Cadenas de Markov

Una cadena de Markov es un proceso estocástico de tiempo discreto: un proceso que ocurre en una serie de pasos de tiempo, en cada uno de los cuales se hace una elección aleatoria. Una cadena de Markov consta de N estados.
Una cadena de Markov se caracteriza por una matriz de probabilidad de transición $N \times N$, P cada una de cuyas entradas está en el intervalo $[0, 1]$; las entradas en cada fila de P suman 1. La cadena de Markov puede estar en uno de los estados N en cualquier paso de tiempo dado; luego, la entrada P_{ij} nos dice la probabilidad de que el estado en el siguiente paso de tiempo sea j, condicionado a que el estado actual sea i. Cada entrada P_{ij} se conoce como probabilidad de transición y depende solo del estado actual i; esto se conoce como propiedad de Markov. Así, la propiedad de Markov puede definirse de la siguiente manera:
Ecuación 1 :

$$\forall i, j, P_{ij} \in [0, 1]$$

Y ecuación 2:

$$\forall i, \sum_{j=1}^{N} P_{ij} = 1.$$

Una matriz con entradas no negativas que satisface la Ecuación 2 se conoce como matriz estocástica. Una propiedad clave de una matriz estocástica es que tiene un vector propio izquierdo principal correspondiente a su valor propio más grande, que es 1.
En una cadena de Markov, la distribución de probabilidad de los siguientes estados para una cadena de Markov depende solo del estado actual y no de cómo llegó la cadena de Markov al estado actual. La siguiente figura muestra una cadena de Markov simple con tres estados. Desde el estado medio A, procedemos con probabilidades (iguales) de 0.5 a B o C. Desde B o C, procedemos con probabilidad 1 a A. La matriz de probabilidad de transición de esta cadena de Markov es entonces:

$$\begin{pmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

La distribución de probabilidad de una cadena de Markov sobre sus estados

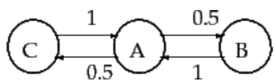

Figure 1: Una cadena de Markov simple con tres estados; los números en los enlaces indican las probabilidades de transición.

puede verse como un vector de probabilidad: un vector en cuyas entradas están en el intervalo $[0, 1]$, y las entradas suman 1. Un vector de probabilidad N -dimensional, cuenta con componentes correspondientes a uno de los estados N de una cadena de Markov, puede verse como una distribución de probabilidad sobre sus estados. Para nuestra cadena de Markov simple de la figura 1, el vector de probabilidad tendría 3 componentes que suman 1.
Podemos ver a un internauta aleatorio en el gráfico web como una cadena de Markov, con un estado para cada página web, y cada probabilidad de transición

representa la probabilidad de pasar de una página web a otra. La operación de movilizarse contribuye a estas probabilidades de transición. La matriz de adyacencia A del gráfico web se define de la siguiente manera: si hay un hipervínculo desde la página i a la página j, entonces $A_{ij} = 1$, de lo contrario $A_{ij} = 0$. Podemos derivar fácilmente la matriz de probabilidad de transición P para nuestra cadena de Markov a partir de la matriz $N \times N$ A:

1. Si una fila de A no tiene 1, reemplace cada elemento por 1/N. Para todas las demás filas, proceda de la siguiente manera.
2. Divida cada 1 en A por el número de 1 en su fila. Por lo tanto, si hay una fila con tres unos, cada uno de ellos se reemplaza por 1/3.
3. Multiplica la matriz resultante por $1 - \alpha$.
4. Agregue α/N a cada entrada de la matriz resultante, para obtener P.

Podemos representar la distribución de probabilidad de la posición del internauta en cualquier momento mediante un vector de probabilidad $\vec{x}$. En $t = 0$, el internauta puede comenzar en un estado cuya entrada correspondiente en $\vec{x}$ es 1 mientras que todos los demás son cero. Por definición, la distribución del internauta en $t = 1$ viene dada por el vector de probabilidad $\vec{x}P$; en $t = 2$ por $(\vec{x}P)P = \vec{x}P^2$, y así sucesivamente. Por lo tanto, podemos calcular la distribución del internauta sobre los estados en cualquier momento, dada solo la distribución inicial y la matriz de probabilidad de transición P.
Si se permite que una cadena de Markov se ejecute durante muchos pasos de tiempo, cada estado se visita a una frecuencia (diferente) que depende de la estructura de la cadena de Markov. En nuestra analogía de ejecución, el internauta visita ciertas páginas web (por ejemplo, páginas de inicio de noticias populares) con más frecuencia que otras páginas. Ahora hacemos esta intuición precisa, estableciendo las condiciones bajo las cuales la frecuencia de visita converge a una cantidad fija en un estado estable.

Este documento fue publicado por: Cambridge University Press (Julio 04, 2009) Markov chains. [En línea]. Disponible: https://nlp.stanford.edu/IR-book/html/htmledition/markov-chains-1.html#fig:figmarkov