Ejercicios 7: Modelos de markov

Departamento de Informática / Department of Computer Science
Universidad Carlos III de Madrid

Inteligencia Artificial
Grado en Ingeniería Informática
2019/20

Ejercicio 1: Evaluación de alumnos

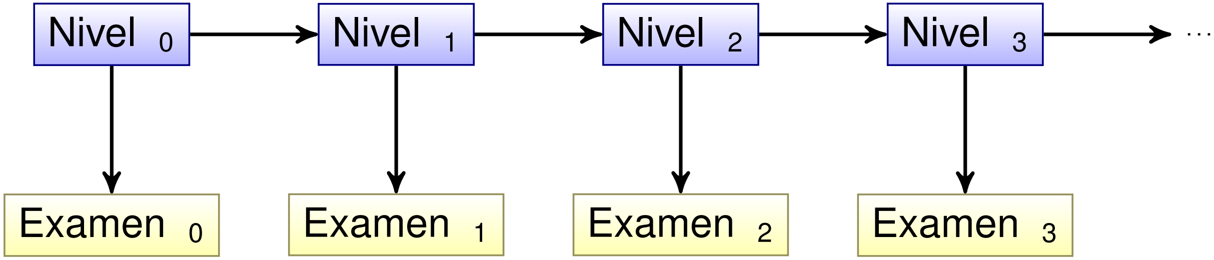
- ¿Cómo podría aplicar HMMs (Modelos Ocultos de Markov) para decidir si un alumno está entendiendo o no una asignatura? Describa todos los componentes del HMM correspondiente, asignando razonadamente probabilidades numéricas.

Evaluación de alumnos, representación

- El HMM debe tener variables de estado no observables (ocultas), y variables que representen las observaciones.
- El estado oculto será el grado de comprensión de la asignatura (puede ser una variable con distintos valores discretos).
- Para que el uso de HMM esté justificado, debe existir una relación temporal entre el nivel de comprensión en los momentos sucesivos. Todo indica que podemos considerarlo así, ya que no sería habitual que el alumno pasara en un solo paso de entender completamente la asignatura a no entender nada, pero sí puede variar en sentido positivo o negativo. Es decir, ambas variables **no** son independientes. Tenemos que hacer la suposición (fundada) de que el grado de comprensión (estado) en un momento dado depende exclusivamente del grado de comprensión en el momento anterior (hipótesis de Markov)
- Las observaciones podrían ser las notas de los sucesivos exámenes de la asignatura.

Evaluación de alumnos, representación

El HMM sería el que muestra la figura. Los momentos del tiempo que se representan corresponden con la realización de exámenes.



Las distribuciones de probabilidad que caracterizan este HMM son:

- P($Nivel_0$): probabilidad a priori del estado inicial
- P($Nivel_i|Nivel_{i-1}$): modelo de transiciones entre estados
- P($Examen_i|Nivel_i$): modelo de observaciones

Evaluación de alumnos, representación

Hay que fijar los dominios de las variables. Por simplificar las tablas podríamos definir $Nivel \in \{I, S, N\}$ (**I** es nivel insuficiente, **S** suficiente y **N** notable), y $Examen \in \{T, F\}$ (**T** es que el examen se aprueba). Suponiendo un modelo estacionario, hay que definir $P(Nivel_0)$ y las dos tablas de probabilidad condicional:

- $P(Nivel_0) = \{0.75(I), 0.25(S), 0(N)\}$, donde entendemos que es más probable tener un nivel bajo al principio.
- $P(Nivel_i|Nivel_{i-1})$ para $i > 0$. En un caso real habría que fundar estos datos en la experiencia. Para este ejemplo vamos a asignarlos simplemente teniendo en cuenta que hay mayor probabilidad de mejorar que de empeorar.
- $P(Examen_i|Nivel_i)$, aquí vamos a entender que si se tiene nivel N se aprueba casi siempre, mientras que con los otros niveles hay mayor incertidumbre.

	$Nivel_i$		
$Nivel_{i-1}$	I	S	N
I	0.50	0.30	0.20
S	0.20	0.50	0.30
N	0.0	0.20	0.80

	$Examen_i$	
$Nivel_i$	T	F
I	0.30	0.70
S	0.70	0.30
N	0.90	0.10

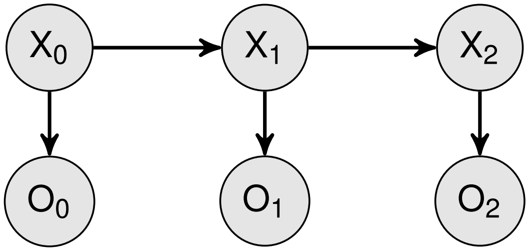
Evaluación de alumnos, representación

La tarea de determinar el nivel del alumno en un momento dado, a partir de las notas de los exámenes realizados, correspondería con la resolución de un problema de filtrado. Habría que calcular:

$$P(Nivel_i | Examen_0, Examen_1, \cdots, Examen_i)$$

Ejercicio 2: HMM

- A veces uno tiene un resfriado (cold) que le hace estornudar (sneeze). También puede tener alergia (allergy) y esto también le puede hacer estornudar. Otras veces uno se siente bien (well) y no estornuda. Se decide modelar esto como un HMM donde:
  - Los estados X_t representan el estado de la persona: $X_t \in \{well(w), allergy(a), cold(c)\}$
  - Las observaciones O_t representan los síntomas observables (si estornuda o no): $O_t \in \{s, \neg s\}$



  - Asumiremos que los instantes de tiempo representan días y que inicialmente la persona está bien

Ejercicio 2: HMM

Las distribuciones de probabilidad que caracterizan el HMM son las siguientes:

- $P(X_0) = (1,0,0)$ (es decir, $P(X_0 = w) = 1$)
- $P(X_{t+1}/X_t)$

X_t	$P(X_{t+1} = w/X_t)$	$P(X_{t+1} = a/X_t)$	$P(X_{t+1} = c/X_t)$
w	0.7	0.2	0.1
a	0.6	0.3	0.1
c	0.2	0.2	0.6

- $P(O_t/X_t)$

X_t	$P(O_t = s/X_t)$	$P(O_t = \neg s/X_t)$
w	0.1	0.9
a	0.8	0.2
c	0.7	0.3

1. ¿Cuál es la probabilidad de que uno esté bien (w) mañana?
2. ¿Cuál es la probabilidad de la secuencia w,c,c,w en los primeros 4 días?
3. ¿Cuál es la probabilidad de que hoy (día 0) no se observen estornudos ($O_0 = \neg$s) y mañana se observen estornudos ($O_1 = s$)?

HMM: Solución

1. ¿Cuál es la probabilidad de que uno esté bien (w) mañana?

$$P(X_1 = w) = P(X_1 = w|X_0 = w)P(X_0 = w) = (0.7)1 = 0.7$$

2. ¿Cuál es la probabilidad de la secuencia w,c,c,w en los primeros 4 días?

$$\begin{aligned} P(X_0 = w, X_1 = c, X_2 = c, X_3 = w) &= P(X_3 = w|X_2 = c)P(X_2 = c|X_1 = c)P(X_1 = c|X_0 = w)P(X_0 = w) \\ &= (0.2)(0.6)(0.1)(1) = 0.012 \end{aligned}$$

3. ¿Cuál es la probabilidad de que hoy (día 0) no se observen estornudos ($O_0 = \neg$s) y mañana se observen estornudos ($O_1 = s$)?

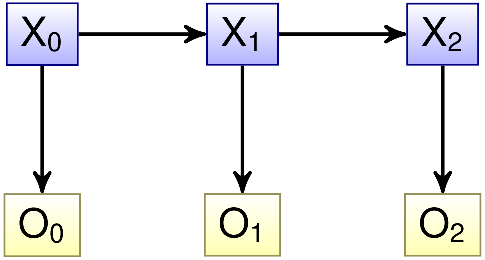
$$\begin{aligned} P(O_0 = \neg s, O_1 = s) &= \sum_{x_0, x_1} P(O_1 = s|X_1 = x1)P(X_1 = x1|X_0 = x_0)P(O_0 = \neg s|X_0 = x_0)P(X_0 = x_0) \\ &= P(O_0 = \neg s|X_0 = w)P(O_1 = s|X_1 = w)P(X_1 = w|X_0 = w)P(X_0 = w) + \\ &= P(O_0 = \neg s|X_0 = w)P(O_1 = s|X_1 = c)P(X_1 = c|X_0 = w)P(X_0 = w) + \\ &= P(O_0 = \neg s|X_0 = w)P(O_1 = s|X_1 = a)P(X_1 = a|X_0 = w)P(X_0 = w) + \\ &= (0.9)(0.1)(0.7)(1) + (0.9)(0.8)(0.2)(1) + (0.9)(0.7)(0.1)(1) \\ &= 0.27 \end{aligned}$$

Ejercicio 3: Reconocimiento de voz

Supongamos que se quiere construir un sistema para reconocimiento automático del voz. En particular, se quieren reconocer palabras dada una señal acústica. Para simplificar, sólo consideraremos palabras de tres letras que contengan las letras (*a*, *r* y *t*).

El proceso se puede modelar con un HMM en el que los estados (ocultos) se definen con una variable que representa la letra, y las observaciones representan la señal acústica, en forma de un fonema extraído de la señal.

- Los estados X_t representan la letra $\{a, r, t\}$
- Las observaciones O_t representan el fonema $\{a, r, t\}$



Ejercicio 3: Reconocimiento de voz

Se conocen las siguientes probabilidades a priori y CPTs

- $P(X_0) = (0.3, 0.3, 0.4)$
- $P(X_{t+1}/X_t)$

X_t	$P(X_{t+1} = a/X_t)$	$P(X_{t+1} = r/X_t)$	$P(X_{t+1} = t/X_t)$
a	0.0	0.7	0.3
r	0.5	0	0.5
t	1.0	0.0	0.0

- $P(O_t/X_t)$

X_t	$P(O_t = a/X_t)$	$P(O_t = r/X_t)$	$P(O_t = t/X_t)$
a	0.9	0.1	0.0
r	0.1	0.8	0.1
t	0.0	0.2	0.8

Ejercicio 3: Reconocimiento de voz

1. ¿Cuál es la probabilidad de que la letra inicial sea *a* cuando el sistema escucha el fonema *a*? ($P(X_0 = a/O_0 = a)$)
2. ¿Cuál es la probabilidad de que el primer fonema sea *a* ($P(O_0 = a)$)?
3. Si se observa la secuencia de fonemas *arr*, ¿Cuál es la palabra más probable?

1. $P(X_0 = a|O_0 = a)$
   - $P(X_0 = a|O_0 = a) = \alpha P(O_0 = a|X_0 = a)P(X_0 = a) = \alpha 0.9 \times 0.3 = \alpha 0.27 = 0.9$
     Para calcular α se puede calcular la misma probabilidad para el resto de valores de X_0:
   - $P(X_0 = r|O_0 = a) = \alpha P(O_0 = a|X_0 = r)P(X_0 = r) = \alpha 0.1 \times 0.3 = \alpha 0.03 = 0.1$
   - $P(X_0 = t|O_0 = a) = \alpha P(O_0 = a|X_0 = t)P(X_0 = t) = \alpha 0.0 \times 0.4 = 0$
     Por tanto, $\alpha = 1/(0.27 + 0.03) = 1/0.3$
2. Equivale a $1|\alpha$ del ejercicio anterior. También se puede calcular como:

$$P(O_0 = a) = \sum_i P(O_0 = a|X_0 = i)P(X_0 = i) = P(O_0 = a|X_0 = a)P(X_0 = a) +$$
$$P(O_0 = a|X_0 = r)P(X_0 = r) + P(O_0 = a|X_0 = t)P(X_0 = t) =$$
$$0.9 \times 0.3 + 0.1 \times 0.3 + 0.0 \times 0.4 = 0.3$$

3. $P(X_0, X_1, X_2|O_0 = a, O_1 = r, O_2 = r)$?

$$P(X_0, X_1, X_2|O_0 = a, O_1 = r, O_2 = r) = \alpha P(X_0, X_1, X_2, O_0 = a, O_1 = r, O_2 = r) =$$
$$= P(O_0 = a|X_0)P(X_0)P(O_1 = r|X_1)P(X_1|X_0)P(O_2 = r|X_2)P(X_2|X_1)$$

Hay que calcular esta probabilidad para todos los casos de X_0, X_1 y X_2 y tomar el mas probable. Por ejemplo, para $X_0 = a, X_1 = r, X_2 = t$:

$$P(X_0 = a, X_1 = r, X_2 = t|O_0 = a, O_1 = r, O_2 = r) = \alpha P(X_0 = a, X_1 = r, X_2 = t, O_0 = a, O_1 = r, O_2 = r) =$$
$$= P(O_0 = a|X_0 = a)P(X_0 = a)P(O_1 = r|X_1 = r)P(X_1 = r|X_0 = a)P(O_2 = r|X_2 = t)P(X_2 = t|X_1 = r) =$$
$$= \alpha 0.9 \times 0.3 \times 0.8 \times 0.7 \times 0.2 \times 0.5$$

Ejercicio 4: Seguimiento aviación

- Se tiene un radar que cada pocos segundos obtiene la posición y velocidad de un avión de combate F-16 Falcon. Denominamos estas observaciones O_t, para el instante de tiempo t. Se quiere construir un sistema para determinar si ese F-16 supone una amenaza inminente o no. Es decir para determinar si nos va a atacar o no.
- ¿Cuál de los modelos vistos en clase es el más adecuado para construir el sistema? ¿Cómo se realizaría el razonamiento para obtener la respuesta?

Seguimiento aviación: Solución

- El problema se puede modelar como un Hidden Markov Model (HMM).
- Observaciones (O_t): representan la información del radar
- Estados ocultos (X_t): representan si el F-16 nos va a atacar o no en el instante t.
- La distribución $P(X_0)$ representa la probabilidad inicial de que un F-16 nos ataque.
- La distribución $P(X_{t+1}|X_t)$ representa la evolución de la amenaza, es decir, la probabilidad de tener un ataque en el instante t dependiendo del estado anterior.
- La distribución $P(O_t|X_t)$ representa el modelo de observaciones que determina la probabilidad de que el F-16 tenga una posición y velocidad dependiendo de si hay amenaza o no.
  La tarea de inferencia a resolver es filtrado: $P(X_t = t|O_0, \ldots, O_t)$

Ejercicio 5: Tenis

Se dispone de los datos de la figura sobre la conveniencia de jugar o no al tenis, en función de datos meteorológicos.

- Construya un clasificador Naive Bayes que nos aconseje sobre hacer o no la actividad, usando datos de los primeros 10 días para predecir los últimos 4.
- Construya Modelos de Markov que modelen de forma separada la evolución temporal de cada una las cuatro variables meteorológicas. ¿Qué pasaría si quisiéramos tener en cuenta todas ellas en un solo modelo?

Day	Clouds	Temp.	Hum.	Wind	Plays?
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Tenis: sin solución

La solución a este ejercicio se deja para el alumno.