Markov Models

Dr Dimitris C. Dracopoulos
email: `d.dracopoulos@westminster.ac.uk`

Course web page:
`http://users.wmin.ac.uk/~dracopd/DOCUM/courses/2ait608/ait608.html`

Markov Models

Used to model dynamic systems.

- Speech/language processing
- Human behaviour (e.g. user modelling)
- Modelling psysical/biological processes
- Stock market
- ...

A class of probability models has to be defined, estimated from the observed behaviour of the dynamic system.

Markov Chains

A finite state Markov chain (stochastic finite state machine) can be defined:

- States: $s \in \{1, \ldots, m\}$, where m is finite.
- Starting state s_0: may be fixed or drawn from some a priori distribution $P_0(so)$.
- Transitions (dynamics): how the system moves from the current state s_t to the next state s_{t+1}.
- The transitions satisfy the first order Markov property:

$$P(s_{t+1}|s_t, s_{t-1}, \ldots, s_0) = P_1(s_{t+1}|s_t) \tag{1}$$

Markov Chains (cont'd)

Markov chains define a stochastic system generates a sequence of states:

$$s_0 \longrightarrow s_1 \longrightarrow s_2 \longrightarrow \ldots$$

where s_0 is drawn from $P_0(s_0)$ and each s_{t+1} from one step transition probabilities $P_1(s_{t+1}|s_t)$.

- A Markov chain can be represented as a state transition diagram.

Transition Probabilities

The conditional probability p_{ij} is defined as the probability that a system which occupies state i, will occupy state j after its next transition.

- Since the system must be in some state after its next transition:

$$\sum_{j=1}^{N} p_{ij} = 1 \tag{2}$$

- Since p_{ij} are probabilities:

$$0 \leq p_{ij} \leq 1 \tag{3}$$

Example - The Toymaker

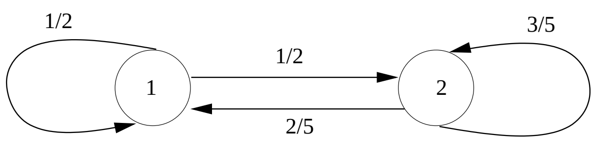
A toymaker is involved in the novelty toy business. He may be in either of two states:

1. The toy he is currently producing has found great favour with the public.
2. The toy is out of favour.

Transition probabilities:

- If in first state 50% chance of remaining in state 1, and 50% chance of unfortunate move to state 2 at following week.
- While in state 2, he experiments with new toys and he may return to state 1 after a week with probability $\frac{2}{5}$, or remain unprofitable in state 2 with probability $\frac{3}{5}$.

$$P = [p_{ij}] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{3}{5} \end{bmatrix} \tag{4}$$



Markov Chain Problems

Two interesting problems to solve:

- *Prediction*: Probabilities that the system will be in state s_k after n transitions, given that at $n = 0$ is it in a known state
- *Estimation*: Calculation of transition probabilities given some observed sequences of state transitions.

The Prediction Problem

Example: What is the probability that the toymaker will be in state 1 after n weeks, given that he is in state 1 at the beginning of the n-week period?

Define $\pi_i(n)$ as the probability that the system will occupy state i after n transitions, if its state at $n = 0$ is known.

Then:

$$\sum_{i=1}^{N} \pi_i(n) = 1 \tag{5}$$

$$\pi_j(n+1) = \sum_{i=1}^{N} \pi_i(n) p_{ij} \qquad n = 0, 1, 2, \ldots \tag{6}$$

Define a row vector of state probabilities $\pi(n)$ with components $\pi_i(n)$.

Then:

$$\pi(n+1) = \pi(n)P \qquad n = 0, 1, 2, \ldots \tag{7}$$

Now:

$$\begin{aligned} \pi(1) &= \pi(0)P \\ \pi(2) &= \pi(1)P = \pi(0)P^2 \\ \pi(3) &= \pi(2)P = \pi(0)P^3 \end{aligned} \tag{8}$$

In general:

$$\pi(n) = \pi(0)P^n \qquad n = 0, 1, 2, \ldots \tag{9}$$

Application to the Toymaker Example

Assume that the toymaker starts with a successful toy, then $\pi_1(0) = 1$, $\pi_2(0) = 0$.

$$\pi(1) = \pi(0)P = [10] \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \frac{1}{2} \end{bmatrix} \tag{10}$$

After 1 week tha toymaker is equally likely to be successful or unsuccessful.

After 2 weeks:

$$\pi(2) = \pi(1)P = [\frac{1}{2}\frac{1}{2}] \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} \frac{9}{20} \frac{11}{20} \end{bmatrix} \tag{11}$$

so that the toymaker is slightly more likely to be unsuccessful.

Example: Successive State Probabilities Starting with a Successful Toy

n 0	1	2	3	4	5	...
$\pi_1(n)$ 1	0.5	0.45	0.445	0.4445	0.44445	...
$\pi_2(n)$ 1	0.5	0.55	0.555	0.5555	0.55555	...

As n becomes very large:

- $\pi_1(n)$ approaches $\frac{4}{9}$
- $\pi_2(n)$ approaches $\frac{5}{9}$

Example: Successive State Probabilities Starting with a Unsuccessful Toy

$\pi_1(0) = 0$, $\pi_2(0) = 1$.

n	0	1	2	3	4	5	...
$\pi_1(n)$	1	0.4	0.44	0.444	0.4444	0.44444	...
$\pi_2(n)$	1	0.6	0.56	0.556	0.5556	0.55556	...

As n becomes very large:

- $\pi_1(n)$ approaches $\frac{4}{9}$
- $\pi_2(n)$ approaches $\frac{5}{9}$

Ergodic Processes

In the previous example of the toymaker, as n becomes large, $\pi_1(n)$ approaches $\frac{4}{9}$ and $\pi_2(n)$ approaches $\frac{5}{9}$ independent of the starting state.

A Markov process is completely *ergodic* if:

- The limiting state probability distribution is independent of starting conditions.

For completely ergodic Markov processes π with components π_i is $\pi(n)$ as n approaches infinity. Also called absolute state probabilities.

Absolute State Probabilities

For ergodic Markov processes from (9):

$$\pi = \pi P \tag{12}$$

We also have:

$$\sum_{i=1}^{N} \pi_i = 1 \tag{13}$$

From (12), (13) the limiting state probabilities for any process can be found.

For the toymaker example, equations (12), (13) yield:

$$\begin{aligned} \pi_1 &= \frac{1}{2}\pi_1 + \frac{2}{5}\pi_2 \\ \pi_2 &= \frac{1}{2}\pi_1 + \frac{3}{5}\pi_2 \\ \pi_1 &+ \pi_2 = 1 \end{aligned} \tag{14}$$

The unique solution of the above is: $\pi_1 = \frac{4}{9}$, $\pi_2 = \frac{5}{9}$.

- Finding the limiting state probabilities involves the solution of a set of N linear equations.

Markov Chain Estimation

- We need to estimate the initial distribution $P_0(s_0)$ and the transition probabilities $P_1(s'|s)$.
- Estimation from L observed sequences of different lengths:

$$
\begin{array}{ccc}
s_0^{(1)} & \longrightarrow & s_1^{(1)} \longrightarrow s_2^{(1)} \longrightarrow \ldots \longrightarrow s_{n_1}^{(1)} \\
\ldots & \ldots & \\
s_0^{(L)} & \longrightarrow & s_1^{(L} \longrightarrow s_2^{(L)} \longrightarrow \ldots \longrightarrow s_{n_L}^{(L)}
\end{array}
\tag{15}
$$

Maximum likelihood estimates (observed fractions):

$$\hat{P}(s_0 = i) = \frac{1}{L} \sum_{l=1}^{L} \delta(s_0^{(l)}, i) \tag{16}$$

where $\delta(x, y) = 1$, if $x = y$, and zero otherwise.

Markov Chain Estimation for Transition Probabilities

$$
\begin{array}{ccc}
s_0^{(1)} & \longrightarrow & s_1^{(1)} \longrightarrow s_2^{(1)} \longrightarrow \ldots \longrightarrow s_{n_1}^{(1)} \\
\ldots & \ldots & \\
s_0^{(L)} & \longrightarrow & s_1^{(L} \longrightarrow s_2^{(L)} \longrightarrow \ldots \longrightarrow s_{n_L}^{(L)}
\end{array}
\tag{17}
$$

- The transition probabilities are obtained as observed fractions of transitions out of a specific state.

Joint estimate over successive states:

$$\hat{P}_{s,s'}(s = i, s' = j) = \frac{1}{(\sum_{l=1}^{L} n_l)} \sum_{l=1}^{L} \sum_{t=0}^{n_l - 1} \delta(s_t^{(l)}, i)\delta(s_{t+1}^{(l)}, j) \tag{18}$$

and the transition probability estimates are:

$$\hat{P}_1(s' = j | s = i) = \frac{\hat{P}_{s,s'}(s = i, s' = j)}{\sum_k \hat{P}_{s,s'}(s = i, s' = k)} \tag{19}$$

Hidden Markov Models

- A hidden Markov model (HMM) is a model which we generate a sequence of outouts in addition to the Markov state sequence:

$$
\begin{array}{ccccccc}
s_0 & \longrightarrow & s_1 & \longrightarrow & s_2 & \longrightarrow & \ldots \\
\downarrow & & \downarrow & & \downarrow & & \\
x_0 & & x_1 & & x_2 & &
\end{array}
$$

- Only the outputs $\{x_0, x_1, \ldots, x_n\}$ are observed. The state sequence remains hidden.

Hidden Markov Problems to Solve

Two of the most interesting problems to solve, related to HMM are:

- Calculate the probability that our model geenrated the observation sequence $\{x_0, x_1, \ldots, x_n\}$?
  - forward-backward algorithm
- How do we find the most likely hidden state sequence corresponding to these observations?
  - dynamic programming