Lección 6: Modelos de cadenas de Markov oculto

Dr. Jesús Savage
Dr. Carlos Rivera

23 de abril de 2020

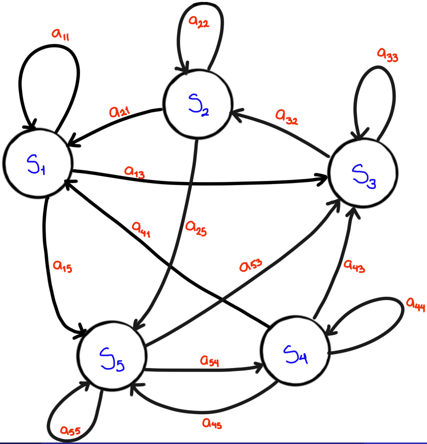
Índice

Modelos de cadenas de Markov

La teoría de cadenas de Markov ha sido extensivamente utilizada para resolver problemas de reconocimiento de patrones. La asumpción principal es que señales de una, dos y tres dimensiones pueden ser caracterizadas como un proceso aleatorio paramétrico, y que los parámetros del proceso estocástico pueden ser determinados (estimados) en una manera precisa y bien definida Se utilizan principalmente para el tipo de procesos que varian en el tiempo.

Procesos de Markov discretos en el tiempo

Una cadena de Markov está formado por un conjunto de estados unidos por ligas. La transición de un estado a otro está determinado por probabilidades.



Procesos de Markov discretos en el tiempo

Los cambios de estados se denominan q_t en el tiempo t. Probabilidad de que la cadena de Markov se encuentre en un estado determinado j

$$P[q_t = j \mid q_{t-1} = i, q_{t-2} = k, ...] = P[q_t = j \mid q_{t-1} = i]$$

$$a_{ij} = P[q_t = j \mid q_{t-1} = i], 1 \leq i\, j \leq N$$

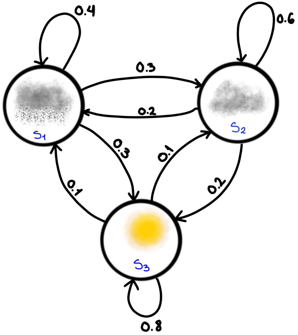
$$0 \leq a_{ij} \leq 1 \;\forall\; i, j$$

$$\sum_{j=1}^{N} a_{ij} = 1 \;\forall i$$

Este modelo es observable puesto que la salida es un conjunto de estados en cada instante donde cada estado representa un evento observable.

Ejemplo

Supóngase que tenemos un modelo del tiempo como se describe a continuación:



Donde: $S_1 = Lluvia$; $S_2 = Nublado$; $S_3 = Soleado$
El tiempo t es tomado a las 12:00 en el día.

Ejemplo

$$A = \{a_{ij}\} = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

Problema: ¿Cuál es la probabilidad que el clima en 8 días consecutivos sea: soleado, soleado, soleado, lluvioso, lluvioso, soleado, nublado, soleado? Definiendo la secuencia de observación $\underline{O}$, como

$$\underline{O} = (Soleado, Soleado, Soleado, Lluvioso,$$

$$Lluvioso, Soleado, Nublado, Soleado)$$

$$\underline{O} = (S_3, S_3, S_3, S_1, S_1, S_3, S_2, S_3)$$

$$P(\underline{O}|modelo) = Probabilidad\ de\ observar\ la\ secuencia\ \underline{O},$$

$$dado\ el\ modelo\ del\ clima$$

Ejemplo

$$= P(S_3, S_3, S_3, S_1, S_1, S_3, S_2, S_3|modelo)$$
$$= P(S_3)P\left[S_3|S_3\right]^2 P\left[S_1|S_3\right] P\left[S_1|S_1\right] P\left[S_3|S_1\right] P\left[S_2|S_3\right] P\left[S_3|S_2\right]$$
$$= \pi_3(a_{33})^2\ a_{31}\ a_{11}\ a_{13}\ a_{32}\ a_{23}$$
$$= (1.0)(0.8)^2(0.1)(0.4)(0.3)(0.1)(0.2) = 1.536x10^{-4}$$

Donde se usa la notación:

$$\pi_3 = P\left[q_1 = 3\right] = 1$$

π_i son las probabilidades iniciales, en este ejemplo cuando se inicio el experimento era ya un día soleado, por lo tanto $\pi_3 = 1$.

Problema

¿Cuál es la probabilidad que el sistema se quede en un estado específico i por d días?
Secuencia de observaciones:

$$\underline{O} = (i, i, i, ..., i, j \neq i)$$

$$dias = (1, 2, 3, ..., d, d + 1)$$

¿ $P(\underline{O}|modelo, q_1 = i)$?
Recordar la definición de probabilidades condicionales:

$$P(A|B) = \frac{P(A \cdot B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cdot B)}{P(A)}$$

Problema

$$P(\underline{O}|modelo, q_1 = i) = \frac{P(\underline{O}, modelo, q_1 = i)}{P(modelo, q_1 = i)} = \frac{P(\underline{O}, modelo, q_1 = i)}{P(modelo)P(q_1 = i)}$$

$$P(\underline{O}, q_1 = i|modelo) = \frac{P(\underline{O}, q_1 = i, modelo)}{P(modelo)}$$

$$P(\underline{O}, q_1 = i|modelo)P(modelo) = P(\underline{O}, q_1 = i, modelo)$$

Substituyendo esta expresión en la primera ecuación:

$$P(\underline{O}|modelo, q_1 = i) = \frac{P(\underline{O}, q_1 = i|modelo)P(model)}{P(modelo)P(q_1 = i)}$$

$$\Rightarrow P(\underline{O}|modelo) = \frac{P(\underline{O}, q_1 = i|modelo)}{P(q_1 = i)}$$

$$= \pi_i \; a_{ii}^{d-1} \cdot (1 - a_{ii})/\pi_i$$

Problema

$$= a_{ii}^{d-1}(1 - a_{ii}) = P_i(d)$$

$P_i(d)$ es la función de probabilidad de estar en el estado i d veces continuas.

$$P_i(d) = P\left[q_t = j|q_{t-1} = i\right] ... P\left[q_1 = i|q_0 = i\right]$$

Se puede calcular el número esperado de durar en el estado i

$$\bar{d}_i = \sum_{d=1}^{\infty} d \ p_i(d)$$

Problema

$$= \sum_{d=1}^{\infty} d a_{ii}^{d-1}(1 - a_{ii})$$

$$= (1 - a_{ii}) \sum_{d=1}^{\infty} d a_{ii}^{d-1} = (1 - a_{ii}) \frac{\delta}{\delta a_{ii}} \sum_{d=1}^{\infty} a_{ii}^{d}$$

Por otra parte:

$$\sum_{d=1}^{\infty} a_{ii}^{d} = \sum_{d=1}^{\infty} a_{ii}^{d} + 1 - 1 = \sum_{d=0}^{\infty} a_{ii}^{d} - 1 = \frac{1}{1 - a_{ii}} - 1$$

$$= \frac{1 - 1 + a_{ii}}{1 - a_{ii}} = \frac{a_{ii}}{1 - a_{ii}}$$

$$= (1 - a_{ii}) \frac{\delta}{\delta a_{ii}} \left(\frac{a_{ii}}{1 - a_{ii}} \right)$$

Problema

$$\frac{\delta}{\delta a_{ii}} a_{ii}(1-a_{ii})^{-1} = (1-a_{ii})^{-1} - a_{ii}(-1)(1-a_{ii})^{-2}$$

$$= \frac{1}{1-a_{ii}} + \frac{a_{ii}}{(1-a_{ii})^2} = \frac{1-a_{ii}+a_{ii}}{(1-a_{ii})^2} = \frac{1}{(1-a_{ii})^2}$$

Finalmente

$$\bar{d}_i = (1-a_{ii}) \cdot \frac{1}{(1-a_{ii})^2} = \frac{1}{(1-a_{ii})}$$

Modelos de Markov Ocultos

Ejemplos de repaso:

Dada una moneda normal $P(sol) = P(aguila) = 0.5$

1 ¿Cuál es la probabilidad que los próximos diez volados den la siguiente secuencia, con S = sol y A = aguila?

   $(S, S, A, S, A, A, S, A, A, S)$

   Para una moneda normal haciendo 10 volados independientes hay 2^{10} secuencias probables, y como cada una tiene la misma probabilidad de ocurrir, entonces:

$$P(SSASAASAAS) = \left(\frac{1}{2}\right)^{10}$$

Modelos de Markov

2 ¿Cuál es la probabilidad de que los próximos 10 volados generen una secuencia ($SSSSSSSSSS$)?

$$P(SSSSSSSSSS) = \left(\frac{1}{2}\right)^{10}$$

3 ¿Cuál es la probabilidad de que 5 de los 10 próximos volados sean soles? Esta probabilidad es igual al número de secuencias observadas con 5 soles y 5 aguilas (en cualquier orden).

$$P(5S, 5A) = \binom{10}{5}\left(\frac{1}{2}\right)^{10} = \frac{252}{1024} \approx 0.25$$

Ya que hay $\binom{10}{5}$ formas de tener 5 soles y 5 águilas en 10 volados, y cada secuencia tiene una probabilidad de $\left(\frac{1}{2}\right)^{10}$

Modelos de Markov Ocultos

4 ¿Cuál es el número esperado de soles en 10 volados?

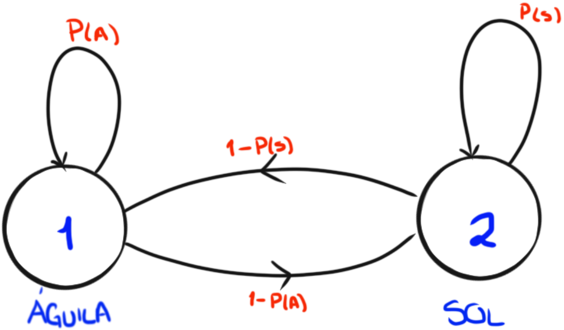
$$E(S\ en\ 10\ volados) = \sum_{d=0}^{10} d \binom{10}{d} \left(\frac{1}{2}\right)^{10} = 5$$

Ejemplo

Supongamos que una persona está echando volados detrás de una cortina, con diferentes monedas cargadas.
La persona solamente dice el resultado obtenido sin decir de qué moneda usó. ¿Cómo se construiría un modelo de cadenas de Markov Ocultos tal que explicara la recurrencia de águilas y soles obtenidos?

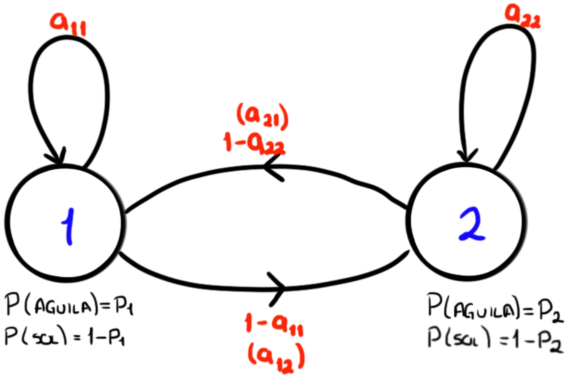
$$\underline{O} = (O_1, O_2, O_3, ..., O_T)$$
$$= (AASSA...S)$$

Modelo con una sola moneda



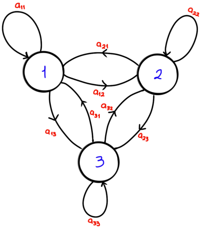
Observaciones = AASSASASS
Estados = 112212122
Los estados correspoden a las observaciones.

Modelo con dos monedas



Observaciones = AASSASASS
Estados = 211122211

Modelo con tres monedas



$$Observaciones = \underline{O} = SSAASASAASSA$$

$$Estados = \underline{q} = 312331111313$$

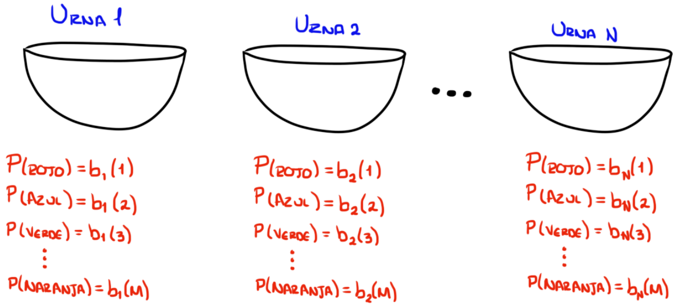
$$P_1(A) = P_1; \quad P_2(A) = P_2; \quad P_3(A) = P_3$$

$$P_1(S) = 1 - P_1; \quad P_2(S) = 1 - P_2; \quad P_3(S) = 1 - P_3$$

Las $a_i s$ representan las probabilidades de cambiar las monedas cuando se hacen los volados

Modelo de las urnas con pelotas

Se tienen N urnas con M bolas de diferentes colores. Se tiene una persona que aleatoriamente escoge una pelota de alguna de las urnas y la muestra, repitiendo esta operación varias veces. El proceso completo corresponde a una salida observable de un HMM.



Modelo de las urnas con pelotas

$$\underline{O} = \{Verde, Verde, Azul, Rojo, Amarillo, ..., Azul\}$$

Elementos de un HMM para observaciones de símbolos discretos.

1 N número de estados del modelo.
A pesar que los estados son ocultos, para muchas aplicaciones prácticas hay evidancia física para cada uno de los estados. Cada estado es denominado como $\{1, 2, ..., N\}$, y el tiempo en el estado como q_t

2 M, el número de observaciones distintas en cada estado, es decir, el alfabeto de tamaño discreto.
Las observaciones corresponden a las salidas físicas del sistema que está siendo modelado. Se representan los símbolos individuales como $V = \{V1, V_2, ..., V_M\}$

Modelo de las urnas con pelotas

3 La distribución de probabilidad es de cambio de estado $A = \{a_{ij}\}$ donde

$$a_{ij} = P\{q_{t+1} = j | q_t = i\}, 1 \leq i, j \leq N;$$

$$0 \leq a_{ij} \leq 1; \quad \forall i, j$$

4 La distribución de probabilidad del símbolo observado, $B = \{b_j(k)\}$

$$b_j(k) = P[O_t = V_k | q_t = j],\ 1 \leq k \leq M$$

5 La distribución del estado inicial $\pi = \pi_j$

$$\pi_i = P[q_1 = i],\ 1 \leq i \leq N$$

Modelos de Markov Ocultos

Entonces el modelo de Markov oculto (en ingles Hidden Markov Model, HMM), completo requiere los valores de N, M, los símbolos de salida V y una especificación de las probabilidades A, B y π. Por conveniencia se usa la siguiente notación:

$$\lambda = (A, B, \pi)$$