CIND-221: Introducción a Procesos Estocásticos

Felipe Osorio

f.osoriosalgado@uandresbello.edu

Facultad de Ingeniería, UNAB

Definición 1 (Proceso estocástico):

Un proceso estocástico es una colección de variables aleatorias $\{X_t, t \in T\}$ definidas en un mismo espacio de probabilidades $(\Omega, \mathcal{F}, \mathsf{P})$ con T el conjunto de índices del proceso.[1]

El conjunto de valores que adopta la variable X_t es llamado espacio de estados del proceso y es denotado por S, y cada elemento de S es un estado del proceso.

[1] También llamado espacio paramétrico.

Observación:

Para cada $\omega \in \Omega$ fijo, tenemos que la transformación definida como:

$$\begin{aligned} X(\omega) : T &\to S \\ t &\to X_t(\omega) \end{aligned}$$

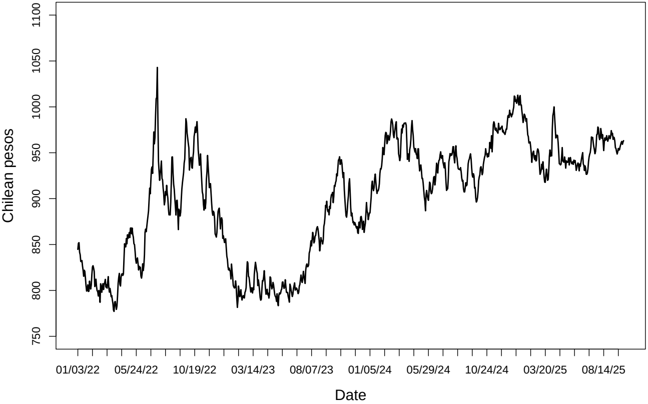
es llamada trayectoria del proceso.[2]

A continuación se presenta ejemplos de trayectorias con un conjunto de datos reales y mediante datos simulados, a saber:

- Dólar observado vs. peso chileno desde Enero-2022 a Octubre-2025.
- Trayectorias simuladas desde una caminata aleatoria.

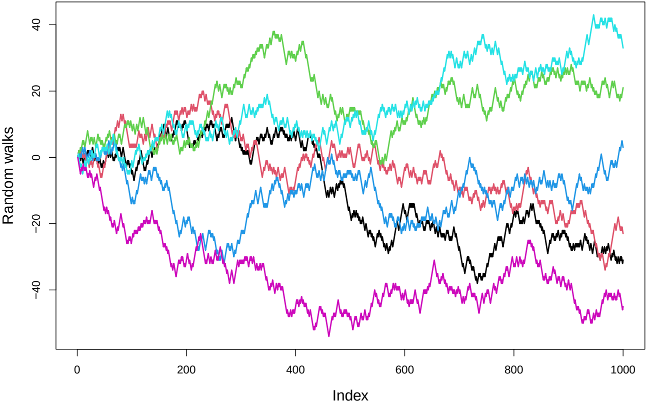
[2] o realización del proceso estocástico.

Trayectoria del dólar observado vs. peso chileno desde Enero-2022 a Octubre-2025[3]



[3] Datos extraídos desde página del Banco Central.

Trayectorias simuladas



Procesos Estocásticos

Procesos estocásticos pueden clasificarse en 4 tipos, dependiendo de la naturaleza del espacio de estados y espacio paramétrico.

- Proceso estocástico de tiempo discreto y espacio de estado discreto
  - El número de individuos en una población al final del año t, que se puede modelar como $\{X_t : t \in T\}$, donde $T = \{0, 1, 2, \dots\}$ y $S = \{0, 1, 2, \dots\}$.
- Proceso estocástico de tiempo continuo y espacio de estado discreto
  - El número de llamadas entrantes X_t en un intervalo $[0, t]$. Es decir, para el proceso estocástico $\{X_t : t \in T\}$ tenemos $T = \{t : 0 \leq t < \infty\}$ y $S = \{0, 1, \dots\}$.

- Proceso estocástico de tiempo discreto y espacio de estado continuo
  - El precio de cierre de una acción en el día t. De este modo el proceso $\{X_t : t \in T\}$, tiene $T = \{0, 1, \dots\}$ y $S = \{x : 0 \leq x < \infty\}$.

- Proceso estocástico de tiempo continuo y espacio de estado continuo
  - El flujo de un río que se observa en un año.[4] En este caso

$$T = \{t : 0 \leq t < \infty\}, \qquad S = \{x : 0 \leq x < \infty\}.$$

[4] X_t es el flujo en el tiempo t.

Definición 2:

Si, para todo $t_0, t_1, \dots, t_n$ tal que $t_0 < t_1 < \cdots < t_n$, las variables aleatorias

$$X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}},$$

son independientes, entonces el proceso $\{X_t : t \in T\}$ se dice un proceso con incrementos independientes.

Definición 3:

Un proceso estocástico $\{X_t : t \in T\}$ se dice con incrementos estacionarios si

$$X_{t+h} - X_{s+h} \stackrel{\mathsf{d}}{=} X_t - X_s,$$

para cualquier $t, s \in T$ y $h > 0$.

Definición 4:

Si, para todo $t_1, t_2, \ldots, t_n$ la distribución conjunta de los vectores

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_n})^\top, \qquad (X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h})^\top,$$

es la misma, para todo $h > 0$. Entonces, El proceso $\{X_t : t \in T\}$ se dice estacionario.

Notación:

Sea $\{X_t : t \in T\}$ un proceso estocástico. Entonces, denotamos:

- $m(t) = \mathsf{E}(X_t)$, $t \in T$ a la función de media del proceso.
- $K(s,t) = \mathrm{cov}(X_s, X_t)$, para $s, t \in T$ como la función de covarianza del proceso.

Definición 5:

Un proceso estocástico $\{X_t, t \in T\}$ se denomina proceso de segundo orden si $\mathsf{E}(X_t^2) < \infty$, para todo $t \in T$.

Ejemplo:

Sean Z_1 y Z_2 variables aleatorias independientes normalmente distribuídas, cada una con media 0 y varianza σ^2. Para $\lambda \in \mathbb{R}$, considere

$$X_t = Z_1 \cos(\lambda t) + Z_2 \sin(\lambda t), \quad t \in \mathbb{R}.$$

Entonces $\{X_t : t \in \mathbb{R}\}$ es un proceso estacionario de segundo orden.

Note que la función de medias es dada por

$$\mathsf{E}(X_t) = \mathsf{E}\{Z_1 \cos(\lambda t) + Z_2 \sin(\lambda t)\} = \mathsf{E}(Z_1)\cos(\lambda t) + \mathsf{E}(Z_2)\sin(\lambda t).$$

Sin embargo, $Z_j \sim \mathsf{N}(0, \sigma^2)$, para $j = 1, 2$. De este modo, $\mathsf{E}(X_t) = 0$.

Por otro lado,

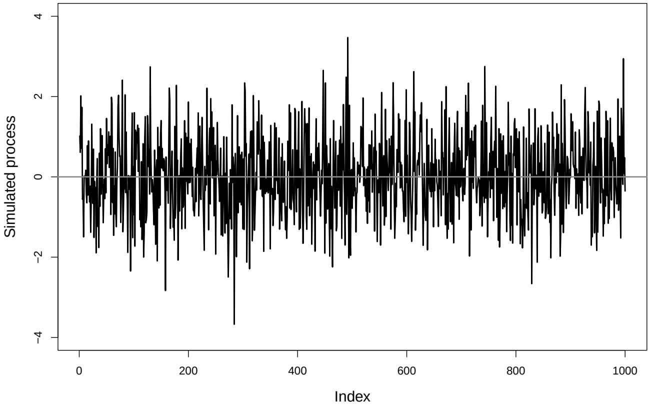
$$\begin{aligned}
\mathsf{E}(X_t^2) &= \mathsf{E}\{[Z_1 \cos(\lambda t) + Z_2 \sin(\lambda t)]^2\} \\
&= \mathsf{E}\{Z_1^2 \cos^2(\lambda t) + Z_2^2 \sin^2(\lambda t) + Z_1 Z_2 \cos(\lambda t)\sin(\lambda t)\} \\
&= \mathsf{E}(Z_1^2)\cos^2(\lambda t) + \mathsf{E}(Z_2^2)\sin^2(\lambda t) + \mathsf{E}(Z_1 Z_2)\cos(\lambda t)\sin(\lambda t),
\end{aligned}$$

como Z_1 y Z_2 son independientes, sigue que $\mathsf{E}(Z_1 Z_2) = \mathsf{E}(Z_1)\,\mathsf{E}(Z_2) = 0$, y como $\mathsf{E}(Z_1^2) = \mathsf{E}(Z_2^2) = \sigma^2$, obtenemos:

$$\mathsf{E}(X_t^2) = \sigma^2\{\cos^2(\lambda t) + \sin^2(\lambda t)\} = \sigma^2 < \infty.$$

Es decir, el proceso $\{X_t : t \in \mathbb{R}\}$ es de segundo orden.

Proceso simulado: $X_t = Z_{1t}\cos(\lambda t) + Z_{2t}\sin(\lambda t)$, con $t \in [1, T]$ donde $T = 1000$, $\{Z_{jt}\}$ es simulado desde $\mathrm{N}(0, 1)$, para $j = 1, 2$, y $\lambda = 1$.



Definición 6:

Un proceso estocástico de segundo orden $\{X_t, t \in T\}$ se dice débilmente estacionario si $m(t) = \mathsf{E}(X_t)$ es independiente de t y su función de covarianza $K(s,t)$ depende sólo de la diferencia $|t-s|$, es decir,

$$\mathrm{cov}(X_s, X_t) = f(|t-s|).$$

Ejemplo:

Sea $\{X_n : n \geq 1\}$ variables aleatorias no correlacionadas con media 0 y varianza 1. Entonces

$$\mathrm{cov}(X_r, X_n) = \mathsf{E}(X_r X_n) = \begin{cases} 0, & r \neq n, \\ 1, & r = n. \end{cases}$$

Entonces $\{X_n : n \geq 1\}$ es un proceso débilmente estacionario.

Definición 7 (Proceso de Markov):

Sea $\{X_t : t \geq 0\}$ un proceso estocástico definido sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathsf{P})$ con espacio de estados S discreto. Se dice que $\{X_t : t \geq 0\}$ es un proceso de Markov si para cualquier $0 \leq t_1 < t_2 < \cdots < t_{n-1} < t_n$ y para todo x_1, $x_2, \ldots, x_{n-1}, y$, se tiene que

$$\mathsf{P}(X_{t_n} = y | X_{t_1} = x_1, \ldots, X_{t_{n-1}} = x_{n_1}) = \mathsf{P}(X_{t_n} = y | X_{t_{n-1}} = x_{n-1}).$$

Observación:

Cualquier proceso estocástico que tenga incrementos independientes es un proceso de Markov.

Definición 8 (Función de transición):

Si $\{X_t : t \geq 0\}$ es un proceso de Markov, a la probabilidad condicional:

$$p_{st}(x, y) = \mathsf{P}(X_t = y | X_s = x),$$

se le denomina función de transición del proceso.[5]

Observación:

A continuación supondremos que la función de transición será estacionaria,

$$p_t(x, y) = p_{uv}(x, y) = \mathsf{P}(X_v = y | X_u = x),$$

con $t = v - u \geq 0$.

[5] Es decir, $p_{st}(x, y)$ es la probabilidad de transición del tiempo s al tiempo t $(s \to t)$.

Definición 9 (Cadena de Markov):

Sea S un conjunto discreto. Una cadena de Markov es una secuencia de variables aleatorias $\{X_n : n \in \mathbb{N}\}$ que satisface

$$\mathsf{P}(X_{n+1} = y | X_1 = x_1, \ldots, X_{n-1} = x_{n-1}, X_n = x) = \mathsf{P}(X_{n+1} = y | X_n = x),$$

para todo $x_1, \ldots, x_{n-1}, x, y \in S$ y $n \geq 1$.

Observación:

Podemos interpretar la propiedad anterior como: *Si conocemos* $X_n = x$, *conocer la historia anterior* $X_{n-1}, X_{n-2}, \ldots, X_1$ *no tiene influencia en el estado futuro* X_{n+1}.

Definición 10 (Probabilidad de transición):

Sea $\{X_n : n \geq 0\}$ una cadena de Markov. Las probabilidades

$$p_{ij} = \mathsf{P}(X_{n+1} = j | X_n = i),$$

son llamadas probabilidades de transición.

Observación:

A continuación caracterizaremos una cadena de Markov organizando las probabilidades p_{ij} en una matriz. Considere la siguiente definición.

Definición 11 (Vector de probabilidades):

Sea $\boldsymbol{p} = (p_1, \dots, p_n)^\top$ se dice un vector de probabilidades, si satisface:

(i) $p_i \geq 0$, $i = 1, \dots, n$.

(ii)[6] $\displaystyle\sum_{i=1}^{n} p_i = 1$.

[6] o equivalentemente $\mathbf{1}^\top \boldsymbol{p} = 1$ con $\mathbf{1} = (1, \dots, 1)^\top$

Definición 12 (Cadena homogénea):

Una cadena de Markov $\{X_n : n \geq 0\}$ es llamada homogénea si las probabilidades de transición no dependen de n. Esto es,

$$p_{ij} = \mathsf{P}(X_1 = j | X_0 = i) = \mathsf{P}(X_{n+1} = j | X_n = i).$$

Definición 13 (Distribución inicial):

La distribución de probabilidades $\pi := \{\pi_i\}_{i \in S}$, con

$$\pi_i = \mathsf{P}(X_0 = i),$$

es llamada distribución inicial.

Definición 13 (Matriz de transición):

La matriz

$$\boldsymbol{P} = (p_{ij}) = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

es llamada matriz de transición.

Note que una matriz de transición es una matriz estocástica[7] que satisface:

(i) $p_{ij} \geq 0$, para todo $i, j \in S$.

(ii) $\sum_j p_{ij} = 1$, para todo $j \in S$.

[7] Una matriz de transición también es llamada matriz de probabilidad.

Ejemplo:

Considere la matriz

$$\boldsymbol{P} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}.$$

En efecto, $p_{ij} > 0$, $\forall\, i, j$, y

$$\begin{aligned} 0.5 + 0.4 + 0.1 &= 1, \\ 0.3 + 0.4 + 0.3 &= 1, \\ 0.2 + 0.3 + 0.5 &= 1. \end{aligned}$$

Ejemplo (Secuencia IID):

Asuma que $X_0, X_2, \ldots$ es una secuencia de variables aleatorias IID que toma valores en $\{1, \ldots, k\}$, con

$$\mathsf{P}(X_n = j) = p_j, \qquad \text{para } j = 1, \ldots, k, \text{ y } n \geq 0,$$

donde $p_1 + \cdots + p_k = 1$. Por independencia, tenemos que

$$\mathsf{P}(X_1 = j | X_0 = i) = \mathsf{P}(X_1 = j) = p_j.$$

La matriz de transición es dada por:

$$\boldsymbol{P} = \begin{pmatrix} p_1 & p_2 & \cdots & p_k \\ p_1 & p_2 & \cdots & p_k \\ \vdots & \vdots & & \vdots \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}.$$

Ejemplo (Caminata aleatoria):

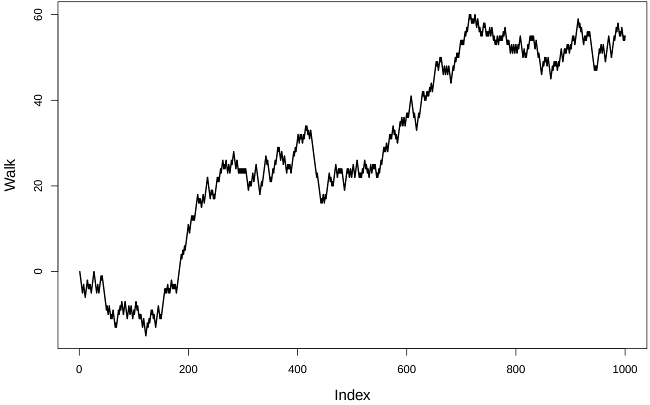
Una cadena de Markov cuyo espacio de estados es dado por los enteros $i = 0, \pm 1, \pm 2, \ldots$ se dice una caminata aleatoria si para $0 < p < 1$, tenemos

$$p_{i,i+1} = p = p_{i,i-1}, \qquad i = 0, \pm 1, \pm 2, \ldots,$$

con matriz de transición

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1-p & 0 & p & \cdots & & & & \cdots & 0 \\ 0 & 0 & 1-p & 0 & p & \cdots & & & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & & & \vdots \\ & & \cdots & 0 & 1-p & 0 & p & 0 & \cdots & 0 \\ & & & \cdots & 0 & 1-p & 0 & p & \cdots & 0 \\ & & & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \end{pmatrix}.$$

Caminata aleatoria 1-D[8]



[8] En inglés, random walk.

Resultado 1:

(a) Considere $\boldsymbol{P}$ matriz de transición, entonces la condición (ii) puede ser expresada como:

$$\boldsymbol{P}\mathbf{1} = \mathbf{1}.$$

(b) Cualquier fila de la matriz $\boldsymbol{P}$ es un vector de probabilidades.

Demostración:

Considere $\boldsymbol{P} \in \mathbb{R}^{n\times n}$. Para verificar (a), note que $\boldsymbol{P}\mathbf{1} = \mathbf{1}$, es decir:

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Es decir,

$$
\begin{gathered}
p_{11} + p_{12} + \cdots + p_{1n} = 1 \\
p_{21} + p_{22} + \cdots + p_{2n} = 1 \\
\vdots \\
p_{n1} + p_{n2} + \cdots + p_{nn} = 1
\end{gathered}
$$

o equivalentemente

$$\sum_{j=1}^{n} p_{1j} = 1,\ \sum_{j=1}^{n} p_{2j} = 1, \quad \ldots \quad \sum_{j=1}^{n} p_{nj} = 1.$$

De este modo,

$$\sum_{j=1}^{n} p_{ij}, \qquad i = 1, \ldots, n.$$

La demostración de (b) es directa.

Resultado 2:

Si $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ es matriz de transición y $\boldsymbol{q}$ es vector de probabilidades n-dimensional. Entonces $\boldsymbol{q}^\top \boldsymbol{P}$ es vector de probabilidades.

Demostración:

Sea $\boldsymbol{r}^\top = \boldsymbol{q}^\top \boldsymbol{P}$, con $\boldsymbol{r} = (r_1, \dots, r_n)^\top$. Evidentemente $r_i \geq 0$ para $i = 1, \dots, n$. Note que $\boldsymbol{r}^\top \mathbf{1} = \sum_{j=1}^n r_j$. Así,

$$\boldsymbol{r}^\top \mathbf{1} = \boldsymbol{q}^\top \boldsymbol{P}\mathbf{1} = \boldsymbol{q}^\top \mathbf{1} = 1,$$

pues $\boldsymbol{q}$ es vector de probabilidades.

Resultado 3:

Si $\boldsymbol{P} = (p_{ij})$ y $\boldsymbol{Q} = (q_{ij})$ son matrices de transición $n \times n$. Entonces el producto $\boldsymbol{PQ}$ es una matriz de transición.

Demostración:

Sabemos que $\boldsymbol{P1} = \boldsymbol{1}$ y $\boldsymbol{Q1} = \boldsymbol{1}$. Deseamos mostrar $\boldsymbol{PQ1} = \boldsymbol{1}$. En efecto,

$$\boldsymbol{PQ1} = \boldsymbol{P1} = \boldsymbol{1},$$

lo que verifica el resultado.

Resultado 4:

Si $\boldsymbol{P}$ es matriz de transición $n \times n$, entonces $\boldsymbol{P}^2, \boldsymbol{P}^3, \dots, \boldsymbol{P}^m$ son matrices de transición.

Demostración:

El resultado sigue inmediatamente desde el Resultado 3. En efecto, tomando $\boldsymbol{Q} = \boldsymbol{P}$, sigue que:

$$\boldsymbol{P}^2\mathbf{1} = \boldsymbol{P}\boldsymbol{Q}\mathbf{1} = \mathbf{1},$$

análogamente para $\boldsymbol{P}^3\mathbf{1} = \boldsymbol{P}^2\boldsymbol{Q}\mathbf{1} = \mathbf{1}$ y así sucesivamente para $\boldsymbol{P}^m$.[9]

[9] Este resultado también se puede mostrar usando inducción.

Resultado 5 (Ecuaciones de Chapman-Kolmogorov):

Si la secuencia de variables aleatorias $\{X_n : n \geq 0\}$ es una cadena de Markov y si $k < m < n$, entonces tenemos que para todo $h, j \in S$,

$$\mathsf{P}(X_n = j | X_k = h) = \sum_{i \in S} \mathsf{P}(X_n = j | X_m = i)\, \mathsf{P}(X_m = i | X_k = k).$$

Definición 14:

Sea $\{X_n : n \in \mathbb{N}\}$ una cadena de Markov, la probabilidad de transición en m pasos desde el estado i al j, es dada por:

$$p_{ij}^{(m)} = \mathsf{P}(X_{n+m} = j | X_n = i).$$

Además,

$$p_{ij}^{(0)} = \mathsf{P}(X_n = j | X_n = i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}.$$

Observación:

La probabilidad $p_{ij}^{(m)}$ es estacionaria si y solo si,

$$p_{ij}^{(m)} = \mathsf{P}(X_{n+m} = j | X_n = i) = \mathsf{P}(X_m = j | X_0 = i).$$

Observación:

Las probabilidades de transición de m-pasos, pueden ser escritas en la matriz de transición:

$$\boldsymbol{P}^{(m)} = (p_{ij}^{(m)}).$$

Es decir, $\boldsymbol{P}^{(0)} = \boldsymbol{I}$.

Definición 15 (Cadena de Markov homogénea):

Una cadena de Markov cuyas probabilidades de transición en m pasos son todas estacionarias es llamada cadena de Markov homogénea.

Proposición 1:

Las ecuaciones de Chapman-Kolmogorov proveen un procedimiento para calcular las probabilidades de transición de n-pasos. En efecto, sigue que:

$$p_{ij}^{(n+m)} = \sum_{k=0}^{\infty} p_{ik}^{(n)} p_{kj}^{(m)},$$

para todo $n, m \geq 0$ y todo $i, j \in S$.

Observación:

En forma matricial la Proposición 1 puede ser escrita como:

$$\boldsymbol{P}^{(n+m)} = \boldsymbol{P}^{(n)} \cdot \boldsymbol{P}^{(m)},$$

donde $\cdot$ denota multiplicación matricial.

Note la condición inicial,

$$\boldsymbol{P}^{(2)} = \boldsymbol{P}^{(1+1)} = \boldsymbol{P} \cdot \boldsymbol{P} = \boldsymbol{P}^2,$$

continuando por inducción tenemos que la matriz de transición de n-pasos puede ser escrita como:

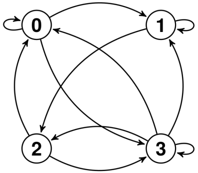
$$\boldsymbol{P}^{(n)} = \boldsymbol{P}^{(n-1+1)} = \boldsymbol{P}^{(n-1)}\boldsymbol{P} = \boldsymbol{P}^n.$$

Observación:

Si $\{X_n : n \in \mathbb{N}\}$ es una cadena de Markov con conjunto de estados $S = \{0, 1, 2, 3\}$, con matriz de transición

$$\boldsymbol{P} = \begin{pmatrix} 1/5 & 1/5 & 0 & 3/5 \\ 0 & 1/3 & 2/3 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix},$$

que puede ser representada por una red con vértices que indican los estados y arcos indicando transiciones.



Corolario 1 (Matriz de transición de n-pasos):

Sea $X_0, X_1, \ldots$ una cadena de Markov con matriz de transición $\boldsymbol{P}$. La matriz de transición de n-pasos de la cadena es dada por:

$$\boldsymbol{P}^{(n)} = \boldsymbol{P}, \qquad n \geq 0,$$

con $p_{ij}^{(n)} = \mathsf{P}(X_n = j | X_0 = i)$, $\forall\, i, j$.

Ejercicio:

Considere la matriz de transición

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad q = 1 - p.$$

Obtenga $\boldsymbol{P}^3$ y verifique que es matriz de transición.

Ejercicio:

Considere una cadena de Markov con matriz de transición

$$\boldsymbol{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix},$$

donde $p, q \in (0,1)$ y $p + q > 0$. Determine $\boldsymbol{P}^n$.

Sugerencia:

- Calcular valores y vectores propios de $\boldsymbol{P}$, esto es $\boldsymbol{P}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$, $i = 1, 2$.
- Sea $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2)$ y $\boldsymbol{U} = (\boldsymbol{u}_1, \boldsymbol{u}_2)$, luego $\boldsymbol{P} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1}$.
- Ahora, $\boldsymbol{P}^n = \boldsymbol{U}\boldsymbol{\Lambda}^n\boldsymbol{U}^{-1}$ (¿por qué?).
- Calcular $\boldsymbol{U}^{-1}$ y $\boldsymbol{\Lambda}^n$ para obtener $\boldsymbol{P}^n$.

Ejercicio:

Considere

$$\boldsymbol{P} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{pmatrix},$$

verifique (por inducción) que la n-ésima potencia de $\boldsymbol{P}$ es dada por:

$$\boldsymbol{P}^n = \begin{pmatrix} \frac{1}{2^n} & 0 & 1-\frac{1}{2^n} \\ 0 & \frac{1}{2^n} & 1-\frac{1}{2^n} \\ 0 & 0 & 1 \end{pmatrix}.$$