Procesos de Markov

Un proceso de Markov es una sucesión de vectores $\mathbf{v_k}$ con $k \in \mathbb{N}_0$ donde pensamos que v_k indica el estado de un sistema en el paso k-ésimo. v_0 es el estado inicial, y en cada paso la transición al estado siguiente está dada por una matriz de Markov M de forma tal que

$$v_{k+1} = Mv_k$$

M es una matriz de Markov si es una matriz estocástica por columnas, es decir:

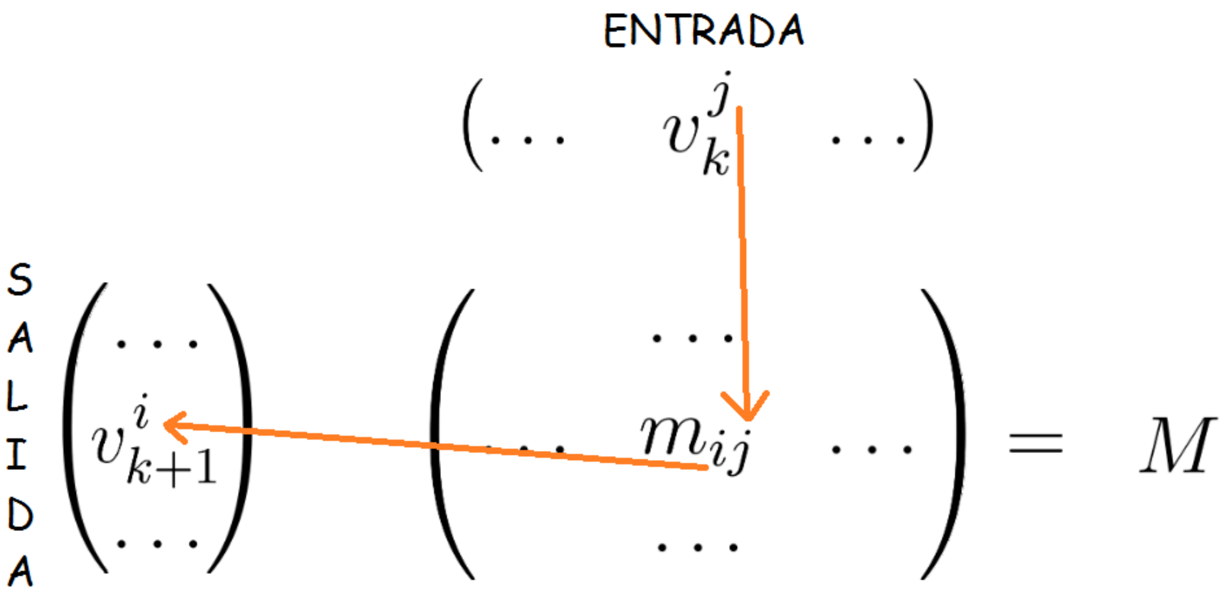
1) **Todas las entradas m_{ij} de M son no negativas.**
2) **Cada columna de M suma 1.**

Aplicaciones de procesos/cadenas de Markov: simulaciones de estructuras de proteínas, teoría de *queueing*, modelos financieros, comportamiento de poblaciones, cinética química, forma de células, modelos de lenguajes (https://medium.com/datos-y-ciencia/oraciones-artificiales-88b8c66a6e3d), teoría de juegos, termodinámica y mecánica estadística, algoritmos Monte Carlo, etc.

Las matrices positivas ($m_{ij} \geq 0$) en general tienen un autovalor de mayor módulo real y positivo.
En el caso de las matrices de Markov, ese autovalor será 1. **Toda matriz de Markov tiene autovalor 1**, y si λ es cualquier otro autovalor, entonces $|\lambda| \leq 1$.
Entonces, los autovalores de M pueden ser 1 (siempre, aunque puede tener multiplicidad mayor a 1), -1, 0 o cualquier número de módulo menor a 1.

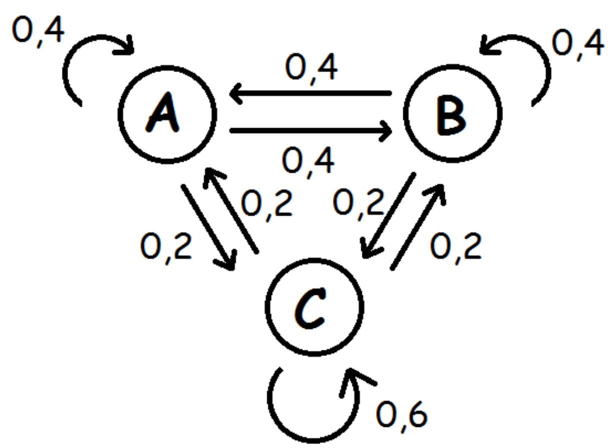
Si un vector $v = (v_1, ..., v_n)$ cumple que sus entradas son no negativas y suman 1, diremos que es un **vector estocástico o de probabilidad**. El estado global del sistema, v, está descrito por los estados v_j, $1 \leq j \leq n$ del vector v. Si v es de probabilidad, podemos pensar que cada v_j corresponde a **la probabilidad de estar en el estado j-ésimo**, o a la **proporción de la población que se encuentra en ese estado**.

Cada entrada en la matriz M representa la probabilidad de pasar de un estado j a un estado i en un paso, es decir, m_{ij} está asociada a la transición del estado j del vector v_k al estado i del vector v_{k+1}.



Problema 1

Consideremos un modelo simplificado para el precio de una acción, donde cada día el precio puede ir en alza (A), en baja (B) o permanecer constante (C). Las probabilidades de pasar de un estado al otro están representadas en el diagrama de la derecha:



a) Determinar la matriz M de transición del proceso.
b) Si la acción comienza el primer día en alza (A), ¿cuál es la probabilidad de que al cuarto día siga en alza?
c) ¿Existe un estado de equilibrio?
d) ¿Existe un vector límite v_∞ para todo vector inicial v_0? ¿Existe la matriz M^∞?

a) Tenemos un vector de estado $v = (A, B, C)$ (primera coordenada corresponde al primer estado A, etc.). Armamos la matriz M de transición:

$$M = \begin{pmatrix} 0,4 & 0,4 & 0,2 \\ 0,4 & 0,4 & 0,2 \\ 0,2 & 0,2 & 0,6 \end{pmatrix}$$

b) Al principio: la acción empieza en A. $v_0 = (1, 0, 0)$.

$v_1 = Av_0$, $v_2 = Av_1 = A^2v_0$, ..., $v_4 = A^4v_0 = (0,3376, 0,3376, 0,3248)$
(lo calculamos en Python)
La probabilidad de estar en A al cabo del cuarto día es 0,3376.

En general: $v_k = M^k v_0$.

c) Un estado de equilibrio es v tal que $v = Mv$ (o sea, v es un autovector de autovalor 1). Estar en equilibrio significa que el estado del sistema v permanece constante en el tiempo. Solemos pedir que ese v sea un vector de probabilidad (o sea que la suma de sus coordenadas dé 1) para hablar de la existencia de *un* estado de equilibrio (si no, cualquier autovector asociado a 1 es de equilibrio). Siempre hay un estado de equilibrio para toda matriz de Markov M (porque 1 siempre es autovalor de una matriz M de Markov). Si la multiplicidad de 1 es mayor a 1, podría haber más de un estado de equilibrio.

d) $v_\infty = lím\,(v_k) = lím(M^k v_0)\ cuando\ k \to \infty$

Si por ejemplo v_0 es de equilibrio, el límite existe y es justamente ese v_0.
Para un v_0 general, la existencia de v_∞ depende de la existencia del límite de la matriz M^k
Si M es diagonalizable existe C cambio de base dado por los autovectores y D matriz diagonal con los autovalores de M: $M = CDC^{-1} \Rightarrow M^k = CD^kC^{-1}$.
Queremos estudiar qué le pasa a D^k.

Calculamos los autovalores de esta M, desarrollando el polinomio característico e igualando a 0.

$$\begin{vmatrix} 0.4-\lambda & 0.4 & 0.2 \\ 0.4 & 0.4-\lambda & 0.2 \\ 0.2 & 0.2 & 0.6-\lambda \end{vmatrix}$$

$$= (0.4-\lambda)\begin{vmatrix} 0.4-\lambda & 0.2 \\ 0.2 & 0.6-\lambda \end{vmatrix} - 0.4\begin{vmatrix} 0.4 & 0.2 \\ 0.2 & 0.6-\lambda \end{vmatrix} + 0.2\begin{vmatrix} 0.4 & 0.4-\lambda \\ 0.2 & 0.2 \end{vmatrix}$$

$$= (0,4-\lambda)((0,4-\lambda)(0,6-\lambda) - 0,2^2) - 0,4(0,4(0,6-\lambda) - 0,2^2) + 0,2(0,08 - 0,2(0,4-\lambda))$$
$$= (0,4-\lambda)(0,24 - \lambda + \lambda^2 - 0,04) - 0,4(0,24 - 0,4\lambda - 0,04) + 0,2^2\lambda$$
$$= (0,4-\lambda)(0,2 - \lambda + \lambda^2) - 0,4(0,2 - 0,4\lambda) + 0,2^2\lambda$$
$$= 0,08 - 0,4\lambda + 0,4\lambda^2 - 0,2\lambda + \lambda^2 - \lambda^3 - 0,08 + 0,16\lambda + 0,04\lambda$$
$$= -0,4\lambda + 0,4\lambda^2 + \lambda^2 - \lambda^3 = -\lambda(\lambda^2 - \lambda + 0,4) = -\lambda(\lambda-1)(\lambda-0,4)$$

O si no, hay una forma más fácil… Sabemos que

- $\lambda_1 = 1$ porque M es de Markov
- $\lambda_2 = 0$ porque M no es inversible (las dos primeras filas/columnas de M están repetidas).
- $tr(M) = 0,4 + 0,4 + 0,6 = 1,4 = \lambda_1 + \lambda_2 + \lambda_3 \Rightarrow \lambda_3 = 0,4$.

Propiedad de la traza y el determinante y los autovalores

(esto es una observación aparte de lo del ejercicio)

Una matriz A nxn puede tener n autovalores reales o menos. Si trabajamos con números complejos, siempre tenemos n autovalores (contados con multiplicidad). El problema es que puede haber polinomios que no tengan raíces reales pero siempre tendrán todas las raíces en los complejos.

Si trabajamos en C, vale que el polinomio característico de A es:

$$\chi_A(x) = \prod_{i=1}^{n}(x - \lambda_i) \Rightarrow \chi_A(x) = x^n + (\lambda_1 +...+ \lambda_n)x^{n-1} + ... + \prod_{i=1}^{n}\lambda_i$$

El término con x^{n-1} es la suma de los monomios que se obtienen multiplicando n-1 veces x por algún valor de λ_i, por lo que su coeficiente será la suma de los λ_i. De la misma forma, el término independiente (sin x) se obtiene multiplicando todos los λ_i.

Por otro lado, si consideramos

$$\chi_A(\lambda) = det(A - \lambda I) = \lambda^n + (\lambda_1 +...+ \lambda_n)\lambda^{n-1} + ... + \prod_{i=1}^{n}\lambda_i$$

evaluando el polinomio en $x = \lambda$, y desarrollamos la expresión de ese determinante, observamos que (esto es un argumento convincente, no una demostración, pero se pueden afinar los detalles) al desarrollar por fila o columna, si empezamos por un elemento a_{ij} con i distinto de j, el determinante que nos queda de la matriz (n-1)x(n-1) eliminando la fila y columna de a_{ij}, tendrá a lo sumo n-2 elementos de la diagonal, es decir, los elementos $a_{ii} - \lambda$. Cuando hagamos el desarrollo para el polinomio característico, quedará un término $O(\lambda^2)$. La única forma de tener un término de la forma λ^{n-1} es tomando justo el desarrollo que incluye todos los elementos de la diagonal, es decir,

$$\prod_{i=1}^{n}(a_{ii} - \lambda) = \lambda^n - (a_{11} +...+ a_{nn})\lambda^{n-1} + O(\lambda^{n-2})$$

Igualando a la expresión de $\chi_A(\lambda)$, queda $\lambda_1 +...+ \lambda_n = a_{11} +...+ a_{nn} = tr(A)$.

Vale que la suma de los autovalores (complejos) es la traza para cualquier matriz A.

Esto, como dijimos en la clase, también se puede probar con la forma de Jordan, que es algo parecido a la diagonalización pero que se puede hacer para cualquier matriz (o sea, cuando no se puede diagonalizar también), y que excede al curso.

Para el determinante: evaluamos el polinomio característico en $\lambda = 0$:

$$\chi_A(0) = det(A - 0I) = det(A) = \prod_{i=1}^{n} \lambda_i$$

y vale que el producto de los autovalores es el determinante. Volvemos al ejercicio.

M tiene 3 autovalores distintos cada uno con al menos un autovector asociado (que serían w_1, w_2, w_3), por lo que es diagonalizable. Si $C = (w_1|w_2|w_3)$,

$$M^k = CD^kC^{-1} = C \begin{pmatrix} 1^k & & \\ & 0^k & \\ & & 0,4^k \end{pmatrix} C^{-1} \to C \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix} C^{-1} = M^\infty$$

cuando $k \to \infty$.

Existirá M^∞ si y sólo si el único autovalor de módulo 1 es el el 1.

(observemos que si tenemos -1 como autovalor queda oscilando $(-1)^k$ en la matriz diagonal y eso hace que no exista el límite).

Esta M^∞ me dice cómo es el estado límite v_∞ para cualquier v_0:

$$v_\infty = M^\infty v_0$$

Si hay un único autovector asociado a 1 normalizado (digamos, w_1), queda $M^\infty = (w_1|\dots|w_1)$

Calculemos, para la matriz M, cuál es el w_1 (autovector de autovalor 1 normalizado). Buscamos w tal que (M-I)w = 0.

$$M - I = \begin{pmatrix} -0,6 & 0,4 & 0,2 \\ 0,4 & -0,6 & 0,2 \\ 0,2 & 0,2 & -0,4 \end{pmatrix}$$

Funciona $w = (1, 1, 1) \Rightarrow w_1 = w/suma(w) = (1/3, 1/3, 1/3)$.
w_1 es el estado de equilibrio del sistema.

Observemos también que v_∞ también tiene que ser un estado de equilibrio del sistema, es decir que debe ser $v_\infty = w_1 = (1/3, 1/3, 1/3)$ (de vuelta, si tomamos los v y el w normalizados).

Problema 2

Desde el 7 de septiembre de 2018 (primer partido del ciclo Scaloni - contra Guatemala), la Scaloneta mostró que:

1. *De 63 partidos ganados (G):*
    - *en 46 ocasiones el siguiente partido fue también G.*
    - *hubo 8 empates (E) a continuación.*
    - *y sólo 9 perdidos (P) después de uno ganado.*

2. *Después de 13 partidos E:*
    - *hubo 11 G.*
    - *2 E.*
    - *ninguno perdido.*

3. *Y a continuación de cada una de las 8 derrotas (P):*
    - *hubo 6 G.*
    - *hubo 2 E.*
    - *no hubo dos derrotas consecutivas.*

1. *Armar la matriz de Markov asociada al proceso.*
2. *¿Existe un estado límite para todo vector inicial? ¿Cuál sería? Dar una interpretación.*

$$M_2 = \begin{pmatrix} 46/63 & 11/13 & 6/8 \\ 8/63 & 2/13 & 2/8 \\ 9/63 & 0 & 0 \end{pmatrix}$$

b) Ver en el Colab. Existe estado límite porque la matriz tiene autovalores 1 y dos autovalores complejos de módulo menor a 1. Para todo vector inicial de probabilidad v_0 es el autovector normalizado asociado a 1, que se puede calcular a partir de un v_0 cualquiera y aplicando la matriz M_2 repetidas veces (método de la potencia, que vale porque 1 es el único autovalor de módulo máximo). Lo hacemos numéricamente porque la cuenta es díficil a mano.

Problema 3

Sea la matriz estocástica por columnas A:

$$A = \begin{pmatrix} \frac{1}{4} - \frac{1}{4}a & \frac{1}{2} + \frac{1}{2}a^2 & 1 - a^2 \\ a & \frac{1}{2} - \frac{1}{2}a^2 & a^2 - \frac{1}{2}a \\ \frac{3}{4} - \frac{3}{4}a & 0 & a^2 - \frac{1}{2}a \end{pmatrix}$$

donde a es un número real.

1. *Hallar los posibles valores de a.*
2. *Para cada valor hallado en 1., decidir si existe estado límite v_∞ para cualquier estado inicial. ¿Existirá una matriz A^∞ tal que $A^k \to A_\infty$ cuando $k \to \infty$?*

1. Matriz estocástica por columnas: entradas sean no negativas y la suma de las columnas tiene que dar 1. Sumamos por columna e igualamos las expresiones a 1; la no negatividad de las entradas la evaluamos después.

Columna 1: $\frac{1}{4} - \frac{1}{4}a + a + \frac{3}{4} - \frac{3}{4}a = 1$

Columna 2: $\frac{1}{2} + \frac{1}{2}a^2 + \frac{1}{2} - \frac{1}{2}a^2 = 1$

Columna 3: $1 - a^2 + a^2 - \frac{1}{2}a + a^2 - \frac{1}{2}a = 1 - a + a^2 = 1$

$\Rightarrow a^2 - a = a(a-1) = 0 \Rightarrow a = 0$ ó $a = 1$

Reemplazamos primero con a=0:

$$A_0 = \begin{pmatrix} 0,25 & 0,5 & 1 \\ 0 & 0,5 & 0 \\ 0,75 & 0 & 0 \end{pmatrix}$$

A_0 efectivamente es de Markov. Existirá estado límite para todo v_0 si $\lambda = 1$ es el único autovalor de módulo 1.

Buscamos los autovalores de A_0:

$$\begin{vmatrix} 0,25-\lambda & 0,5 & 1 \\ 0 & 0,5-\lambda & 0 \\ 0,75 & 0 & -\lambda \end{vmatrix}$$

$$= 1 \begin{vmatrix} 0 & 0,5-\lambda \\ 0,75 & 0 \end{vmatrix} - \lambda \begin{vmatrix} 0,25-\lambda & 0,5 \\ 0 & 0,5-\lambda \end{vmatrix}$$

$= \ -0,75(0,5-\lambda) - \lambda(0,25-\lambda)(0,5-\lambda)$
$= (0,5-\lambda)(-0,75-\lambda(0,25-\lambda))$
$= (0,5-\lambda)(-0,75-0,25\lambda+\lambda^2) = (0,5-\lambda)(\lambda-1)(\lambda+0,75)$

$\Rightarrow \lambda_1 = 1;\ \lambda_2 = 0,5;\ \lambda_3 =-0,75$

Como el único autovalor de módulo 1 es el 1, existirá A_0^{∞} y por lo tanto existirá un vector límite v_{∞} para todo v_0. Con esto ya terminamos la parte b) para A_0.

A_0 es diagonalizable. Calculamos w_1 vector de equilibrio para ver cómo da v_{∞}:

$$A_0 - I = \begin{pmatrix} -0,75 & 0,5 & 1 \\ 0 & -0,5 & 0 \\ 0,75 & 0 & -1 \end{pmatrix}$$

Busco $(A_0 - I)w = 0\ con\ w = (x,y,z) \Rightarrow\ y = 0,\ -0,75x + z = 0$
$Tomo\ w = (1,\ 0, 0,75)$ ó $(4,\ 0,\ 3) \Rightarrow w_1 = (4/7,\ 0,\ 3/7)$
Debería ser $v_{\infty} = w_1$ para todo v_0.

Ahora para a=1:

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0,5 \\ 0 & 0 & 0,5 \end{pmatrix}$$

Buscamos los autovalores con a = 1

$$det(A_1 - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0,5 \\ 0 & 0 & 0,5-\lambda \end{vmatrix}$$

$= (0,5 - \lambda)(\lambda^2 - 1) \Rightarrow \lambda_1 = 1,\ \lambda_2 = -1,\ \lambda_3 = 0,5$

La matriz A_1 es diagonalizable, y valdrá

$$M^k = CD^kC^{-1} = C\begin{pmatrix} 1^k & & \\ & (-1)^k & \\ & & 0,5^k \end{pmatrix}C^{-1}$$

M^k no converge (no existe el límite M^∞). Eso no quiere que en ningún exista estado límite; puede existir pero dependiendo del v_0. Si v_0 justo es un estado de equilibrio, $M^k v_0 = v_0$ que converge a un estado límite (porque siempre da v_0). Pero si v_0 no es de equilibrio, v_k puede oscilar.

Podemos calcular el vector de equilibrio:

$$A_1 - I = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0,5 \\ 0 & 0 & -0,5 \end{pmatrix}$$

Funciona w = (1, 1, 0) $\Rightarrow$ $w_1 = (1/2,\ 1/2,\ 0)$. w_1 cumple que $A_1^k w_1 = w_1$, o sea, existe un estado límite si empezamos con $v_0 = w_1$. Pero para otros v_0 no necesariamente existirá este estado límite, y no existe A_1^∞.

Problema 4

Un estudio sobre preferencias de series televisivas probó que si una persona mira una vez las temporadas nuevas de los Simpsons (N), lo que hace a continuación es, con igual probabilidad, seguir mirando eso, decidir mirar de ahora en más las temporadas viejas (V), o pasar a mirar otros programas de televisión (O) y nunca volver atrás. La matriz de Markov asociada al proceso es:

$$S = \begin{pmatrix} 1 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix}$$

¿Existe un único estado de equilibrio?

Calculamos los autovectores de la matriz S:

$$det(S - \lambda I) = \begin{vmatrix} 1-\lambda & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} - \lambda & 0 \\ 0 & \frac{1}{3} & 1-\lambda \end{vmatrix} = (1-\lambda)^2(\frac{1}{3} - \lambda)$$

Obtenemos $\lambda_1 = 1$, $\lambda_2 = 1$ $(1\ es\ de\ multiplicidad\ 2)$, $\lambda_3 = 1/3$.

Calculemos los autovectores asociados a 1:

$$S - I = \begin{pmatrix} 0 & \frac{1}{3} & 0 \\ 0 & \frac{-2}{3} & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}$$

Busco $w = (x, y, z)\ tal\ que\ (S - I)w = 0 \Rightarrow y = 0$, pero x, z quedan libres
Elegimos $w_1 = (1, 0, 0)\ y\ w_2 = (0, 0, 1)$. Estos son dos vectores de equilibrio distintos, el primero está "asociado a N" y el segundo, "asociado al estado O".

Si ahora tenemos un $v_0 = a\,w_1 + bw_2 + cw_3$ (con la base de autovectores dada por w_1, w_2, w_3) y aplicamos la matriz S:

$$Sv_0 = a\,Sw_1 + b\,Sw_2 + c\,Sw_3 = a.1\,w_1 + b.1\,w_2 + c.(1/3)w_3$$

$$v_k = S^k v_0 = a\,S^k w_1 + b\,S^k w_2 + c\,S^k w_3 = a.1^k\,w_1 + b.1^k\,w_2 + c.(1/3)^k\,w_3$$

$$v_k = a\,w_1 + b\,w_2 + c.(1/3)^k\,w_3 \to aw_1 + bw_2 = v_\infty(v_0)$$

El estado límite depende de la condición inicial v_0.

Vale igual que $v_\infty = S^\infty v_0$, con $C = (w_1|w_2|w_3)$. Podemos hacer la cuenta y calcular (hicimos la cuenta en el Colab):

$$S^\infty = C\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}C^{-1} = \begin{pmatrix}1 & 0,5 & 0\\ 0 & 0 & 0\\ 0 & 0,5 & 1\end{pmatrix}$$

S^∞ ya no es de la forma $(w_1|w_1|w_1)$, ahora hay que calcularla.