

Tema 2. Aplicaciones del cálculo matricial

Consideremos un *proceso* que se desarrolla en *etapas* sucesivas. En cada etapa el *estado* del proceso viene dado por un *vector de estado*. Se supone que el vector de estado en cada etapa solamente depende del vector de estado de la etapa anterior, es decir, se trata de procesos “sin memoria”. Representaremos por $\mathbf{X}(n)$ el vector de estado en la etapa n .

Ejemplo 1. Consideremos una población de animales hembras que tienen una esperanza máxima de vida de 24 meses. Dicha población está dividida en tres grupos de edad: las crías que tienen menos de 8 meses, las jóvenes que tienen al menos 8 meses y menos de 16 meses y las adultas que tienen al menos 16 meses y menos de 24. Se hacen recuentos periódicos de la población cada 8 meses. Se supone que todas las adultas mueren al pasar de uno a otro recuento. Se sabe que en cada período $1/4$ de las crías llegan a jóvenes y $2/3$ de las jóvenes llegan a adultas. Las crías no se reproducen, el número medio de crías hembras por cada hembra joven es de 2 y por cada hembra adulta es de 3. Inicialmente hay 200 crías, 100 jóvenes y 80 adultas. Queremos estudiar la evolución de dicha población a medio y largo plazo.

En este modelo las etapas son de 8 meses. El estado de la población está descrito por un vector $\mathbf{X}(n) = (x(n), y(n), z(n))^t$ que nos da el número de crías, $x(n)$, jóvenes, $y(n)$, y adultos, $z(n)$, en la etapa n . Tenemos las siguientes relaciones.

$$\begin{aligned}x(n+1) &= 2y(n) + 3z(n) \\y(n+1) &= \frac{1}{4}x(n) \\z(n+1) &= \frac{2}{3}y(n)\end{aligned}$$

Que podemos escribir en forma matricial

$$\begin{pmatrix} x(n+1) \\ y(n+1) \\ z(n+1) \end{pmatrix} = \begin{pmatrix} 0 & 2 & 3 \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x(n) \\ y(n) \\ z(n) \end{pmatrix}$$

O bien

$$\mathbf{X}(n+1) = \mathbf{A} \cdot \mathbf{X}(n) \quad (n = 0, 1, 2, \dots) \quad (1)$$

Donde

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 3 \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \end{pmatrix}$$

Pongamos $\mathbf{X}(0) = (200, 100, 80)^t$ que nos da la distribución inicial de la población. Tenemos que

$$\mathbf{X}(1) = \mathbf{A} \cdot \mathbf{X}(0), \quad \mathbf{X}(2) = \mathbf{A} \cdot \mathbf{X}(1) = \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{X}(0) = \mathbf{A}^2 \cdot \mathbf{X}(0), \quad \mathbf{X}(3) = \mathbf{A} \cdot \mathbf{X}(2) = \mathbf{A} \cdot \mathbf{A}^2 \cdot \mathbf{X}(0) = \mathbf{A}^3 \cdot \mathbf{X}(0)$$

Y en general

$$\mathbf{X}(n) = \mathbf{A}^n \cdot \mathbf{X}(0) \quad (n = 1, 2, 3, \dots) \quad (2)$$

Esta igualdad permite calcular $\mathbf{X}(n)$ para valores pequeños de n calculando la potencia correspondiente de la matriz \mathbf{A} . Veremos más adelante cómo pueden obtenerse fácilmente resultados aproximados para valores grandes de n .

Ejemplo 2. Supongamos que al realizar estudios climáticos en una determinada zona obtenemos los siguientes datos. Si un día es caluroso, entonces la probabilidad de que el día siguiente sea también caluroso es $3/5$, y la probabilidad de que haga frío $2/5$. Por otro lado, si un día es frío, entonces $1/5$ es la probabilidad de que el día siguiente siga siendo frío y $4/5$ de que sea un día caluroso. Si hoy es un día caluroso queremos calcular la probabilidad de que dentro de cinco días sea frío. También queremos calcular la probabilidad a largo plazo de que un día sea frío o caluroso.

En este proceso las etapas son de un día y el vector de estado viene dado por $\mathbf{X}(n) = (x(n), y(n))^t$, donde $x(n)$ representa la probabilidad de que el día n sea caluroso e $y(n)$ la probabilidad de que sea frío. Inicialmente tenemos que $\mathbf{X}(0) = (1, 0)^t$. Tenemos que

$$\begin{aligned} x(n+1) &= \frac{3}{5}x(n) + \frac{4}{5}y(n) \\ y(n+1) &= \frac{2}{5}x(n) + \frac{1}{5}y(n) \end{aligned}$$

O bien

$$\mathbf{X}(n+1) = \mathbf{B} \cdot \mathbf{X}(n) \quad (n = 0, 1, 2, \dots) \quad (3)$$

Donde

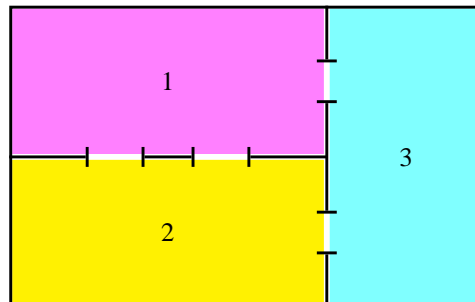
$$\mathbf{B} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix}$$

Y, al igual que en el ejemplo anterior, obtenemos que

$$\mathbf{X}(n) = \mathbf{B}^n \cdot \mathbf{X}(0) \quad (n = 1, 2, 3, \dots) \quad (4)$$

Lo que queremos calcular es $y(5)$ para lo cual se calcula $\mathbf{X}(5) = \mathbf{B}^5 \cdot \mathbf{X}(0)$. Para calcular la probabilidad a largo plazo de que un día sea frío o caluroso hay que calcular la potencia \mathbf{B}^n para valores grandes de n . Más adelante veremos cómo puede hacerse dicho cálculo con facilidad.

Ejemplo 3. En un laboratorio se coloca un grupo de ratones en una caja dividida en tres compartimentos comunicados y con la misma facilidad de acceso, como se indica en la figura.



Los compartimentos permanecen cerrados y se abren cada lunes. Cada semana todos los ratones cambian de compartimento y eligen al azar otro. Suponiendo que la distribución inicial del número de ratones en los compartimentos 1, 2 y 3 viene dada por $\mathbf{X}(0) = (x_1(0), x_2(0), x_3(0))^t$, queremos calcular la distribución de ratones en los distintos compartimentos cuando han pasado cuatro semanas así como su distribución a largo plazo.

Notemos por p_{ij} la probabilidad de que un ratón que está en el compartimento j pase al compartimento i . Tenemos que

$$p_{31} = p_{32} = \frac{1}{3}, \quad p_{21} = p_{12} = \frac{2}{3}, \quad p_{13} = p_{23} = \frac{1}{2}, \quad p_{11} = p_{22} = p_{33} = 0,$$

El número de ratones que hay en la semana $n+1$ en cada compartimento viene dado por

$$\begin{aligned} x_1(n+1) &= x_2(n)p_{12} + x_3(n)p_{13} = \frac{2}{3}x_2(n) + \frac{1}{2}x_3(n) \\ x_2(n+1) &= x_1(n)p_{21} + x_3(n)p_{23} = \frac{2}{3}x_1(n) + \frac{1}{2}x_3(n) \\ x_3(n+1) &= x_1(n)p_{31} + x_2(n)p_{32} = \frac{1}{3}x_1(n) + \frac{1}{3}x_2(n) \end{aligned}$$

que podemos escribir matricialmente en la forma

$$\mathbf{X}(n+1) = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{2} \\ \frac{2}{3} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix} \mathbf{X}(n) \quad (n = 0, 1, 2, \dots) \quad (5)$$

Llamando \mathbf{T} a la matriz que aparece en la igualdad anterior, resulta que la dinámica de la población está descrita por la ecuación matricial $\mathbf{X}(n+1) = \mathbf{T} \cdot \mathbf{X}(n)$ y, al igual que en los ejemplos anteriores, obtenemos que

$$\mathbf{X}(n) = \mathbf{T}^n \cdot \mathbf{X}(0) \quad (n = 1, 2, 3, \dots) \quad (6)$$

Lo que queremos calcular es $\mathbf{X}(4)$, que puede hacerse calculando la potencia \mathbf{T}^4 , y también $\mathbf{X}(n)$ para valores grandes de n , lo que más adelante veremos cómo puede hacerse con facilidad de forma aproximada.

Valores y vectores propios. Diagonalización de matrices

La dinámica de los procesos considerados en los ejemplos anteriores, así como la de otros que veremos más adelante, responde a una ecuación matricial del tipo

$$\mathbf{X}(n+1) = \mathbf{M} \cdot \mathbf{X}(n) \quad (n = 0, 1, 2, \dots) \quad (7)$$

Donde \mathbf{M} es una matriz cuadrada de orden q . Partiendo de un vector de estado inicial $\mathbf{X}(0)$, el vector de estado en cada etapa viene dado por

$$\mathbf{X}(n) = \mathbf{M}^n \cdot \mathbf{X}(0) \quad (n = 1, 2, 3, \dots) \quad (8)$$

Interesan por lo tanto métodos que permitan calcular con facilidad las potencias de una matriz o que nos informen del comportamiento de dichas potencias para valores grandes del exponente.

Para calcular las potencias de una matriz cuadrada, \mathbf{M} , de orden q , vamos a tratar de escribir dicha matriz de la forma

$$\mathbf{M} = \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} \quad (9)$$

donde \mathbf{P} y \mathbf{D} son matrices cuadradas de orden q , \mathbf{P} es **invertible** y \mathbf{D} es **diagonal**, es decir, los únicos elementos no nulos de \mathbf{D} están en la diagonal principal; si dichos elementos los representamos por $\lambda_1, \lambda_2, \dots, \lambda_q$, escribiremos $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_q)$.

Vamos a calcular las matrices \mathbf{P} y \mathbf{D} , cuando ello sea posible, y a estudiar el comportamiento de \mathbf{M}^n y de $\mathbf{X}(n) = \mathbf{M}^n \mathbf{X}(0)$ para valores grandes de n .

De la igualdad (9) se deduce

$$\begin{aligned} \mathbf{M}^2 &= \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} \cdot \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} = \mathbf{P} \cdot \mathbf{D}^2 \cdot \mathbf{P}^{-1} \\ \mathbf{M}^3 &= \mathbf{M}^2 \cdot \mathbf{M} = \mathbf{P} \cdot \mathbf{D}^2 \cdot \mathbf{P}^{-1} \cdot \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} = \mathbf{P} \cdot \mathbf{D}^3 \cdot \mathbf{P}^{-1} \end{aligned}$$

Y, en general

$$\mathbf{M}^n = \mathbf{P} \cdot \mathbf{D}^n \cdot \mathbf{P}^{-1} \quad (10)$$

Observa que el cálculo de \mathbf{D}^n es inmediato pues $\mathbf{D}^n = \text{diag}(\lambda_1^n, \lambda_2^n, \dots, \lambda_q^n)$.

Cuando existen matrices \mathbf{P} y \mathbf{D} en las condiciones indicadas que verifican la igualdad (9) se dice que la matriz \mathbf{M} es **diagonalizable**. Por tanto, si la matriz \mathbf{M} es diagonalizable y sabemos calcular las matrices \mathbf{P} y \mathbf{D} podemos calcular fácilmente sus potencias. Veamos cómo se calculan dichas matrices.

Supuesto que \mathbf{M} es diagonalizable, se debe verificar que $\mathbf{M} \cdot \mathbf{P} = \mathbf{P} \cdot \mathbf{D}$ y, llamando \mathbf{P}_k al vector columna k -ésima de la matriz \mathbf{P} , deducimos que $\mathbf{M} \cdot \mathbf{P}_k = \lambda_k \mathbf{P}_k$. Observa que $\mathbf{P}_k \neq \mathbf{0}$ porque la matriz \mathbf{P} es invertible.

Se dice que un número λ es un **valor propio** de la matriz \mathbf{M} cuando hay algún vector *no nulo* \mathbf{X} tal que $\mathbf{M} \cdot \mathbf{X} = \lambda \mathbf{X}$, en tal caso también se dice que \mathbf{X} es un **vector propio** asociado al valor propio λ .

Por tanto, *si la matriz \mathbf{M} es diagonalizable, las columnas de la matriz \mathbf{P} son vectores propios y los elementos de la diagonal de \mathbf{D} son valores propios de \mathbf{M} .*

Deducimos que **una condición necesaria y suficiente para que \mathbf{M} sea diagonalizable es que tenga q vectores propios linealmente independientes.**

Si λ es un valor propio de \mathbf{M} , entonces el sistema de ecuaciones lineales $\mathbf{M} \cdot \mathbf{X} = \lambda \mathbf{X}$, es decir $\mathbf{M} \cdot \mathbf{X} - \lambda \mathbf{X} = \mathbf{0}$, que podemos escribir en la forma $(\mathbf{M} - \lambda \mathbf{I}) \cdot \mathbf{X} = \mathbf{0}$, donde \mathbf{I} es la matriz identidad de orden q , debe tener soluciones no nulas y, como es un sistema homogéneo, el determinante de la matriz $\mathbf{M} - \lambda \mathbf{I}$ debe ser nulo. Recíprocamente, si esta condición se cumple entonces el sistema de ecuaciones lineales $(\mathbf{M} - \lambda \mathbf{I}) \cdot \mathbf{X} = \mathbf{0}$ tiene soluciones distintas de cero y, por tanto, λ es un valor propio de \mathbf{M} . En consecuencia, los valores propios de \mathbf{M} son las soluciones de la **ecuación característica** de \mathbf{M} :

$$\det(\mathbf{M} - \lambda \mathbf{I}) = 0$$

Se trata de una ecuación polinómica de grado q en la variable λ . El Teorema Fundamental del Álgebra nos dice que dicha ecuación tiene q soluciones (contando cada solución tantas veces como su multiplicidad) reales o complejas¹.

Como consecuencia de que el determinante de una matriz cuadrada es igual al determinante de la matriz transpuesta, se verifica que \mathbf{M}^t tiene la misma ecuación característica que \mathbf{M} y, por tanto, los mismos valores propios que \mathbf{M} con las mismas multiplicidades.

Si λ es un valor propio de \mathbf{M} entonces las soluciones del sistema de ecuaciones lineales

$$(\mathbf{M} - \lambda \mathbf{I}) \cdot \mathbf{X} = \mathbf{0}$$

son un subespacio vectorial no nulo que se llama el **espacio propio** asociado al valor propio λ . Se verifica el siguiente teorema.

Teorema 1. *Una matriz cuadrada \mathbf{M} es diagonalizable si, y sólo si, se verifica que la dimensión del espacio propio asociado a cada valor propio es igual a su multiplicidad algebraica.*

Como el espacio propio asociado a un valor propio simple es una recta vectorial cuya dimensión es uno, deducimos, como caso particular de este teorema, que *si todos los valores propios de una matriz son simples se verifica que dicha matriz es diagonalizable.*

Ejemplo 4. Consideremos la matriz

$$\mathbf{M} = \begin{pmatrix} -1 & 3 & 3 \\ 2 & 0 & 2 \\ 1 & -1 & -3 \end{pmatrix}$$

Para saber si es diagonalizable lo primero es calcular sus valores propios

$$\det(\mathbf{M} - \lambda \mathbf{I}) = 16 + 4\lambda - 4\lambda^2 - \lambda^3$$

La ecuación característica es $16 + 4\lambda - 4\lambda^2 - \lambda^3 = 0$. Como tiene coeficientes enteros y el coeficiente líder es 1, comprobamos si los divisores del término independiente son soluciones. Obtenemos fácilmente que $\lambda = 2$ es solución. Con lo cual ya es posible calcular las otras soluciones que resultan ser -4 y -2 . Hay tres valores propios distintos, por lo que la matriz es diagonalizable. Como los valores propios son simples, cada uno de ellos tiene un espacio propio que será de dimensión uno y estará engendrado por un vector propio. El espacio propio asociado al valor propio $\lambda = 2$ son las soluciones

¹Aunque los valores propios de una matriz real pueden ser números complejos, en los ejercicios de este tema solamente trabajaremos con números reales.

del sistema de ecuaciones lineal homogéneo $(\mathbf{M} - 2\mathbf{I}) \cdot \mathbf{X} = \mathbf{0}$

$$\begin{pmatrix} -3 & 3 & 3 \\ 2 & -2 & 2 \\ 1 & -1 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} -3x + 3y + 3z = 0 \\ 2x - 2y + 2z = 0 \\ x - y - 5z = 0 \end{cases}$$

Este sistema debe tener rango dos. Observa que la tercera fila es la opuesta de la suma de las otras dos por lo que podemos eliminarla y, una vez hecho eso, en el sistema de dos ecuaciones que nos queda el determinante de los coeficientes de y y de z es distinto de cero, por lo que hacemos $x = t$ y calculamos y y z en función de t obteniendo que las soluciones son los vectores de la forma $(t, t, 0)$ y haciendo $t = 1$ obtenemos el vector propio $(1, 1, 0)$. De forma análoga se calculan vectores propios para $\lambda = -4$ y para $\lambda = -2$, obteniendo, respectivamente los vectores $(-1, 0, 1)$ y $(0, -1, 1)$. Por tanto

$$\mathbf{P} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Debemos comprobar que $\mathbf{M} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ pero, para ahorrarnos el cálculo de \mathbf{P}^{-1} , es más cómodo comprobar la igualdad equivalente $\mathbf{M}\mathbf{P} = \mathbf{P}\mathbf{D}$.

$$\mathbf{M}\mathbf{P} = \begin{pmatrix} -1 & 3 & 3 \\ 2 & 0 & 2 \\ 1 & -1 & -3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 0 \\ 2 & 0 & 2 \\ 0 & -4 & -2 \end{pmatrix}$$

$$\mathbf{P}\mathbf{D} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 0 \\ 2 & 0 & 2 \\ 0 & -4 & -2 \end{pmatrix}$$

El siguiente resultado proporciona información del comportamiento de las potencias de una matriz para valores grandes del exponente.

Se dice que un valor propio λ de una matriz cuadrada \mathbf{M} es un **valor propio dominante** si para cualquier otro valor propio $\alpha \neq \lambda$ se verifica que $|\lambda| > |\alpha|$. Recuerda que representamos por $\langle \mathbf{X} | \mathbf{Y} \rangle$ el producto escalar de los vectores \mathbf{X} e \mathbf{Y} .

Teorema 2. Sea \mathbf{M} una matriz cuadrada de orden q que tiene un valor propio dominante λ que es simple. Sea \mathbf{P} un vector propio de \mathbf{M} y $\mathbf{Q} = (\beta_1, \beta_2, \dots, \beta_q)$ un vector propio de \mathbf{M}^t asociados al valor propio λ y tales que $\langle \mathbf{P} | \mathbf{Q} \rangle = 1$. Entonces se verifica que

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda^n} \mathbf{M}^n = (\beta_1 \mathbf{P} | \beta_2 \mathbf{P} | \dots | \beta_q \mathbf{P})^2 \tag{11}$$

Dado $\mathbf{X}(0) \in \mathbb{R}^q$, pongamos $\mathbf{X}(n) = \mathbf{M}^n \mathbf{X}(0)$. Se verifica que

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda^n} \mathbf{X}(n) = \alpha \mathbf{P} \tag{12}$$

donde $\alpha = \langle \mathbf{X}(0) | \mathbf{Q} \rangle$.

Puesto que no hemos dado una definición formal del significado de los símbolos que figuran a la izquierda en las igualdades (11) y (12), interpretaremos dichas igualdades en el sentido de que para valores grandes de n se verifica que

$$\frac{1}{\lambda^n} \mathbf{M}^n \approx (\beta_1 \mathbf{P} | \beta_2 \mathbf{P} | \dots | \beta_q \mathbf{P}), \quad \frac{1}{\lambda^n} \mathbf{X}(n) \approx \alpha \mathbf{P} \tag{13}$$

² $(\beta_1 \mathbf{P} | \beta_2 \mathbf{P} | \dots | \beta_q \mathbf{P})$ es la matriz cuyas columnas son los vectores $\beta_i \mathbf{P}$, $1 \leq i \leq q$.

Ejemplo 5. Sea

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 1 \\ -3 & 4 & 1 \\ -2 & 1 & 3 \end{pmatrix}$$

La ecuación característica $\det(\mathbf{M} - \lambda \mathbf{I}) = -\lambda^3 + 7\lambda^2 - 16\lambda + 12 = 0$ tiene dos raíces, $\lambda = 3$ simple, y $\lambda = 2$ doble. El espacio propio asociado al valor propio $\lambda = 2$ viene dado por las soluciones del sistema $(\mathbf{M} - 2\mathbf{I}) \cdot \mathbf{X} = \mathbf{0}$, esto es

$$\begin{aligned} -2x + y + z &= 0 \\ -3x + 2y + z &= 0 \\ -2x + y + z &= 0 \end{aligned}$$

Podemos eliminar la tercera ecuación y se obtiene fácilmente que las soluciones son los vectores de la forma (t, t, t) con $t \in \mathbb{R}$, es decir, se trata de una recta vectorial, por lo que la dimensión geométrica del valor propio $\lambda = 2$ es 1. Por tanto, la matriz no es diagonalizable. Pero como hay un valor propio dominante $\lambda = 3$ y simple, podemos calcular el límite $\lim_{n \rightarrow \infty} \frac{1}{3^n} \mathbf{M}^n$. Para ello debemos calcular vectores propios $\mathbf{P} = (x, y, z)$ y $\mathbf{Q} = (a, b, c)$ de \mathbf{M} y de \mathbf{M}^t respectivamente, asociados al valor propio $\lambda = 3$ y tales que $\langle \mathbf{P} | \mathbf{Q} \rangle = 1$. Dichos vectores deben verificar las ecuaciones lineales $(\mathbf{M} - 3\mathbf{I}) \cdot \mathbf{P} = \mathbf{0}$ y $(\mathbf{M}^t - 3\mathbf{I}) \cdot \mathbf{Q} = \mathbf{0}$, esto es

$$\begin{cases} -3x + y + z = 0 \\ -3x + y + z = 0 \\ -2x + y = 0 \end{cases} \quad \begin{cases} -3a - 3b - 2c = 0 \\ a + b + c = 0 \\ a + b = 0 \end{cases}$$

En el primer sistema podemos eliminar la primera ecuación y se obtiene fácilmente que las soluciones son los vectores de la forma $(t, 2t, t)$ con $t \in \mathbb{R}$, y las soluciones del segundo son de la forma $(t, -t, 0)$. Podemos tomar $\mathbf{P} = (1, 2, 1)$ y calculamos \mathbf{Q} por la condición $\langle (t, -t, 0) | (1, 2, 1) \rangle = 1$, es decir, $t = -1$, con lo que $\mathbf{Q} = (-1, 1, 0)$. Por tanto

$$\lim_{n \rightarrow \infty} \frac{1}{3^n} \mathbf{M}^n = \begin{pmatrix} -1 & 1 & 0 \\ -2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

Interpretamos esta igualdad en el sentido de que para valores grandes de n se verifica que

$$\frac{1}{3^n} \mathbf{M}^n \cong \begin{pmatrix} -1 & 1 & 0 \\ -2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

Con el programa *wx-Maxima* se obtiene que

$$\frac{1}{3^{30}} \mathbf{M}^{30} = \begin{pmatrix} -1,00007 & 0,999995 & 0,0000782264 \\ -2,00007 & 1,99999 & 0,0000782264 \\ -1,00007 & 0,999995 & 0,0000834415 \end{pmatrix}$$

Poblaciones estructuradas por grupos de edad. Modelo de Leslie

Vamos a estudiar un modelo, propuesto por Patrick Leslie (1900 - 1974), que describe la dinámica, en intervalos de tiempo de longitud k , de los individuos hembra de una población que están clasificados por edades en intervalos de la misma longitud k . El ejemplo 1 es típico de este modelo. La hipótesis básica para poder aplicar este modelo es que en la población estudiada el número de machos es un porcentaje fijo del de hembras, usualmente se supone que hay el mismo número de ambos. El modelo

se aplica a poblaciones cerradas (sin inmigración ni emigración) y no se consideran factores que limiten el crecimiento.

Consideremos, para fijar ideas, una población en la que las hembras tienen una esperanza máxima de vida de E años (podría ser otra unidad de tiempo) y las clasificamos en 4 grupos de edades correspondientes a los intervalos $I_k = [(k-1)E/4, kE/4]$, $k = 1, 2, 3, 4$ de forma que *cada hembra esté solamente en uno de los mismos* (recién nacidas, jóvenes, adultas y viejas).

En este modelo la distancia temporal entre dos recuentos consecutivos debe ser igual a la duración de los intervalos de edades; en este caso, los recuentos se harían cada $E/4$ años.

Llamando $x_k(n)$ al número de individuos en el grupo I_k en la etapa $n = 0, 1, 2, \dots$, para describir la evolución de la población hay que conocer las **tasas de fertilidad** f_k , $k = 1, 2, 3, 4$, que representan el promedio del número de hijas que tiene una hembra del grupo I_k durante el tiempo que permanece en este grupo, y las **tasas de supervivencia** s_k , $k = 1, 2, 3$, que representan la fracción de las hembras que están en el grupo I_k que pasarán al grupo I_{k+1} .

Representemos por $\mathbf{X}(n) = (x_1(n), x_2(n), x_3(n), x_4(n))^t$ el vector de distribución de las edades en la etapa n . El número de hembras en el grupo I_1 en la etapa $n+1$, es decir $x_1(n+1)$, será la suma de las hembras nacidas entre las etapas n y $n+1$, que viene dado por

$$x_1(n+1) = f_1x_1(n) + f_2x_2(n) + f_3x_3(n) + f_4x_4(n) \quad n = 0, 1, 2, \dots \quad (14)$$

El número de hembras en el grupo I_{k+1} , $k = 1, 2, 3$, en la etapa $n+1$ será igual al de hembras que estaban en I_k en la etapa n que han sobrevivido

$$x_{k+1}(n+1) = s_kx_k(n) \quad k = 1, 2, 3. \quad n = 0, 1, 2, \dots \quad (15)$$

Hemos tenido en cuenta que en cada etapa todos los individuos pasan de un grupo al siguiente, incluidos los viejos, que se supone que han muerto de uno a otro recuento. Podemos expresar matricialmente las ecuaciones (14) y (15) en la forma

$$\begin{pmatrix} x_1(n+1) \\ x_2(n+1) \\ x_3(n+1) \\ x_4(n+1) \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 & f_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{pmatrix} \begin{pmatrix} x_1(n) \\ x_2(n) \\ x_3(n) \\ x_4(n) \end{pmatrix} \quad n = 0, 1, 2, \dots$$

Y en forma vectorial

$$\mathbf{X}(n+1) = \mathbf{M} \cdot \mathbf{X}(n) \quad (n = 0, 1, 2, \dots) \quad (16)$$

La matriz

$$\mathbf{M} = \begin{pmatrix} f_1 & f_2 & f_3 & f_4 \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \end{pmatrix}$$

se llama **matriz de Leslie** y su característica principal es que sus coeficientes son nulos con la excepción de la primera fila cuyos elementos son todos mayores o igual que cero, $f_k \geq 0$, $k = 1, 2, 3, 4$, y los de la subdiagonal que son positivos $s_k > 0$, $k = 1, 2, 3$.

Suponiendo conocidas las poblaciones iniciales de cada grupo, $\mathbf{X}(0) = (x_1(0), x_2(0), x_3(0), x_4(0))^t$, tenemos que

$$\mathbf{X}(n) = \mathbf{M}^n \mathbf{X}(0) \quad (n = 1, 2, 3, \dots) \quad (17)$$

Por tanto, conociendo la distribución inicial de los grupos y la matriz de Leslie, podemos conocer la distribución de hembras en cada etapa y con ello el tamaño (machos y hembras) de cada grupo de la población.

La ecuación (17) es cómoda para estudiar la evolución a corto plazo de la población, basta para ello calcular las correspondientes potencias de la matriz de Leslie, pero también estamos interesados en estudiar la evolución a largo plazo de la población.

El siguiente resultado describe la evolución a largo plazo de una población modelada por una matriz de Leslie.

Teorema 3. *Sea \mathbf{M} una matriz de Leslie en cuya primera fila hay al menos dos elementos consecutivos estrictamente positivos³, entonces se verifica que \mathbf{M} tiene un único valor propio positivo, λ , que es simple y dominante, y tiene un vector propio asociado $\mathbf{P} = (p_1, p_2, \dots, p_q)$ con todas sus componentes positivas.*

A la vista de este teorema, y del Teorema 2, deducimos que

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda^n} \mathbf{X}(n) = \alpha \mathbf{P} \quad (18)$$

Donde $\alpha = \langle \mathbf{X}(0) | \mathbf{Q} \rangle$, y \mathbf{Q} es un vector propio de \mathbf{M}^t asociado al valor propio dominante λ tal que $\langle \mathbf{P} | \mathbf{Q} \rangle = 1$.

Vamos a analizar la información que proporciona la igualdad (18). Poniendo $\mathbf{P} = (p_1, p_2, \dots, p_q)$, dicha igualdad significa que

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda^n} (x_1(n), x_2(n), \dots, x_q(n)) = \alpha (p_1, p_2, \dots, p_q) \quad (19)$$

es decir

$$\lim_{n \rightarrow \infty} \frac{x_i(n)}{\lambda^n} = \alpha p_i \quad 1 \leq i \leq q \quad (20)$$

Por tanto, para valores grandes de n , se tiene que

$$x_i(n) \approx \lambda^n \alpha p_i, \quad x_i(n+1) \approx \lambda x_i(n), \quad \mathbf{X}(n+1) \approx \lambda \mathbf{X}(n) \quad (21)$$

Lo que nos dice que, para valores grandes de n , la razón de crecimiento de cada grupo poblacional, al igual que la razón de crecimiento de la población total, se estabilizan y son aproximadamente iguales a λ . Es decir, el número λ representa la razón con que aumentan o disminuyen todos los grupos de edad, así como la población total, al pasar de una etapa a la siguiente. En consecuencia, para valores grandes de n , se verifica que

- $\frac{x_i(n+1)}{x_i(n)} \approx \lambda$ para $1 \leq i \leq q$. Es decir, la **tasa de crecimiento** de cada grupo, expresada en tanto por uno, es la misma para todos ellos e igual a λ .
- $\frac{x_1(n+1) + x_2(n+1) + \dots + x_q(n+1)}{x_1(n) + x_2(n) + \dots + x_q(n)} \approx \lambda$. Es decir, la **tasa de crecimiento** de toda la población, expresada en tanto por uno, es igual a λ .
- $\frac{x_i(n+1) - x_i(n)}{x_i(n)} \approx \lambda - 1$ para $1 \leq i \leq q$. Es decir, la **tasa de crecimiento neta** de cada grupo, expresada en tanto por uno, es la misma para todos ellos e igual a $\lambda - 1$. Lo mismo sucede con la tasa de crecimiento neta de toda la población.

En consecuencia

- Si $0 < \lambda < 1$ la población terminará por extinguirse.
- Si $\lambda > 1$ la población aumentará indefinidamente y tendrá una tasa neta de crecimiento igual a $100(\lambda - 1)$ expresada en tanto por ciento.
- Si $\lambda = 1$ la población alcanzará un equilibrio y a largo plazo $\mathbf{X}(n)$ se mantendrá constante aproximadamente igual a $\alpha \mathbf{P}$.

³Esto siempre puede conseguirse haciendo los intervalos de edad suficientemente pequeños.

Deducimos también de las igualdades (20) que

$$\lim_{n \rightarrow \infty} \frac{x_1(n) + x_2(n) + \dots + x_q(n)}{\lambda^n} = \alpha(p_1 + p_2 + \dots + p_q) \quad (22)$$

Y, haciendo el cociente de los límites en (20) y (22), obtenemos que

$$\lim_{n \rightarrow \infty} \frac{x_i(n)}{x_1(n) + x_2(n) + \dots + x_q(n)} = \frac{p_i}{p_1 + p_2 + \dots + p_q} \quad (23)$$

Por tanto

$$\frac{x_i(n)}{x_1(n) + x_2(n) + \dots + x_q(n)} \approx \frac{p_i}{p_1 + p_2 + \dots + p_q} \quad (24)$$

Es decir, el vector \mathbf{P} nos informa del tamaño relativo que **a largo plazo** tendrá cada grupo respecto al total de la población. Es decir, el vector de proporciones relativas de la población se aproxima al vector de poblaciones relativas dado por \mathbf{P} .

Por tanto, para estudiar la dinámica de una población que sigue un modelo de Leslie en el que hay dos tasas de fertilidad consecutivas positivas, todo lo que necesitamos es calcular el único valor propio positivo, λ , de la matriz y un vector propio asociado al mismo, \mathbf{P} , con componentes positivas. Si además queremos una estimación aproximada de la distribución de la población para valores grandes de n necesitamos calcular el número $\alpha = \langle \mathbf{X}(0) | \mathbf{Q} \rangle$, donde \mathbf{Q} es un vector propio de \mathbf{M}^t asociado al valor propio λ tal que $\langle \mathbf{P} | \mathbf{Q} \rangle = 1$, pues sabemos que, para valores grandes de n , $\mathbf{X}(n) \approx \lambda^n \alpha \mathbf{P}$.

Ejemplo 1 de nuevo. La matriz \mathbf{A} del ejemplo 1 es una matriz de Leslie que cumple las condiciones del teorema 3 y tiene como ecuación característica

$$\det(\mathbf{A} - \lambda \mathbf{I}) = -\lambda^3 + \frac{1}{2}\lambda + \frac{1}{2}$$

que, evidentemente, tiene la raíz $\lambda = 1$. El teorema 3 nos dice que dicho valor propio es el único positivo y es simple y dominante (los otros dos valores propios, como puedes comprobar, son complejos). Un vector propio asociado al mismo con componentes positivas es $\mathbf{P} = (12, 3, 2)$. Por tanto, podemos afirmar que, a largo plazo, la población se estabilizará y las proporciones relativas de cada grupo respecto al tamaño total de la población viene dadas por el vector $(12/17, 3/17, 2/17)$.

Los vectores propios de la matriz \mathbf{A}^t asociados al valor propio $\lambda = 1$ son de la forma $(t, 4t, 3t)$ con $t \neq 0$. Imponiendo la condición $\langle (12, 3, 2) | (t, 4t, 3t) \rangle = 30t = 1$, obtenemos $t = 1/30$. Por tanto $\mathbf{Q} = (1/30, 2/15, 1/10)$ y, como $\mathbf{X}(0) = (200, 100, 80)$, obtenemos que $\alpha = \langle \mathbf{X}(0) | \mathbf{Q} \rangle = 28$. Por tanto, para valores grandes de n , tenemos que $\mathbf{X}(n) \approx 28(12, 3, 2) = (336, 84, 56)$. Por tanto, la población total se estabilizaría en torno a 476 hembras, de las cuales 336 serían crías, 84 jóvenes y 56 adultas.

El teorema 2 también nos dice que

$$\lim_{n \rightarrow \infty} \mathbf{A}^n = \left(\frac{1}{30} \mathbf{P} \mid \frac{2}{15} \mathbf{P} \mid \frac{1}{10} \mathbf{P} \right) = \begin{pmatrix} \frac{2}{5} & \frac{8}{5} & \frac{6}{5} \\ \frac{1}{10} & \frac{2}{5} & \frac{3}{10} \\ \frac{1}{15} & \frac{4}{15} & \frac{1}{5} \end{pmatrix}$$

Naturalmente, se cumple que

$$\begin{pmatrix} \frac{2}{5} & \frac{8}{5} & \frac{6}{5} \\ \frac{1}{10} & \frac{2}{5} & \frac{3}{10} \\ \frac{1}{15} & \frac{4}{15} & \frac{1}{5} \end{pmatrix} \cdot \begin{pmatrix} 200 \\ 100 \\ 80 \end{pmatrix} = \begin{pmatrix} 336 \\ 84 \\ 56 \end{pmatrix}$$

Ejemplo 6. Dividimos las hembras de una población animal, cuya edad máxima es de 15 años, en tres grupos “crías”, “jóvenes” y “adultos”. Las crías no se reproducen, cada hembra joven tiene por promedio 4 crías y las adultas 3. La mitad de las crías pasan a jóvenes y la cuarta parte de jóvenes pasan a adultas. La matriz de Leslie es

$$\mathbf{L} = \begin{pmatrix} 0 & 4 & 3 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{pmatrix}$$

Se supone una distribución inicial de 1000 hembras por grupo. Como la población debe dividirse en grupos de edad de la misma duración, las crías son las hembras menores de 5 años, las jóvenes son las que tienen al menos 5 años y menos de 10 y las adultas las que tienen al menos 10 años hasta 15. Los recuentos deben hacerse cada 5 años. Representando por $\mathbf{X}(n) = (x(n), y(n), z(n))$ la distribución de la población en la etapa n -ésima (cuando han pasado $5n$ años). Tenemos que $\mathbf{X}(n) = \mathbf{L}^n \mathbf{X}(0)$.

Si, por ejemplo, queremos saber la distribución de la población pasados 15 años debemos calcular $\mathbf{X}(3) = \mathbf{L}^3 \mathbf{X}(0)$.

Para estudiar la evolución a largo plazo, como hay dos tasas de fertilidad positivas consecutivas, podemos aplicar los resultados que acabamos de ver. Estos nos dicen que hay un único valor propio positivo, en este caso dicho valor es $\lambda = \frac{3}{2}$, que es dominante al que podemos asociar un vector propio con componentes positivas, que calculándolo, resulta ser (salvo múltiplos positivos) $\mathbf{P} = (18, 6, 1)$. Por tanto, a largo plazo, la población aumenta siendo su tasa neta de crecimiento el 50% pues $\mathbf{X}(n+1) \approx \frac{3}{2} \mathbf{X}(n) = (1+0,5) \mathbf{X}(n)$, además, a largo plazo, cada grupo tendrá esa misma tasa neta de crecimiento. El vector $(18, 6, 1)$ indica que las proporciones, siempre a largo plazo, de cada grupo respecto del total son respectivamente $\frac{18}{25} = 0,72$, $\frac{6}{25} = 0,24$ y $\frac{1}{25} = 0,04$, por tanto un 72% de las hembras serán crías, el 24% jóvenes y el 4% adultas.

Sabemos también que $\mathbf{X}(n) \approx \alpha \left(\frac{3}{2}\right)^n (18, 6, 1)$ donde $\alpha = \langle (1000, 1000, 1000) | \mathbf{Q} \rangle$ siendo \mathbf{Q} un vector propio de \mathbf{L}^t asociado al valor propio $\lambda = \frac{3}{2}$ y que verifica $\langle (18, 6, 1) | \mathbf{Q} \rangle = 1$. Poniendo $\mathbf{Q} = (a, b, c)$ debe verificarse que $\mathbf{L}^t(a, b, c)^t = \frac{3}{2}(a, b, c)^t$, esto es

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 4 & 0 & \frac{1}{4} \\ 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{3}{2} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Resolviendo obtenemos que \mathbf{Q} es de la forma $(t/2, 3t/2, t)$, la condición $\langle (18, 6, 1) | (t/2, 3t/2, t) \rangle = 1$ nos da $t = \frac{1}{19}$. Por tanto $\mathbf{Q} = (\frac{1}{38}, \frac{3}{38}, \frac{1}{19})$, y $\alpha = \frac{6000}{38}$. Por tanto, para valores grandes de n , se verifica que $\mathbf{X}(n) \approx \frac{6000}{38} \left(\frac{3}{2}\right)^n (18, 6, 1)$.

Poblaciones estructuradas en estados

Consideraremos ahora poblaciones aisladas constantes, sin emigración ni inmigración, clasificadas en grupos por criterios distintos a la edad. La diferencia principal con el modelo de Leslie es que ahora algunos individuos pueden permanecer en el mismo grupo o estado en diferentes etapas o recuentos. El ejemplo 3 es un caso típico.

Ejemplo 3 de nuevo. Aunque la matriz \mathbf{T} de dicho ejemplo

$$\mathbf{T} = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{2} \\ \frac{2}{3} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

no es una matriz de Leslie, sus valores propios son $\lambda_1 = 1$, $\lambda_2 = -\frac{2}{3}$ y $\lambda_3 = -\frac{1}{3}$. Por lo que $\lambda_1 = 1$ es un valor propio dominante y simple. En consecuencia sabemos, por (12), que

$$\lim_{n \rightarrow \infty} \mathbf{X}(n) = \alpha \mathbf{P}$$

donde \mathbf{P} es un vector propio asociado al valor propio $\lambda_1 = 1$. Un sencillo cálculo nos da que $\mathbf{P} = (3, 3, 2)$, luego dicho vector tiene sus componentes positivas. Por tanto, aunque \mathbf{T} no es una matriz de Leslie, sí se cumple la tesis del Teorema 3 por lo que también son válidas las consecuencias que se deducen de las mismas. Por tanto, el vector \mathbf{P} nos informa del tamaño relativo que, a largo plazo, tendrá cada grupo respecto al total de la población. Deducimos que, a largo plazo, $\frac{3}{8}$ del total de los ratones estarán en los compartimentos 1 y 2, y $\frac{1}{4}$ de los ratones estarán en el compartimento 3.

Observa que este resultado es independiente del número inicial $\mathbf{X}(0) = (x_1(0), x_2(0), x_3(0))^t$ de ratones que hay en cada compartimento y sólo depende del total de la población inicial. De hecho, como la población total permanece constante, debe verificarse que $x_1(0) + x_2(0) + x_3(0) = 8\alpha$, luego $\alpha = \frac{x_1(0)+x_2(0)+x_3(0)}{8}$ sólo depende del total de la población pero no de su distribución inicial.

También sabemos por, (11), que

$$\lim_{n \rightarrow \infty} \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{2} \\ \frac{2}{3} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}^n = \left(a\mathbf{P} \mid b\mathbf{P} \mid c\mathbf{P} \right)$$

Donde $\mathbf{Q} = (a, b, c)$ es un vector propio de la matriz \mathbf{T}^t asociado al valor propio $\lambda_1 = 1$. Como las filas de la matriz transpuesta \mathbf{T}^t suman 1, se tiene que $\mathbf{Q} = (t, t, t)$ y como debe cumplirse que $3t + 3t + 2t = 1$ se sigue que $t = \frac{1}{8}$. Por tanto

$$\lim_{n \rightarrow \infty} \mathbf{T}^n = \begin{pmatrix} \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \\ \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

Matrices positivas y grafos. Teorema de Perron–Frobenius

Se dice que una matriz \mathbf{M} es **no negativa**, y escribiremos $\mathbf{M} \geq 0$, si todos sus elementos son mayores o iguales que 0, y se dice que una matriz \mathbf{M} es **positiva**, y escribiremos $\mathbf{M} > 0$, si todos sus elementos son estrictamente mayores que 0.

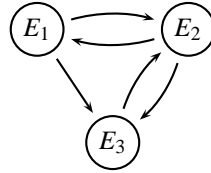
Grafos dirigidos.

Un **grafo dirigido** es un par $G = (E, F)$ formado por un conjunto E de **nodos** o **vértices**, y un conjunto $F \subset E \times E$ cuyos elementos se llaman **segmentos dirigidos** o **flechas** o **arcos** que unen algunos vértices con otros. Si un par $(E_i, E_j) \in F$, eso significa que se puede pasar directamente del nodo E_i al nodo E_j y gráficamente lo representamos como una flecha o arco que sale de E_i y llega a E_j . Un **camino** de E_i a E_j es una sucesión finita de arcos que pasan de E_i a E_j , el número de arcos que forman el camino se llama **longitud** del camino. Un camino que empieza y termina en el mismo nodo se llama un **ciclo**. Se dice que un grafo es **fuertemente conexo** si para cada par de nodos (E_i, E_j) con $1 \leq i, j \leq q$, hay un camino que va de E_i a E_j .

Matriz de incidencia de un grafo dirigido.

A un grafo dirigido $G = (E, F)$ podemos asociar una matriz cuadrada, $\mathbf{M}_G = (m_{ij})$, llamada **matriz de incidencia**, cuyo orden es igual al número de nodos del grafo y $m_{ij} = 1$ cuando $(E_j, E_i) \in F$, es decir, hay una flecha en el grafo que va de E_j a E_i , y $m_{ij} = 0$ en otro caso. Un 1 en el lugar (i, j) de la matriz de incidencia refleja que el nodo i es accesible *directamente* desde el nodo j . Un valor $k > 0$ en el lugar (i, j) de la potencia n -ésima de dicha matriz indica que hay k caminos distintos de longitud n que van del nodo j al nodo i . Un valor $k = 0$ en el lugar (i, j) de la potencia n -ésima de dicha matriz indica que no hay ningún camino de longitud n que vaya del nodo j al nodo i . Por tanto, el grafo es fuertemente conexo, si para cada par (i, j) hay algún n tal que $[(\mathbf{M}_G)^n](i, j) > 0$, donde hemos notado $[(\mathbf{M}_G)^n](i, j)$ el elemento (i, j) en la potencia n -ésima de la matriz de incidencia del grafo.

Ejemplo 7. Consideremos el siguiente grafo ordenado



Cuya matriz de incidencia es

$$\mathbf{M} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Tenemos que

$$\mathbf{M}^2 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

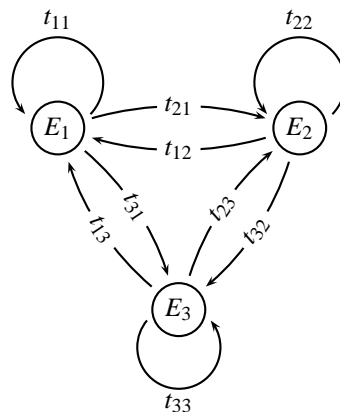
La primera columna de esta matriz nos dice que desde E_1 podemos alcanzar los tres estados por caminos de longitud 2 ($E_1 \rightarrow E_2 \rightarrow E_3$, $E_1 \rightarrow E_3 \rightarrow E_2$, $E_1 \rightarrow E_2 \rightarrow E_1$). El primer elemento en la segunda columna nos dice que desde E_2 no podemos ir a E_1 por un camino de longitud 2, el 2 en la segunda columna nos dice que hay dos caminos de longitud 2 que van de E_2 a E_2 ($E_2 \rightarrow E_1 \rightarrow E_2$ y $E_2 \rightarrow E_3 \rightarrow E_2$). Análoga interpretación tienen los restantes elementos de dicha matriz.

A toda matriz cuadrada no negativa $\mathbf{T} = (t_{ij}) \in \mathcal{M}_{q \times q}$ podemos asociar un grafo dirigido, $G_{\mathbf{T}}$, cuyo conjunto de nodos está formado por los q estados E_i que representan las filas de la matriz, y el estado E_j está unido por una flecha hacia al estado E_i , $E_j \rightarrow E_i$, si, y sólo si, $t_{ij} > 0$. Se dice que una matriz cuadrada no negativa es **irreducible** cuando su grafo asociado es fuertemente conexo.

Ejemplo 8. A una matriz cuadrada de orden 3

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix}$$

asociamos el siguiente grafo dirigido

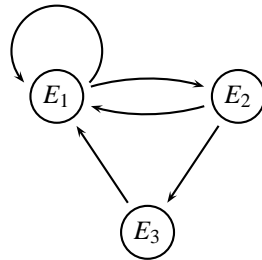


Se entiende que en este grafo **solamente deben aparecer las flechas para las que $t_{ij} > 0$.**

Ejemplo 9. La matriz

$$\mathbf{T} = \begin{pmatrix} 0,7 & 0,2 & 1 \\ 0,3 & 0 & 0 \\ 0 & 0,8 & 0 \end{pmatrix}$$

tiene el siguiente grafo asociado



Como partiendo de un estado cualquiera puede llegarse a cualquier estado, se trata de un grafo fuertemente conexo por lo que la matriz es irreducible. Podemos expresar esto como sigue.

$$\begin{aligned}
 E_1 &\rightarrow E_1 + E_2 \rightarrow E_1 + E_2 + E_3 \\
 E_2 &\rightarrow E_1 + E_3 \rightarrow E_1 + E_2 + E_3 \\
 E_3 &\rightarrow E_1 \rightarrow E_1 + E_2 \rightarrow E_1 + E_2 + E_3
 \end{aligned}$$

Cuya interpretación es la siguiente, $E_1 \rightarrow E_1 + E_2$ indica que desde E_1 podemos alcanzar en un solo paso el propio E_1 y E_2 , y como desde E_2 se puede alcanzar en un paso E_3 , escribimos $E_1 \rightarrow E_1 + E_2 \rightarrow E_1 + E_2 + E_3$ lo que nos dice que desde E_1 se pueden alcanzar los tres estados en dos pasos. Análoga interpretación tiene la segunda línea. Para la tercera línea $E_3 \rightarrow E_1$ nos dice que desde E_3 en un paso solamente puede alcanzarse E_1 , y ya encadenamos con la primera línea resultando $E_3 \rightarrow E_1 \rightarrow E_1 + E_2 \rightarrow E_1 + E_2 + E_3$ que nos dice que desde E_3 se pueden alcanzar los tres estados en 3 pasos. Se verifica que

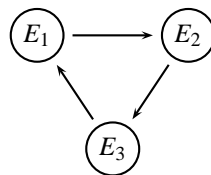
$$\mathbf{T}^3 = \begin{pmatrix} 0,667 & 0,67 & 0,55 \\ 0,165 & 0,282 & 0,21 \\ 0,168 & 0,048 & 0,24 \end{pmatrix}$$

es una matriz positiva.

Esto no pasa siempre. Consideremos la matriz

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 6 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}$$

Cuyo grafo asociado es



Dicha matriz es irreducible pero ninguna potencia suya es positiva pues

$$\mathbf{M}^2 = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ \frac{1}{6} & 0 & 0 \end{pmatrix}$$

y $\mathbf{M}^3 = \mathbf{I}$ es la matriz identidad de orden 3, por lo que cualquier potencia de \mathbf{M} o bien es igual a \mathbf{M} o a \mathbf{M}^2 o a la identidad, en cualquier caso ninguna potencia de \mathbf{M} tiene todos sus elementos estrictamente positivos.

Se dice que una matriz cuadrada no negativa es una **matriz primitiva** si alguna potencia suya es positiva. Si en una matriz cuadrada no negativa, \mathbf{M} , sustituimos sus elementos positivos por unos

obtenemos una matriz $\tilde{\mathbf{M}}$ que es la matriz de incidencia del grafo asociado a la matriz \mathbf{M} . Que \mathbf{M} sea primitiva equivale a que lo sea $\tilde{\mathbf{M}}$, pero eso implica que hay un número n tal que en el grafo asociado a la matriz puede irse de un nodo a cualquier nodo por un camino de longitud n . Por tanto, toda matriz primitiva es irreducible pero, según acabamos de ver, el recíproco no tiene por qué ser cierto. Para saber si una matriz es primitiva se utilizan los siguientes criterios.

Teorema 4. *Sea \mathbf{T} una matriz cuadrada no negativa irreducible. Si se verifica que algún elemento en la diagonal principal de \mathbf{T} es estrictamente positivo, o si \mathbf{T} tiene un valor propio dominante, entonces \mathbf{T} es primitiva.*

El siguiente teorema es uno de los más útiles de la teoría de matrices no negativas.

Teorema de Perron–Frobenius. *Sea \mathbf{T} una matriz cuadrada primitiva. Entonces se verifica que \mathbf{T} tiene un valor propio positivo, $\lambda > 0$, que es simple y dominante, dicho valor propio tiene asociado un vector propio con componentes positivas y es el único valor propio que tiene dicha propiedad.*

Por tanto, lo establecido en el teorema 2 puede aplicarse a matrices primitivas.

Cadenas de Markov y matrices de probabilidad.

Supongamos que un proceso tiene q resultados posibles, o estados, $\{E_1, E_2, \dots, E_q\}$. El proceso se repite de forma que su evolución solamente depende de su estado actual, es decir, es un proceso “sin memoria”. Sea $p_{ij} = P(E_i|E_j)$ la probabilidad de que se produzca el paso del estado E_j al E_i . Estas probabilidades se llaman **probabilidades de transición**, y la matriz $\mathbf{T} = (p_{ij})$ se llama **matriz de transición**. Como cada vez que se repite el proceso algunos de los estados debe ocurrir, se verifica que

$$p_{1j} + p_{2j} + \dots + p_{qj} = 1 \quad (1 \leq j \leq q)$$

Representemos por $p_i(n)$ la probabilidad de que en la repetición n -ésima del proceso estemos en el estado E_i ($1 \leq i \leq q$). Como alguno de los estados debe ocurrir en dicha etapa n , tenemos que

$$p_1(n) + p_2(n) + \dots + p_q(n) = 1 \quad (n = 1, 2, \dots)$$

La probabilidad $p_i(n+1)$ de obtener E_i en la etapa $n+1$ viene dada por

$$p_i(n+1) = p_{i1}p_1(n) + p_{i2}p_2(n) + \dots + p_{iq}p_q(n) \quad (1 \leq i \leq q)$$

Por tanto, definiendo $\mathbf{P}(n) = (p_1(n), p_2(n), \dots, p_q(n))^t$ y $\mathbf{T} = (p_{ij})$, se verifica que

$$\mathbf{P}(n+1) = \mathbf{T} \cdot \mathbf{P}(n) \quad n = 0, 1, 2, \dots$$

Algo con lo que ya debes estar familiarizado y que en el contexto probabilístico que estamos considerando recibe el nombre de **Cadena de Markov**. El ejemplo 2 es un caso típico de cadena de Markov.

Cada fila de \mathbf{T} se corresponde con un estado E_i . Observa que *las entradas positivas de la columna j indican los estados que pueden alcanzarse directamente desde el estado j* . Los elementos de la matriz \mathbf{T} son números mayores o iguales que 0 y la suma de cada columna de la matriz \mathbf{T} es igual a 1. Una matriz que cumple estas condiciones se dice que es una **matriz estocástica o una matriz de probabilidad (por columnas)** como, por ejemplo, las consideradas en los ejemplos 2 y 3. Una tal matriz *siempre tiene como valor propio a $\lambda = 1$* . Para convencerte, considera la matriz $\mathbf{T} - \mathbf{I}$ y suma a la primera fila todas las demás, con ello en la primera fila solamente hay ceros por lo que su determinante es nulo.

El siguiente resultado, que se deduce con facilidad del teorema 2 y del teorema de Perron–Frobenius, describe el comportamiento a largo plazo de una cadena de Markov.

Teorema 5. *Sea \mathbf{T} una matriz de probabilidad primitiva. Entonces $\lambda = 1$ es un valor propio simple dominante de \mathbf{T} que tiene un único vector propio asociado con coordenadas positivas $\mathbf{P} = (p_1, p_2, \dots, p_q)$ tal que $p_1 + p_2 + \dots + p_q = 1$. Se verifica que*

$$\lim_{n \rightarrow \infty} \mathbf{T}^n = (\mathbf{P}|\mathbf{P}| \dots |\mathbf{P}|) \quad (25)$$

Para $\mathbf{X}(0) \in \mathbb{R}^q$, poniendo $\mathbf{X}(n) = \mathbf{T}^n \mathbf{X}(0)$, se verifica que

$$\lim_{n \rightarrow \infty} \mathbf{T}^n \mathbf{X}(0) = \lim_{n \rightarrow \infty} \mathbf{X}(n) = (x_1(0) + x_2(0) + \dots + x_q(0)) \mathbf{P} \quad (26)$$

El vector \mathbf{P} se llama **distribución de estado estacionario**, sus componentes representan la probabilidad de encontrarse en cada uno de los estados a largo plazo.

Ejemplo 2 de nuevo. La matriz de este ejemplo

$$\mathbf{B} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix}$$

es una matriz de probabilidad primitiva. Por tanto, sabemos que $\lambda = 1$ es un valor propio dominante. Fácilmente se calcula que $(2, 1)$ es un vector propio asociado a dicho valor propio. Normalizamos dicho vector de forma que sus coordenadas sumen 1 y obtenemos $\mathbf{P} = (2/3, 1/3)$, lo que nos dice que, a largo plazo, la probabilidad de que un día sea caluroso es de $2/3$ y de que sea frío es de $1/3$. Se verifica también que

$$\lim_{n \rightarrow \infty} \mathbf{B}^n = \begin{pmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Ejemplo 10. Supongamos que en cierto país tres partidos políticos A, B y C se presentan cada año a las elecciones. Se dispone de los siguientes datos

| | A | B | C |
|-----|-----|-----|-----|
| A | 0,6 | 0,5 | 0,3 |
| B | 0,2 | 0,3 | 0,2 |
| C | 0,2 | 0,2 | 0,5 |

Cada columna representa un partido e indica cómo votarán en la próxima convocatoria los votantes de dicho partido; por ejemplo, la columna primera se lee: el 60% de los votantes del partido A volverán a votarlo, el 20% de los mismos votarán al partido B y el restante 20% al partido C . Se supone que el número total de votantes permanece constante cada año. Supongamos que $\mathbf{X}(n) = (x_1(n), x_2(n), x_3(n))^t$ representa los porcentajes de votos obtenidos en el año n por los partidos A, B y C respectivamente. Tenemos entonces que

$$\begin{aligned} x_1(n+1) &= 0,6x_1(n) + 0,5x_2(n) + 0,3x_3(n) \\ x_2(n+1) &= 0,2x_1(n) + 0,3x_2(n) + 0,2x_3(n) \\ x_3(n+1) &= 0,2x_1(n) + 0,2x_2(n) + 0,5x_3(n) \end{aligned}$$

Que podemos escribir

$$\mathbf{X}(n+1) = \mathbf{M}\mathbf{X}(n) \quad \text{donde} \quad \mathbf{M} = \begin{pmatrix} 0,6 & 0,5 & 0,3 \\ 0,2 & 0,3 & 0,2 \\ 0,2 & 0,2 & 0,5 \end{pmatrix} \quad (27)$$

La matriz \mathbf{M} tiene $\lambda = 1$ como valor propio dominante con vector propio asociado *normalizado* (es decir, con componentes positivas que suman 1) el vector $(0,492, 0,222, 0,286)$ que nos da los porcentajes de votantes a cada uno de los partidos a largo plazo si se mantienen las condiciones invariables.

También se verifica que

$$\lim_{n \rightarrow \infty} \mathbf{M}^n = \begin{pmatrix} 0,492 & 0,492 & 0,492 \\ 0,222 & 0,222 & 0,222 \\ 0,286 & 0,286 & 0,286 \end{pmatrix}$$

una matriz cuyas columnas son iguales al vector propio normalizado asociado al valor propio dominante.

El ejemplo 3 de nuevo. El ejemplo 3 es otro ejemplo de un proceso estocástico. Observa que la matriz de dicho ejemplo es una matriz de probabilidad, y se comprueba fácilmente que es irreducible; además $\lambda = 1$ es un valor propio dominante por lo que, en virtud del teorema 4, dicha matriz es primitiva y, por tanto, podemos usar el teorema 5 para obtener de manera más directa los resultados que ya vimos en su momento.

Herencia autosómica

Vamos a analizar matemáticamente cómo se transmite un carácter no ligado al sexo de una generación a la siguiente. Conviene recordar unos conceptos.

Un gen, la unidad básica de la herencia, es un segmento de DNA que contiene toda la información necesaria para sintetizar una proteína. Los seres humanos tenemos unos 20000 a 23000 genes, contenidos en los cromosomas dentro del núcleo de la célula. Las células somáticas (no reproductoras), tienen 46 cromosomas en 23 pares. Cada par consiste en un cromosoma de la madre y otro del padre. Veintidós de los pares, los *autosomas*, son homólogos (idénticos en tamaño, forma y posición y número de genes). El par 23, los *cromosomas sexuales* (X e Y), determina el sexo de una persona. El par 23 en las mujeres consta de dos cromosomas X y en los hombres de un cromosoma X y otro Y.

Los genes están dispuestos en forma lineal a lo largo del DNA de los cromosomas. Cada gen tiene una localización específica llamada *locus*. Los genes que ocupan el mismo locus en cada cromosoma de un par se denominan *alelos*. Cada individuo hereda un alelo del padre y otro de la madre para el mismo gen. Cada gen consta de una secuencia específica de DNA; dos alelos pueden tener secuencias de DNA levemente distintas lo que se puede manifestar en modificaciones concretas de la función de ese gen (producen variaciones en características heredadas como, por ejemplo, el color de ojos, de cabello o el grupo sanguíneo). La forma en que se expresa un gen depende de la combinación de alelos heredada.

Consideremos dos alelos, que representaremos con las letras A y a , de un gen autosómico, es decir, que codifican caracteres no ligados al sexo. En tal caso, los individuos de la población, con independencia del sexo, poseen dos alelos y pueden darse las siguientes combinaciones AA , Aa o aa . Cada una de ellas se llama un **genotipo** y determina la forma en que se expresarán los caracteres ligados a los alelos. Los posibles emparejamientos que pueden darse son

$$AA \times AA, \quad AA \times Aa, \quad AA \times aa, \quad Aa \times Aa, \quad Aa \times aa, \quad aa \times aa$$

En la siguiente tabla figuran las probabilidades de los distintos genotipos para los descendientes.

| | $AA \times AA$ | $AA \times Aa$ | $AA \times aa$ | $Aa \times Aa$ | $Aa \times aa$ | $aa \times aa$ |
|------|----------------|----------------|----------------|----------------|----------------|----------------|
| AA | 1 | 1/2 | 0 | 1/4 | 0 | 0 |
| Aa | 0 | 1/2 | 1 | 1/2 | 1/2 | 0 |
| aa | 0 | 0 | 0 | 1/4 | 1/2 | 1 |

Un alelo A se dice **dominante** frente a otro a cuando su sola presencia hace que se exprese el carácter al que va asociado y se dice **recesivo** en otro caso. Por ejemplo si A es tener ojos castaños y a tener ojos azules, los genotipos AA y Aa tienen ojos castaños y aa azules: el alelo A es dominante. Suele reservarse la letra A para los alelos dominantes.

Ejemplo 11.⁴ En un gran cultivo de plantas con genotipos AA , Aa y aa se va a iniciar un proceso de polinización con plantas del genotipo AA . Queremos estudiar la evolución de los tres genotipos de la población en las sucesivas generaciones.

⁴Tomado del texto de los profesores J.Navas, F.J. Esteban y J.M. Quesada *Modelos Matemáticos en Biología*. Universidad de Jaén 2009, que puede descargarse en http://matema.ujaen.es/jnavas/web_modelos/pdf_mmb08_09/texto%20completo.pdf.

Sea $\mathbf{X}(n) = (x_1(n), x_2(n), x_3(n))^t$ el vector que representa la proporción de plantas de los genotipos AA , Aa y aa respectivamente en la generación n , siendo $\mathbf{X}(0)$ las proporciones iniciales. Las tres primeras filas de la tabla anterior permiten obtener $\mathbf{X}(n+1)$ en función de $\mathbf{X}(n)$

$$\begin{aligned}x_1(n+1) &= x_1(n) + \frac{1}{2}x_2(n) \\x_2(n+1) &= \frac{1}{2}x_2(n) + x_3(n) \\x_3(n+1) &= 0\end{aligned}$$

ecuaciones que podemos escribir matricialmente $\mathbf{X}(n+1) = \mathbf{TX}(n)$, $n = 0, 1, 2, \dots$. Siendo

$$\mathbf{T} = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Se trata de una matriz de probabilidad y podemos considerar que es la matriz de transición de un proceso de Markov con tres posibles estados $E_1 = AA$, $E_2 = Aa$ y $E_3 = aa$. Observa que, evidentemente, no es una matriz regular porque el estado $E_3 = aa$ no es accesible. Naturalmente, como $\mathbf{X}(n) = \mathbf{T}^n \mathbf{X}(0)$, necesitamos calcular \mathbf{T}^n . Se calculan fácilmente los valores propios de \mathbf{T} que son $\lambda = 1$, $\lambda_2 = \frac{1}{2}$ y $\lambda_3 = 0$. Al ser distintos, la matriz es diagonalizable. Como vectores propios asociados podemos tomar

$$\mathbf{P}_1 = (1, 0, 0)^t, \quad \mathbf{P}_2 = (1, -1, 0)^t, \quad \mathbf{P}_3 = (1, -2, 1)^t.$$

Observa que el mero hecho de que \mathbf{P}_1 no tenga todas sus coordenadas positivas confirma que \mathbf{T} no es primitiva. Tenemos que $\mathbf{T} = \mathbf{PDP}^{-1}$ con

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Luego

$$\mathbf{T}^n = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 - \frac{1}{2^n} & 1 - \frac{1}{2^{n-1}} \\ 0 & \frac{1}{2^n} & \frac{1}{2^{n-1}} \\ 0 & 0 & 0 \end{pmatrix}$$

Teniendo en cuenta que $x_1(0) + x_2(0) + x_3(0) = 1$, obtenemos

$$\begin{aligned}x_1(n) &= 1 - \frac{1}{2^n}x_2(0) - \frac{1}{2^{n-1}}x_3(0) \\x_2(n) &= \frac{1}{2^n}x_2(0) + \frac{1}{2^{n-1}}x_3(0) \\x_3(n) &= 0\end{aligned}$$

Lo que nos da las proporciones de los tres genotipos en la generación n en función de las proporciones iniciales de genotipos. En el largo plazo se tiene que $\lim_{n \rightarrow \infty} \mathbf{X}(n) = (1, 0, 0)^t$.

Observa que este último resultado podíamos haberlo obtenido sin necesidad de calcular \mathbf{T}^n pues $\lambda = 1$ es un valor propio dominante simple y, aunque no podemos aplicar el teorema 5 porque la matriz \mathbf{T} no es primitiva, sí podemos aplicar el teorema 2, lo que nos da

$$\lim_{n \rightarrow \infty} \mathbf{X}(n) = \lim_{n \rightarrow \infty} \mathbf{T}^n \mathbf{X}(0) = \alpha \mathbf{P}_1$$

donde $\alpha = \langle \mathbf{X}(0) | \mathbf{Q} \rangle$, siendo \mathbf{Q} un vector propio de la matriz \mathbf{T}^t asociado al valor propio $\lambda = 1$ con $\langle \mathbf{P}_1 | \mathbf{Q} \rangle = 1$.

Como las filas de \mathbf{T}^t suman 1 se tiene que $\mathbf{Q} = (1, 1, 1)$ y, por tanto, $\alpha = 1$. También, por el teorema 2, se verifica que

$$\lim_{n \rightarrow \infty} \mathbf{T}^n = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Lo cual también se deduce directamente a la vista de cómo es \mathbf{T}^n .

Google: The page rank algorithm⁵

El universo (que otros llaman la Biblioteca) se compone de un número indefinido, y tal vez infinito, de galerías hexagonales,... Así empieza la inmortal narración de Jorge Luis Borges *La Biblioteca de Babel*. Escribe Borges, que la Biblioteca es eterna, siempre ha existido y perdurará, y entre sus innumerables libros hay uno que encierra todo el saber universal en el que está escrito todo; está escrito, por ejemplo, que en este momento tú estás leyendo estas palabras. Claro está, la probabilidad de que alguien encuentre ese libro es cero porque en la biblioteca reina el caos.

Esta narración que, entre otras cosas, es una hermosa e inabarcable metáfora del infinito, la escribió Borges en el año 1941. Hoy día esa mítica biblioteca de Babel existe y se llama *World Wide Web*, esto es, *Internet*. ¿De qué serviría toda la información que hay en Internet si no hubiera forma de ordenarla? Algo así debieron pensar Lawrence Page y Sergey Brin, dos estudiantes de doctorado de la Universidad de Stanford, cuando en el año 1998 estaban diseñando *Google*. Los buscadores que había en aquellos días no ordenaban con criterios eficientes las respuestas a las consultas que recibían, de forma que si en *Altavista* (el motor de búsqueda más popular en 1998) buscabas, por ejemplo, “Universidad”, podías obtener una lista de direcciones Web perfectamente inútiles, en las que aunque aparecía la palabra “Universidad” las razones para ello podían ser de todo tipo. No es de extrañar que *Altavista* desapareciera como tal pocos años después del nacimiento de *Google*. Porque *Google* vino a poner orden en el Universo, es decir, en Internet.

El curioso nombre del buscador está relacionado con la palabra *googol*, que alguien inventó para referirse al número inimaginable 10^{100} . Lawrence Page y Sergey Brin sabían que su buscador sería algo grandioso, y para ponerle nombre les vino la idea de esa palabreja y así surgió “Google”. En 1998 había unos 100 millones de sitios Web y *Altavista* atendía unos 20 millones de consultas diarias. En el año 2020 había unos 2000 millones de páginas Web y *Google* atendía más de 200 millones de consultas diarias. Y esto todavía está empezando.

El diseño de un motor de búsqueda en Internet plantea muy interesantes problemas: de almacenamiento, organización y actualización de datos, de búsqueda de los mismos, de seguridad frente a posibles ataques informáticos, y otros en los que aquí no vamos a entrar para centrarnos en lo que comentábamos antes ¿cómo ordenar los resultados de una búsqueda en Internet? Los creadores de *Google* querían que en un número suficientemente grande de casos, al menos una de las diez primeras páginas que se muestren contuviera información útil para quien realiza la consulta. Tú mismo, que seguramente eres *googleadicto*, puedes juzgar si han alcanzado este objetivo. Vamos a las matemáticas del asunto.

Puedes ver Internet como un enorme grafo dirigido, cuyos vértices son las páginas Web, que suponemos numeradas P_1, P_2, \dots, P_n , y hay una flecha dirigida de P_i a P_j cuando hay un enlace de la página P_i a la P_j . Vamos a llamar \mathbf{M} a la matriz de incidencia del grafo. Es una matriz cuadrada de orden n y $m_{ij} = 1$ cuando hay un enlace de P_j a P_i y $m_{ij} = 0$ en otro caso. Se trata, claro está, de asignar una “importancia” o “rango” a cada página reflejada en un valor numérico. El número de enlaces que una página recibe son un claro indicio de su valor, por tanto podríamos pensar en atribuir a cada página un rango proporcional al número de enlaces que recibe. Ese número es fácil de calcular: los enlaces que recibe la página i es la suma de la fila i de la matriz \mathbf{M} . Observa que aunque \mathbf{M} es enorme, sus elementos son casi todos ceros, es lo que se llama una **matriz dispersa**.

⁵Basado en el trabajo *El secreto de Google y el álgebra lineal*, del profesor Pablo Fernández Gallardo de la Universidad Autónoma de Madrid, que puede descargarse en <https://sctmates.webs.ull.es/modulo1lp/8/pfernandez.pdf>.

Si lo piensas un poco, esta primera idea no es muy buena porque le estamos dando el mismo valor a todos los enlaces, lo que no es lógico pues si, por ejemplo, una página recibe pocos enlaces pero desde páginas muy importantes, digamos desde Microsoft, Amazon o del World Wide Web Consortium (W3C), a esos enlaces hay que darles más valor que a otros. Por tanto, cambiamos la idea inicial, y ahora queremos *asignar a cada página un rango proporcional a la suma de los rangos de las páginas que la enlazan*. Llamemos x_i al rango de la página P_i y sea $K > 0$ la constante de proporcionalidad. Queremos que se cumpla la igualdad

$$x_i = K \sum_{j=1}^n m_{ij} x_j \quad (1 \leq i \leq n)$$

¡Vaya! Tenemos un sistema de n ecuaciones lineales con n incógnitas, las x_i . Date cuenta de que, definiendo el vector $X = (x_1, x_2, \dots, x_n)^t$ y $\lambda = \frac{1}{K}$, podemos escribir matricialmente este sistema en la forma

$$\mathbf{M}\mathbf{X} = \lambda\mathbf{X}$$

lo que nos dice que λ es un valor propio de \mathbf{M} y \mathbf{X} es un vector propio asociado al valor propio λ . ¡Esto ya es familiar!

Llegados aquí, lo mejor que podría pasar es que hubiera un único valor propio $\lambda > 0$, con un vector propio asociado positivo \mathbf{X} que fueran solución de nuestro problema. Si la matriz \mathbf{M} fuera una matriz de probabilidad primitiva esto estaría asegurado, pero \mathbf{M} está lejos de cumplir esas condiciones. El siguiente paso es modificarla de forma conveniente para que las cumpla.

Es razonable considerar que la importancia de una página es inversamente proporcional al número de enlaces que parten de ella, es decir, si desde una página salen muchos enlaces eso disminuye el valor global de cada uno de los mismos. El número de enlaces que salen de P_j es la suma, N_j , de la columna j de la matriz \mathbf{M} . Por tanto sustituimos la matriz \mathbf{M} por otra \mathbf{H} cuyos elementos son

$$h_{ij} = \frac{m_{ij}}{N_j}$$

Pero aún quedan algunos detalles que resolver. Por ejemplo, en \mathbf{H} habrá muchas columnas cuyos elementos sean todos nulos, por ejemplo, las que correspondan a páginas Web que contengan un documento, archivos pdf y parecidos, pues de esas páginas no sale ningún enlace. Lo que hicieron Page y Brin fue sustituir dichas columnas por un vector cuyas entradas son todas iguales a $\frac{1}{n}$, un número muy pequeño que no afecta casi nada a la importancia de las demás páginas, y que viene a decir que cuando un navegante sale de una de esas páginas puede ir aleatoriamente a cualquier otra. Llamemos \mathbf{a} al vector fila $1 \times n$ que en los lugares correspondientes a las columnas nulas de \mathbf{H} tiene un 1 y ceros en los demás lugares, y sea \mathbf{e} el vector columna $n \times 1$ cuyas componentes son todas 1. Entonces lo que hacemos es considerar la matriz

$$\mathbf{S} = \mathbf{H} + \frac{1}{n} \mathbf{e} \cdot \mathbf{a}$$

que es la misma \mathbf{H} con las columnas nulas sustituidas en la forma indicada. Observa que \mathbf{S} ya es una matriz de probabilidad pues sus elementos son no negativos y la suma de cada columna es igual a 1. Observa que también podemos considerar a \mathbf{S} como la matriz de transición de una cadena de Markov cuyos posibles estados son las páginas de la Red.

Naturalmente, queremos asegurarnos de que podemos aplicar el teorema de Perron – Frobenius, para ello la matriz \mathbf{S} tendría que ser irreducible o, lo que es igual, el grafo asociado, tendría que ser fuertemente conexo. Pero la Red no lo es y, además, cambia permanentemente, es casi como un organismo vivo, por lo que no tiene mucho sentido fijarse en su conexión. La solución que encontraron Page y Brin fue sustituir la matriz \mathbf{S} por la matriz

$$\mathbf{G} = \alpha\mathbf{S} + (1 - \alpha)\frac{1}{n}\mathbf{J}$$

donde $0 < \alpha < 1$ (ellos tomaron $\alpha = 0,85$) y \mathbf{J} es una matriz cuyas entradas son todas iguales a 1. Lo que se hace es sumar una muy pequeña cantidad positiva a todas las entradas de la matriz \mathbf{S} con lo que se consigue que todas las páginas distribuyan rango, de esta manera se evitan situaciones en que un grupo

de páginas se citan entre ellas pero no tienen ningún enlace hacia el exterior de ellas mismas, con lo cual en las sucesivas iteraciones de \mathbf{S} acumulan rango pero no lo distribuyen. Además, de esta forma la matriz \mathbf{G} sigue siendo una matriz de probabilidad pero ahora sus elementos son todos estrictamente positivos, por lo que es primitiva, y el teorema de Perron–Frobenius, junto con el teorema 5, nos garantiza que $\lambda = 1$ es un valor propio simple dominante con un vector propio asociado de componentes estrictamente positivas que, además, es el único vector propio de \mathbf{G} con esta propiedad.

De hecho, el teorema 5 nos dice que el vector que permite asignar rango a las páginas de la Red es el vector \mathbf{P} de la distribución de estado estacionario. Y dicho vector puede calcularse partiendo, por ejemplo del vector $\mathbf{X}(0) = \frac{1}{n}\mathbf{e}$ mediante iteraciones de la matriz \mathbf{G} , es decir, calculando sucesivamente $\mathbf{X}(k+1) = \mathbf{G}\mathbf{X}(k)$, pues sabemos que $\mathbf{P} = \lim_{k \rightarrow \infty} \mathbf{X}(k)$. La matriz \mathbf{G} no es buena para calcular con ella directamente porque todos sus elementos son positivos, pero volviendo hacia atrás tenemos que

$$\begin{aligned} \mathbf{X}(k+1) &= \mathbf{G}\mathbf{X}(k) = \left(\alpha\mathbf{S} + (1-\alpha)\frac{1}{n}\mathbf{J} \right) \mathbf{X}(k) = \alpha\mathbf{S}\mathbf{X}(k) + (1-\alpha)\frac{1}{n}\mathbf{J}\mathbf{X}(k) = \\ &= \alpha\mathbf{S}\mathbf{X}(k) + (1-\alpha)\frac{1}{n}\mathbf{e} = \alpha\mathbf{H}\mathbf{X}(k) + \frac{\alpha}{n}\mathbf{e}\mathbf{a}\mathbf{X}(k) + (1-\alpha)\frac{1}{n}\mathbf{e} = \\ &= \alpha\mathbf{H}\mathbf{X}(k) + \frac{1}{n}\mathbf{e}(\alpha\mathbf{a}\mathbf{X}(k) + (1-\alpha)) \end{aligned}$$

Como la matriz \mathbf{H} es una matriz dispersa, los cálculos con ella son mucho más rápidos que directamente con \mathbf{G} . Además, la velocidad de convergencia es bastante rápida de forma que no se precisan muchas iteraciones para obtener resultados aproximados aceptables. Los cálculos, según afirman en la página de Google, llevan un par de horas de trabajo de ordenador.