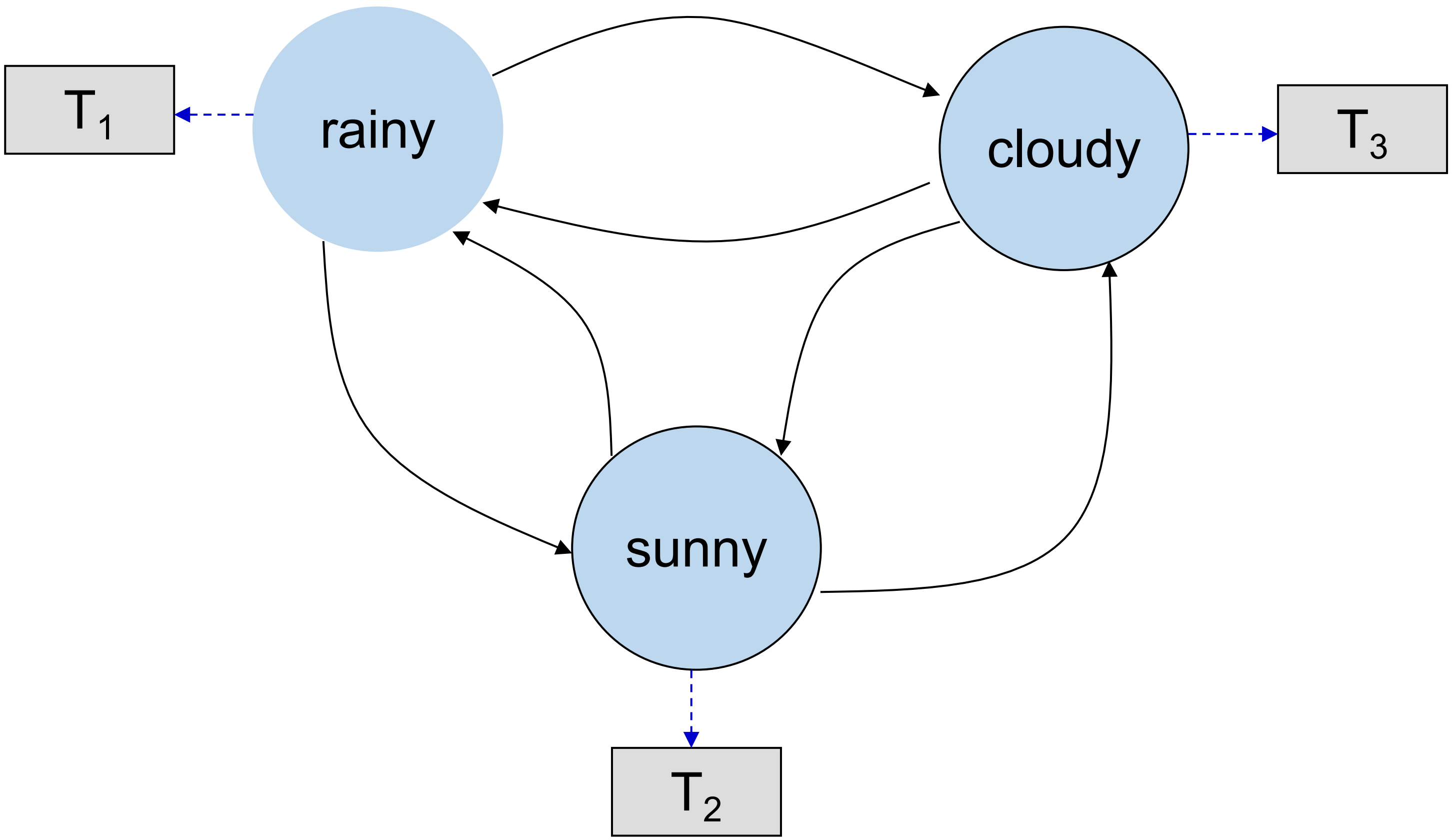
- Many questions we may ask, based on this model.
  - E.g., Given today is cloudy, what is the probability it remains to be cloudy for next 5 days?

Extending to “Hidden” States

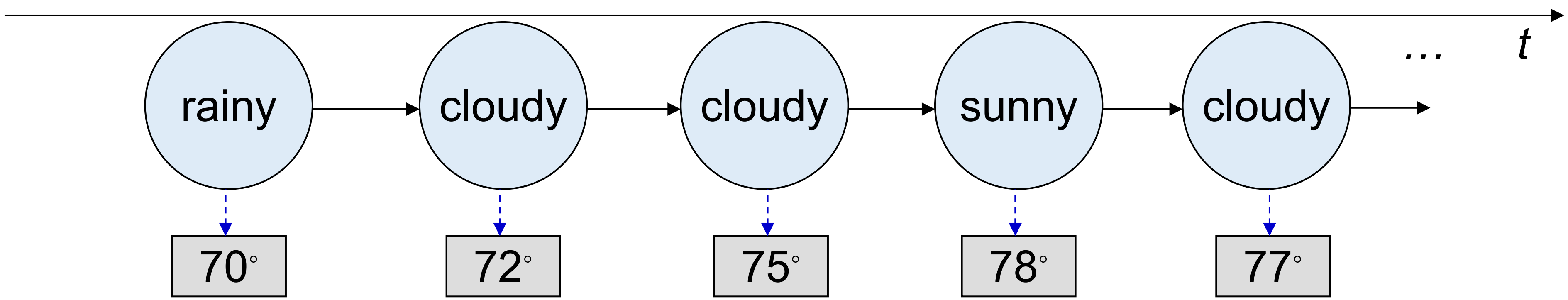
- The previous example is an “observable” Markov model: the output of the system/process is the states of interest.

- Now assume that we can only measure the (average) temperature of a day
  - Further assume this measurement is useful for predicting the weather states (rainy, cloudy, sunny).
  - We can view the temperature values as being produced by the *hidden states* of interest, i.e., the weather.

A Simple HMM



A Specific Process from the Model



Specifying an HMM

- Θ: the set of hidden states.

- The state transition probabilities $a_{ij} = P(s^t=S_j \mid s^{t-1}=S_i)$, $1 \le i,j \le N$
  - Let $A=\{a_{ij}\}$ be the transition probability matrix

- Ω: the set of outputs (observations).

Specifying an HMM (cont'd)

– The observation probabilities: $P(o^t|s^t)$, where o^t stands for the observation at time t, given the state s^t. This is also called the emission probability.

  - ➢ For discrete observation space, we can define $B=\{b_{jk}\}=P(o^t=\mathrm{v}_\mathrm{k} \text{ at } t|s^t=S_j)$ as the emission probability matrix, where v_k is the k^{th} symbol in Ω

– The initial state distribution $\pi = \{\pi_i\}$, $\pi_i=P(s^1=S_i)$

  - ➢ Sometimes we are given an initial state, i.e., $P(s^1=S_i)=1$ for certain i.

Basic Problems in HMM

For a given HMM $\Lambda=\{\Theta,\Omega,A,B,\pi\}$

- Problem 1: Given an observation (sequence) $\boldsymbol{O}=\{o^1,o^2, \ldots ,o^k\}$, what is the most likely state sequence $\boldsymbol{S}=\{s^1,s^2, \ldots ,s^k\}$ that has produced $\boldsymbol{O}$?

- Problem 2: How likely is an observation $\boldsymbol{O}$ (i.e., what is $P(\boldsymbol{O})$) ?

- Problem 3: How to estimate the model parameters (A,B,π)?