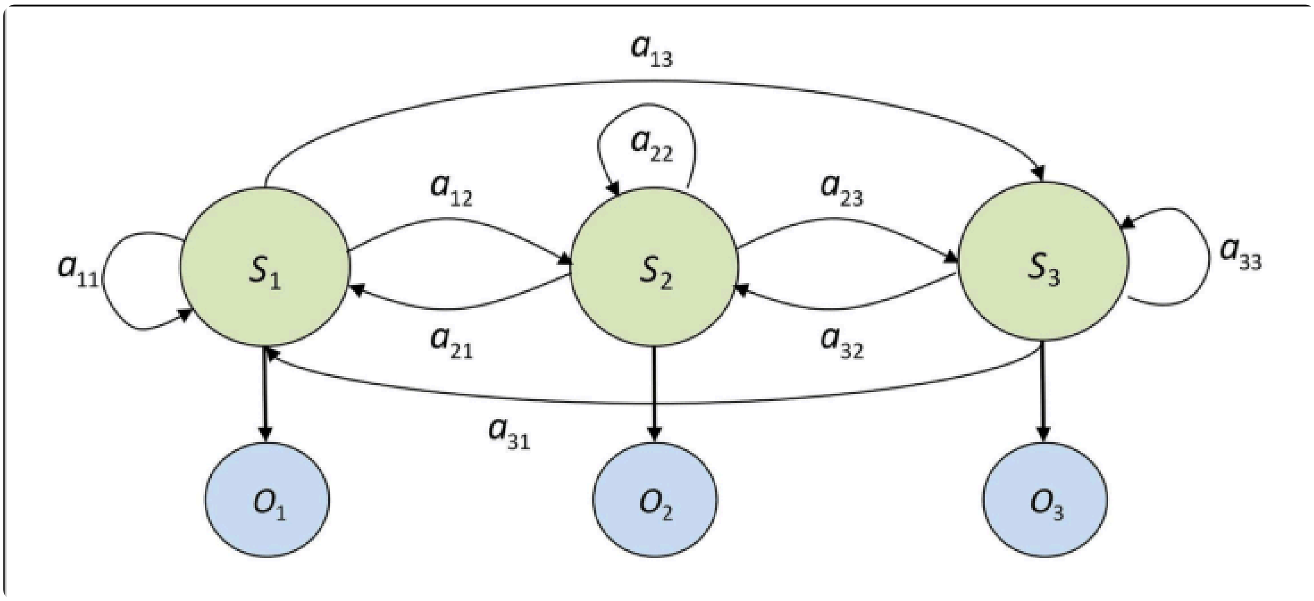
Modelos ocultos de Markov

Tags: Reconocimiento de Patrones

Modelos ocultos de Márkov

Un modelo oculto de Márkov (HMM por sus siglas en inglés), es un autómata de estados finitos capaz de producir a su salida, una secuencia de símbolos observables. El autómata está formado por un conjunto de estados y evoluciona pasando de un estado a otro de forma. Los estados están conectados unos a otros por arcos de transición con probabilidades asociadas a cada arco.

Cada estado tiene asociada una función de densidad de probabilidad, que define la probabilidad que tiene cada símbolo de ser emitido cada vez que se produce una transición desde dicho estado de HMM. Por lo tanto, un HMM consta de dos procesos estocásticos, la producción de símbolos y la secuencia de los estados en la evolución del mismo. De ellos, solo la producción de símbolos es observable.



Los elementos que constituyen un HMM son cinco:

- Un conjunto de N estados $S = \{s_1, s_2, \ldots, s_n\}$
- Los estados deben de estar conectados entre sí, de forma que cualquiera de ellos pueda ser alcanzable desde al menos un estado.
- Un conjunto de M símbolos observables que pueden ser producidos por el HMM. $O = \{o_1, o_2, \ldots, 0_M\}$

- Una matriz de probabilidades de transición de estados, $A = \{a_{ij}\}$. Esta matriz es cuadrada de dimensión N y cada elemento a_{ij} corresponde a la probabilidad de transición del estado s_i al s_j.

$$a_{ij} = p(q_t = s_j | q_{t-1} = s_i) \qquad 1 \leq i, j, \leq N$$

donde q_t indica el estado en el que se encuentra el modelo en el instante t. Debido a su naturaleza probabilística, cada elemento a_{ij} debe cumplir:

$$0 \leq a_{ij} \leq 1$$

$$1 \leq i, j \leq N$$

Por otra parte, también debe cumplirse que las probabilidades con origen en el mismo estado deben estar normalizadas

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

- Un conjunto de parámetros $B = \{b_i(k), 1 \leq i \leq N, 1 \leq k \leq M\}$ que definen, para cada estado la función de densidad de probabilidad de producciones, si las observaciones son magnitudes continuas; o las distribuciones de probabilidad si las observaciones son discretas. Cada b_i se define de la siguiente forma:

  $$b_i(o) = p(x_t = o | q_t = s_i) \qquad 1 \leq i \leq N$$

  donde x_t representa el valor de la observación en el instante de tiempo t
- Un conjunto de probabilidades de estado inicial $\Pi = \{\pi_i\}$, siendo π_i la probabilidad de que el estado inicial de modelo sea s_i

  $$\pi_i = p(q_i = s_i) \qquad 1 \leq i \leq N$$

Las probabilidades de estado inicial π_i deben verificar

$$0 \leq \pi_i \leq 1, \qquad 1 \leq i \leq N$$

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

De esta forma un HMM queda definido mediante

$$\lambda = \{\Pi, A, B\}$$

En el caso general, la variable estocástica asociada a la producción de observaciones es continua y multivariada. Sea X_1^T una secuencia observable, donde el subíndice 1 denota la observación en el instante de tiempo $t = T$, así, la secuencia está compuesta por T vectores continuos observados, de forma que $X_1^T = x_1, x_2, \ldots, x_T$

La utilización de los HMM dentro de un sistema de reconocimiento de patrones requiere la resolución de 3 problemas:

- Evaluación: Dada una secuencia de observaciones $X_1^T = x_1, x_2, \ldots, x_T$ y un modelo λ, se busca cómo evaluar la probabilidad $p(x_1^T|\lambda)$ de que la secuencia observada sea producida por dicho modelo.
- Estimación: Dada una secuencia de observaciones $X_1^T = x_1, x_2, \ldots, x_T$ y un modelo λ, cómo elegir los parámetros del modelo $\lambda = (\Pi, A, B)$ para que la probabilidad de generación de dicha secuencia por el modelo sea óptima.
- Decodificación: Dada una secuencia de observaciones $X_1^T = x_1, x_2, \ldots, x_T$ cómo obtener la secuencia de estados $Q_1^T = q_1, q_2, \ldots, q_T$ que mejor explica la generación de la secuencia por parte del modelo.

References

- Independencia de función de densidad de probabilidad
- Stanford University - Hidden Markov Model
- a tutorial on hidden markov models and selected applications in speech recognition, Lawrence Rabiner