Introducción a las cadenas de Markov

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
    - la temperatura,

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
    - la temperatura,
    - los precios de las acciones,

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,

Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,

Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
    - la temperatura,
    - los precios de las acciones,
    - popularidad de los políticos,
    - el uso de la CPU de un ordenador,
    - la velocidad de una determinada conexión a internet,
    - etc.

Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,
  - etc.
- De cara a modelar una determinada situación, necesitamos que las variables aleatorias cambien con el tiempo.

Introducción a los procesos estocásticos

- El mundo que nos rodea es dinámico, va cambiando con el tiempo:
  - la temperatura,
  - los precios de las acciones,
  - popularidad de los políticos,
  - el uso de la CPU de un ordenador,
  - la velocidad de una determinada conexión a internet,
  - etc.
- De cara a modelar una determinada situación, necesitamos que las variables aleatorias cambien con el tiempo.
- Por dicho motivo, vamos a introducir los procesos estocásticos, que son variables aleatorias que además de depender de los elementos del espacio muestral, dependen del tiempo.

Procesos estocásticos

Un proceso estocástico es una variable aleatoria $X(w, t)$ que depende de dos argumentos:

- un elemento del espacio muestral $w \in \Omega$,
- el tiempo $t \in \mathcal{T}$, donde el conjunto $\mathcal{T}$ puede ser
  - discreto: $\mathcal{T} = \{0, 1, 2, \ldots\}$, o $\mathcal{T} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$,
  - continuo: $\mathcal{T} = [0, \infty)$ o $\mathcal{T} = [-\infty, \infty)$.

- Si fijamos el tiempo t, la variable aleatoria $X(w, t) = X_t(w)$ sería una variable aleatoria que modelizaría lo que pasa en el sistema en el instante t.

Procesos estocásticos

Procesos estocásticos

Un proceso estocástico es una variable aleatoria $X(w,t)$ que depende de dos argumentos:

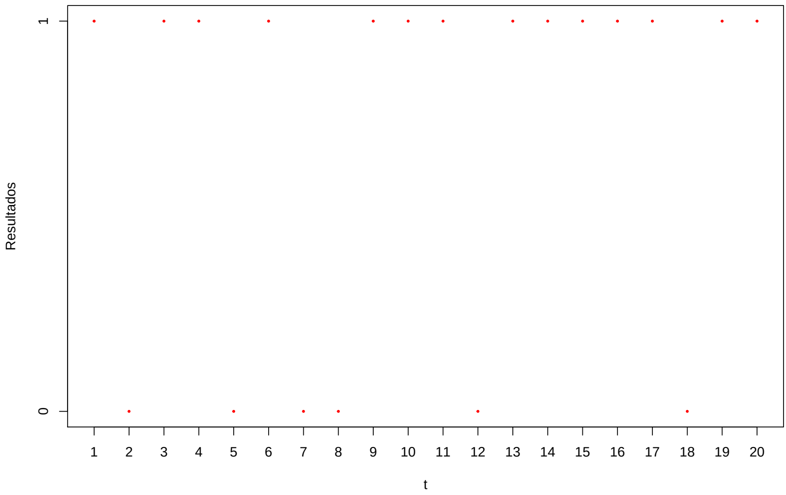
- un elemento del espacio muestral $w \in \Omega$,
- el tiempo $t \in \mathcal{T}$, donde el conjunto $\mathcal{T}$ puede ser
  - discreto: $\mathcal{T} = \{0,1,2,\ldots\}$, o $\mathcal{T} = \{\ldots,-2,-1,0,1,2,\ldots\}$,
  - continuo: $\mathcal{T} = [0,\infty)$ o $\mathcal{T} = [-\infty,\infty)$.

- Si fijamos el tiempo t, la variable aleatoria $X(w,t) = X_t(w)$ sería una variable aleatoria que modelizaría lo que pasa en el sistema en el instante t.
- Si fijamos el elemento w del espacio muestral Ω, tenemos la función dependiendo del tiempo $X(w,t) = X_w(t)$. Dicha función se denomina trayectoria del proceso estocástico.

Ejemplo: lanzamiento de una moneda

Un proceso estocástico sencillo es considerar lanzar una moneda cada cierto espacio de tiempo, por ejemplo cada minuto, y observar el resultado.
En este caso $\Omega = \{c, +\}$ y $\mathcal{T} = \{0, 1, 2, \ldots\}$.
$X(w, t)$ sería la variable aleatoria que nos dice el comportamiento de la moneda en el lanzamiento t-ésimo.
La distribución de dicha variable será de Bernoulli de parámetro p para cualquier instante t donde p es la probabilidad de sacar cara.

Procesos estocásticos



Tipos de procesos estocásticos

Diremos que el proceso estocástico es de estado discreto si la variable aleatoria $X_t(w)$ es discreta y es de estado continuo si la variable aleatoria $X_t(w)$ es continua para todo valor del tiempo t. Diremos que el proceso estocástico es de tiempo discreto si el conjunto $\mathcal{T}$ de valores del tiempo es discreto y es de tiempo continuo si $\mathcal{T}$ es continuo.

- Ejemplo anterior del lanzamiento de una moneda de parámetro p: estado discreto y tiempo discreto.

Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro p: estado discreto y tiempo discreto.
- Número de infectados diarios por una pandemia: estado discreto y tiempo discreto.

Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro p: estado discreto y tiempo discreto.
- Número de infectados diarios por una pandemia: estado discreto y tiempo discreto.
- Temperatura en un lugar determinado: estado continuo y tiempo continuo.

Procesos estocásticos. Ejemplos

- Ejemplo anterior del lanzamiento de una moneda de parámetro p: estado discreto y tiempo discreto.
- Número de infectados diarios por una pandemia: estado discreto y tiempo discreto.
- Temperatura en un lugar determinado: estado continuo y tiempo continuo.
- Tiempo de espera del n-ésimo cliente de una cola en el supermercado modelado como $X(n, w)$. Este ejemplo es "lioso" ya que el tiempo sería el número del cliente y el espacio muestral sería precisamente el tiempo. Pensar que la variable aleatoria "tiempo" dependerá del número de cliente n: estado continuo y tiempo discreto.

Procesos de Markov

Cadena de Markov

Un proceso estocástico $X(t)$ es un proceso de Markov si para cualquier secuencia de valores $t_1 < \cdots < t_n < t$ y para cualquier secuencia de sucesos $A, A_1, \ldots, A_n$,

$$P(X(t) \in A | X(t_1) \in A_1, \ldots, X(t_n) \in A_n)$$
$$= P(X(t) \in A | X(t_n) \in A_n).$$

- Es decir, que la distribución condicionada de la variable $X(t)$ condicionada a los valores del proceso estocástico en n instantes cualesquiera del pasado sólo depende de la distribución condicionada de la variable $X(t)$ condicionada al proceso correspondiente al último instante de la secuencia.

- Esquemáticamente, podemos escribir:

$$
\begin{aligned}
&P(\text{FUTURO}|\text{PASADO},\ \text{PRESENTE})\\
&= P(\text{FUTURO}|\text{PRESENTE}).
\end{aligned}
$$

- Esquemáticamente, podemos escribir:

$$
\begin{aligned}
& P(\text{FUTURO}|\text{PASADO},\ \text{PRESENTE}) \\
& = P(\text{FUTURO}|\text{PRESENTE}).
\end{aligned}
$$

- Ejemplos:

Procesos de Markov

- Esquemáticamente, podemos escribir:

$$P(\text{FUTURO}|\text{PASADO},\ \text{PRESENTE})$$
$$= P(\text{FUTURO}|\text{PRESENTE}).$$

- Ejemplos:
  - Temperatura en un día determinado: no es un proceso de Markov.

Procesos de Markov

- Esquemáticamente, podemos escribir:

$$P(\text{FUTURO}|\text{PASADO},\ \text{PRESENTE})$$
$$= P(\text{FUTURO}|\text{PRESENTE}).$$

- Ejemplos:
  - Temperatura en un día determinado: no es un proceso de Markov.
  - Número de conexiones registradas en un router de internet en un instante determinado: sí es un proceso de Markov ya que la gente se conecta aleatoriamente.

- Esquemáticamente, podemos escribir:

$$P(\text{FUTURO}|\text{PASADO, PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}).$$

- Ejemplos:
  - Temperatura en un día determinado: no es un proceso de Markov.
  - Número de conexiones registradas en un router de internet en un instante determinado: sí es un proceso de Markov ya que la gente se conecta aleatoriamente.
  - Precio de un stock en la bolsa en un día determinado: no es un proceso de Markov.

- Esquemáticamente, podemos escribir:

$$P(\text{FUTURO}|\text{PASADO, PRESENTE}) \\ = P(\text{FUTURO}|\text{PRESENTE}).$$

- Ejemplos:
  - Temperatura en un día determinado: no es un proceso de Markov.
  - Número de conexiones registradas en un router de internet en un instante determinado: sí es un proceso de Markov ya que la gente se conecta aleatoriamente.
  - Precio de un stock en la bolsa en un día determinado: no es un proceso de Markov.
  - Ejemplo del lanzamiento de la moneda: sí es un proceso de Markov.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos metereológicos básicos.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos metereológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos metereológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos metereológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.
  - Economía y finanzas: modelización del colapso de una bolsa de valores.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos metereológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.
  - Economía y finanzas: modelización del colapso de una bolsa de valores.
  - Genética: teoría genética de poblaciones.

Cadenas de Markov

- Dentro de los procesos de Markov están las denominadas cadenas de Markov:

Cadenas de Markov

Una cadena de Markov es un proceso de Markov de estado discreto y de tiempo discreto.

- Las cadenas de Markov se aplican a:
  - Meteorología: modelización de modelos metereológicos básicos.
  - Modelos epidemiológicos: modelización de una epidemia.
  - Internet: el pagerank usado por Google para dar un peso a las páginas web.
  - Economía y finanzas: modelización del colapso de una bolsa de valores.
  - Genética: teoría genética de poblaciones.
  - Redes neuronales: se utilizan en las máquinas de Boltzmann.

- Vamos a introducir algunas simplificaciones:

- Vamos a introducir algunas simplificaciones:
  - Definiremos $\mathcal{T} = \{0, 1, 2 \ldots\}$. Entonces la cadena de Markov sería una secuencia aleatoria de variables aleatorias $X(0), X(1), X(2), \ldots$.

- Vamos a introducir algunas simplificaciones:
  - Definiremos $\mathcal{T} = \{0, 1, 2 \ldots\}$. Entonces la cadena de Markov sería una secuencia aleatoria de variables aleatorias $X(0), X(1), X(2), \ldots$.
  - Llamaremos al conjunto Ω conjunto de estados. Como es discreto, lo enumeraremos de la forma siguiente $\Omega = \{1, 2, \ldots, n\}$.

- Por ser una cadena de Markov, un proceso de Markov, tenemos la denominada propiedad de Markov que nos dice que la variable aleatoria en un instante $t+1$ sólo depende de los valores que toma la variable aleatoria $X(t)$ o el proceso en el instante t.

- Por ser una cadena de Markov, un proceso de Markov, tenemos la denominada propiedad de Markov que nos dice que la variable aleatoria en un instante $t+1$ sólo depende de los valores que toma la variable aleatoria $X(t)$ o el proceso en el instante t.
- Es decir, $p_{ij}(t) = P(X(t+1) = j | X(t) = i)$ vale:

$$\begin{aligned} p_{ij}(t) =& P(X(t+1) = j | X(t) = i) \\ =& P(X(t+1) = j | X(t) = i, X(t-1) = h, X(t-2) = g, \ldots) \end{aligned}$$

Probabilidades de transición

Las probabilidades $p_{ij}(t)$ se llaman probabilidades de transición.
La probabilidad

$$p_{ij}^{(h)}(t) = P(X(t+h) = j | X(t) = i),$$

es la probabilidad de ir desde el estado i hasta el estado j usando h transiciones. Dicha probabilidad se llama probabilidad de transición de h pasos.

Cadenas de Markov homogéneas

Una cadena de Markov es homogénea cuando las probabilidades de transición no dependen del tiempo t.
Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:

Cadenas de Markov homogéneas

Una cadena de Markov es homogénea cuando las probabilidades de transición no dependen del tiempo t.
Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:
  - $p_{ij} = P(X(t+1) = j | X(t) = i)$: probabilidad de transición.

Cadenas de Markov homogéneas

Una cadena de Markov es homogénea cuando las probabilidades de transición no dependen del tiempo t.
Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:
  - $p_{ij} = P(X(t+1) = j|X(t) = i)$: probabilidad de transición.
  - $p_{ij}^{(h)} = P(X(t+h) = j|X(t) = i)$: probabilidad de transición en h pasos.

Cadenas de Markov homogéneas

Una cadena de Markov es homogénea cuando las probabilidades de transición no dependen del tiempo t.
Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:
  - $p_{ij} = P(X(t+1) = j | X(t) = i)$: probabilidad de transición.
  - $p_{ij}^{(h)} = P(X(t+h) = j | X(t) = i)$: probabilidad de transición en h pasos.
  - $P_t(x) = P(X(t) = x)$: distribución de la variable aleatoria $X(t)$.

Cadenas de Markov homogéneas

Una cadena de Markov es homogénea cuando las probabilidades de transición no dependen del tiempo t.
Es decir:

$$p_{ij}(t) = p_{ij}, \quad p_{ij}^{(h)} = p_{ij}^{(h)}.$$

- Notación:
  - $p_{ij} = P(X(t+1) = j|X(t) = i)$: probabilidad de transición.
  - $p_{ij}^{(h)} = P(X(t+h) = j|X(t) = i)$: probabilidad de transición en h pasos.
  - $P_t(x) = P(X(t) = x)$: distribución de la variable aleatoria $X(t)$.
  - $P_0(t) = P(X(0) = x)$: distribución inicial o en el tiempo 0.

Ejemplo

Un ordenador es compartido por dos usuarios. Estudiamos cuantos usuarios hay conectados cada minuto. Nos dicen que un usuario puede desconectarse con probabilidad 0.3 y se puede conectar con probabilidad 0.4.

Modelizamos la situación anterior con una cadena de Markov $X(t)$ que nos da el número de usuarios conectados al cabo de t minutos.

Los valores de $X(t)$ pueden ser $0, 1$ y 2. Por tanto, el conjunto de estados será: $\Omega = \{0, 1, 2\}$.

A continuación, calculemos las probabilidad de transición:

Ejemplo

Si $X(0) = 0$, o no hay ningún usuario conectado, el número de usuarios que habrá en el siguiente minuto es una variable aleatoria binomial de parámetros $n = 2$ y $p = 0.4$:

$$p_{00} = \binom{2}{0} \cdot 0.4^0 \cdot 0.6^2 = 0.36,$$

$$p_{01} = \binom{2}{1} \cdot 0.4^1 \cdot 0.6^1 = 0.48,$$

$$p_{02} = \binom{2}{2} \cdot 0.4^2 \cdot 0.6^0 = 0.16.$$

Ejemplo

Si $X(0) = 1$, o hay un usuario conectado, el número de usuarios nuevos que habrá en el siguiente minuto es una variable aleatoria binomial de parámetros $n = 1$ y $p = 0.4$ y el número de desconexiones en el minuto siguiente será una variable aleatoria binomial de parámetros $n = 1$ y $p = 0.3$:

$$\begin{aligned} p_{10} &= 0.6 \cdot 0.3 = 0.18, \\ p_{11} &= 0.3 \cdot 0.4 + 0.7 \cdot 0.6 = 0.54, \\ p_{12} &= 0.7 \cdot 0.4 = 0.28. \end{aligned}$$

Ejemplo

Si $X(0) = 2$, o hay los dos usuarios conectados, el número de usuarios desconectados que habrá en el siguiente minuto es una variable aleatoria binomial de parámetros $n = 2$ y $p = 0.3$.

$$p_{20} = \binom{2}{0} \cdot 0.7^0 \cdot 0.3^2 = 0.09,$$

$$p_{21} = \binom{2}{1} \cdot 0.7^1 \cdot 0.3^1 = 0.42,$$

$$p_{22} = \binom{2}{2} \cdot 0.7^2 \cdot 0.3^0 = 0.49.$$

Diagrama de transición del ejemplo.