

BITS F464: Machine Learning

11

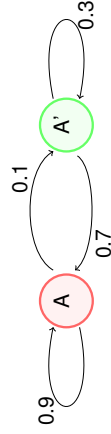
Hidden Markov Modal



Dr. Kamlesh Tiwari
 Assistant Professor, Department of CSIS,
 BITS Pilani, Pilani Campus, Rajasthan-333031 INDIA
 Feb 10, 2021 **ONLINE** (Campus @ BITS-Pilani Jan-May 2021)
<http://kti.wari.in/ml>

Markov Modal

Transition diagram



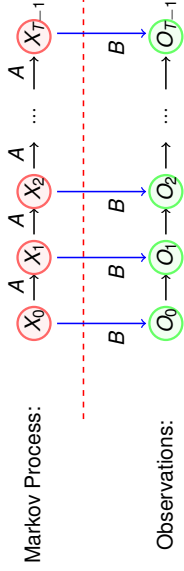
Initial State

- $S_0 = [0.2 \ 0.8]$
- $S_1 = [0.2 \ 0.8]$
- $S_1 = S_0 \times A = [0.74 \ 0.26]$
- $S_2 = S_1 \times A = [0.848 \ 0.152]$
- $S_3 = S_2 \times A = [0.8696 \ 0.1304]$

Transition matrix

$$A = \begin{bmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{bmatrix}$$

Hidden Markov Modal (HMM)



Assume we observe news coverage (S/M/L) of some article, to know whether a day was Hot or Cold?

$$B = \begin{matrix} S & M & L \\ H & \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.7 & 0.2 & 0.1 \end{bmatrix} \\ C & \end{matrix}$$

$$A = \begin{matrix} H & C \\ H & \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \\ C & \end{matrix}$$

Markov Modal

- Andrev Markov: A canonical probabilistic model for temporal or sequential data. $X_0 \xrightarrow{A} X_1 \xrightarrow{A} \dots \xrightarrow{A} X_n$
- Future is independent of past given the present. Assumption is that the present state encode all the history
- Order specifies how many evidences are important. Order three Markov Modal takes last three data
- iid¹ don't work.
- Temporal data, weather prediction, speech recognition, automatic music generation and handwriting recognition are some of the few applications

Example:

Suppose a company selling a product A (has market share of 20%), launches a advertise campaign that is expected to retain 90% old customers and attract 70% new. What maximum market share the product A can get?

¹independent and identically distributed

Is it going to saturate?

Stationary matrix

$$\begin{bmatrix} a & b \end{bmatrix} \times A = \begin{bmatrix} a & b \end{bmatrix}$$

$$\begin{bmatrix} a & b \end{bmatrix} \times \begin{bmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{bmatrix} = \begin{bmatrix} a & b \end{bmatrix}$$

what are a and b? 0.875 and 0.125

- Does it always happen? No, only if matrix is **regular**
- When some power of the matrix has all positive values
- Which of these are regular?

$$\begin{bmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0.2 & 0.8 \\ 1 & 0 \end{bmatrix}$$

Hidden Markov Modal (HMM)

$$B = \begin{matrix} S & M & L \\ H & \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.7 & 0.2 & 0.1 \end{bmatrix} \\ C & \end{matrix}$$

$$A = \begin{matrix} H & C \\ H & \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \\ C & \end{matrix}$$

- Assume initial configuration for H and C be $\pi = [0.6 \ 0.4]$
- And let observations be S, M, S, L
- Then what is $P(HHCC)$?
 $0.6 \times 0.1 \times (0.7 \times 0.4) \times (0.3 \times 0.7) \times (0.6 \times 0.1) = 0.000212$

Hidden Markov Modal (HMM)

Thank You!

State	Probability	Normalized Probability
HHHH	0.000412	0.042787
HHHC	0.000655	0.003655
HHCH	0.000655	0.003655
HHCC	0.000212	0.022017
HCHH	0.000650	0.005193
HGHC	0.000004	0.000415
HGCC	0.000302	0.031384
CHHH	0.000091	0.009451
CHHC	0.001089	0.114091
CHCH	0.001082	0.113426
CHCC	0.001882	0.195451
CCHH	0.000562	0.058573
CCHC	0.000470	0.048811
CCCH	0.000040	0.004154
CCCC	0.002822	0.289073
	0.006847	0.067593

Optimum state sequence

- In dynamic programming is CCCH
- HMM choses most probable symbol at each position. (by summation)

Thank you very much for your attention!

Queries ? Ref

	0	1	2	3
P(H)	0.188182	0.519576	0.228788	0.804029
P(C)	0.811818	0.480424	0.771212	0.195971

Optimum state sequence in HMM is ? CHCH