

Notas para el curso de Matemática II, módulo 1
Facultad de Ciencias
Universidad de la República

Andrés Abella

13 de octubre de 2021

Índice

1. Sistemas de ecuaciones	3
1.1. Sistemas 2×2	3
1.2. Sistemas en general	4
1.3. Sistemas homogéneos	7
2. Matrices	8
2.1. Definiciones	8
2.2. Operaciones con matrices.	10
2.3. Modelo de Leslie (primera parte)	13
3. Determinantes	16
3.1. Definición	16
3.2. Técnicas de cálculo	18
3.3. Propiedades	20
3.4. Aplicación a sistemas $n \times n$	21
3.5. Matriz inversa	22
4. Diagonalización	27
5. Aplicaciones	32
5.1. Genética, herencia autosómica.	32
5.2. Modelo de Leslie (segunda parte)	34
5.3. Un modelo estocástico	39
6. Apéndice 1. Diagonalización	42
7. Apéndice 2. Modelo de Leslie	44

1. Sistemas de ecuaciones

Un *sistema de ecuaciones (lineales)* es un conjunto de ecuaciones. Los siguientes son ejemplos de sistemas.

$$\begin{cases} 2x + y - z = 3 \\ 3x + 2y + z = -5 \end{cases}, \begin{cases} x + y = 1 \\ 2x + 3y = 0 \\ -x + 3y = -2 \end{cases}, \begin{cases} 2x + y = 1 \\ x + 3y = 2 \end{cases}, \begin{cases} x + y + z = 0 \\ 2x + 3y - 2z = 0 \\ -x + 3y - z = 0 \end{cases}. \quad (1)$$

El primero es un ejemplo de un sistema 2×3 , ya que tiene dos ecuaciones y tres incógnitas, el segundo es 3×2 , el tercero es 2×2 y el cuarto es 3×3 . En general un sistema puede tener una cantidad arbitraria de ecuaciones y de incógnitas. Una *solución* del sistema es una asignación de valores a las variables, de forma tal que se verifiquen todas las ecuaciones del sistema. Por ejemplo, si consideramos el primer sistema en (1), entonces una solución es $x = -1$, $y = 1$ y $z = -4$. *Resolver* el sistema es hallar todas las soluciones. En lo que sigue veremos primero cómo se resuelven los sistemas 2×2 y luego estudiaremos el caso general.

1.1. Sistemas 2×2

En esta sección veremos cómo resolver sistemas 2×2 . Lo mostraremos usando ejemplos. Consideremos

$$\begin{cases} 2x + y = 3 \\ 3x - 5y = 11 \end{cases}.$$

Un primer método de resolución es el de *sustitución*. Notar que en la primera ecuación es fácil despejar la segunda variable, obteniendo $y = 3 - 2x$. Si sustituimos este valor en la segunda ecuación, obtenemos una ecuación en una variable, la cual es fácil de resolver

$$3x - 5(3 - 2x) = 11 \quad \Rightarrow \quad 13x - 15 = 11 \quad \Rightarrow \quad 13x = 26 \quad \Rightarrow \quad x = 2.$$

Ahora el valor que obtuvimos de x lo sustituimos en $y = 3 - 2x$, obteniendo $y = 3 - 2(2) = -1$. Luego la solución es $x = 2$ e $y = -1$.

El método anterior funciona bien cuando alguna de las variables es fácil de despejar, en otros casos es mejor el *método de reducción*, que veremos a continuación. Consideremos el siguiente sistema

$$\begin{cases} 7x + 2y = 16 \\ 3x + 5y = 11 \end{cases}. \quad (2)$$

Notar que, en general, si multiplicamos una ecuación por una constante no nula, entonces obtenemos otra ecuación que tiene las mismas soluciones. La técnica acá consiste en multiplicar cada ecuación por una constante adecuada y luego sumarlas, de forma tal que en el resultado final desaparezca una de las incógnitas. Por ejemplo, si en el sistema anterior queremos eliminar la incógnita x , entonces multiplicando la primera ecuación por 3 y la segunda por -7 , y después sumándolas, obtenemos

$$\begin{cases} 21x + 6y = 48 \\ -21x + -35y = -77 \end{cases} \xrightarrow{(+)} -29y = -29 \quad \Rightarrow \quad y = 1.$$

Si ahora sustituimos ese valor de y en la primera ecuación del sistema (2), obtenemos

$$7x + 2(1) = 16 \quad \Rightarrow \quad 7x + 2 = 16 \quad \Rightarrow \quad 7x = 14 \quad \Rightarrow \quad x = 2.$$

Luego la solución es $x = 2$ e $y = 1$. Lo mismo obtendríamos si sustituimos y en la segunda ecuación.

Otra forma de resolver el sistema consiste en eliminar la y en vez de la x . Para eso en (2) mutiplicamos la primer ecuación por 5 y la segunda por -2 y las sumamos:

$$\begin{cases} 35x + 10y = 80 \\ -6x + -10y = -22 \end{cases} \Rightarrow 29x = 58 \Rightarrow x = 2.$$

Para obtener y , sustituimos el valor de x en cualquiera de las ecuaciones del sistema. Por ejemplo, sustituyendo en la primera obtenemos

$$7(2) + 2y = 16 \Rightarrow 14 + 2y = 16 \Rightarrow 2y = 2 \Rightarrow y = 1,$$

luego de nuevo llegamos a la solución $x = 2$ e $y = 1$.

Observación 1.1. Si en el plano introducimos un sistema de coordenadas, entonces podemos identificarlo con el conjunto $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$, siendo \mathbb{R} el conjunto de los números reales. Pensado \mathbb{R}^2 de esta forma, los puntos (x, y) que verifican una ecuación del tipo $ax + by = c$ (con a y b no nulos simultáneamente) son puntos que están alineados en una recta r ; en ese caso decimos que $ax + by = c$ es la *ecuación* de r . Luego cuando queremos resolver un sistema del tipo

$$\begin{cases} ax + by = c \\ dx + ey = f \end{cases}$$

lo que estamos haciendo es buscar los puntos comunes a dos rectas, es decir hallar su intersección. Como dos rectas pueden ser secantes, paralelas (y distintas), o coincidentes, entonces el sistema anterior puede tener una solución, ninguna, o infinitas, respectivamente. Ejemplos de lo anterior son los siguientes, respectivamente.

$$\begin{cases} x + y = 1 \\ x - y = 1 \end{cases}, \quad \begin{cases} x + y = 1 \\ x + y = -1 \end{cases}, \quad \begin{cases} x + y = 1 \\ 2x + 2y = 2 \end{cases}.$$

1.2. Sistemas en general

En general un *sistema de ecuaciones lineales* es un conjunto de ecuaciones de la forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}. \quad (3)$$

Los números a_{ij} , con $i = 1, \dots, m$, $j = 1, \dots, n$, son los *coeficientes*, b_1, \dots, b_m son los *términos independientes* y x_1, \dots, x_n son las *variables* o *incógnitas*, del sistema. Si queremos explicitar que hay m ecuaciones y n incógnitas, diremos que es un sistema $m \times n$. Si hay pocas variables, se suele escribir x, y, z, t en vez de x_1, x_2, x_3, x_4 .

Una *solución* del sistema (3) es una n -upla¹ de números $(\alpha_1, \dots, \alpha_n)$ tales que si en (3) sustituimos x_1 por α_1 , x_2 por α_2 y así seguimos hasta x_n por α_n , entonces se verifican todas las igualdades del sistema. Un sistema se dice *compatible* si tiene solución, en caso contrario se dice *incompatible*. Un sistema compatible se dice *determinado* si tiene una única solución e *indeterminado* en caso contrario. Se prueba que un sistema indeterminado siempre tiene infinitas soluciones. Por *resolver* un sistema entendemos que es clasificarlo en incompatible, compatible determinado o indeterminado, y en estos últimos casos hallar todas sus soluciones.

¹Una n -upla es un conjunto ordenado de n elementos, pudiendo ser $n = 1, 2, 3, \dots$

Un sistema del tipo

$$\begin{cases} x + y + z = 2 \\ 2y + z = 4 \\ 3z = 6 \end{cases},$$

se dice que es un sistema *escalerizado*. Este sistema es muy fácil de resolver, puesto que de la última ecuación se deduce que es $z = 2$, sustituyendo este valor de z en la segunda ecuación deducimos que es $y = 1$, y finalmente sustituyendo estos valores de y y z en la primera ecuación deducimos que es $x = -1$. Luego el sistema es compatible determinado y su solución es $x = -1$, $y = 1$ y $z = 2$. Otros ejemplos de sistemas escalerizados son

$$\begin{cases} x + y + z = 2 \\ 2y + z = 4 \\ 0 = 2 \end{cases} \quad \text{y} \quad \begin{cases} x + y + z = 2 \\ 2y + z = 4 \\ 0 = 0 \end{cases}.$$

Claramente el primer sistema es incompatible y el segundo es compatible indeterminado. Las soluciones de este último se pueden escribir de la forma, $x = y - 2$, $z = -2y + 4$, mientras que la variable y queda libre. Otra forma de escribir esas soluciones es

$$x = \lambda - 2, \quad y = \lambda, \quad z = -2\lambda + 4,$$

siendo λ un número real arbitrario. Este último sistema es compatible indeterminado con *un grado de libertad* (la variable y está libre, si le damos un valor, entonces x y z quedan determinadas).

En general, los sistemas de ecuaciones que más aparecen son los *cuadrados*, que son los que tienen tantas ecuaciones como variables. Nosotros estudiaremos principalmente este tipo de sistemas de ecuaciones y al final de la sección veremos un par de ejemplos en que la cantidad de variables y de ecuaciones es distinta.

Dos sistemas de ecuaciones se dicen *equivalentes* si tienen las mismas soluciones². Usaremos el símbolo \sim para indicar la equivalencia de sistemas.

Proposición 1.2. *Operaciones con las ecuaciones de un sistema que dan lugar a sistemas equivalentes.*

- Intercambiar el orden de las variables (es decir, el orden de los sumandos).
- Intercambiar el orden de las ecuaciones.
- Sustituir una ecuación por un múltiplo no nulo de la misma.
- Sumarle a una ecuación un múltiplo de otra. □

El método para resolver sistemas de ecuaciones conocido como de *eliminación Gaussiana* o *escalerización*, consiste en aplicar reiteradamente las operaciones descritas en la proposición anterior hasta obtener un sistema equivalente que esté en forma escalerizada y por lo tanto sea fácil de resolver. Mostraremos este método mediante los siguientes ejemplos.

Ejemplo 1.3. Queremos resolver el siguiente sistema

$$\begin{cases} 2x + 3y + 5z = 9 \\ x + 2y + 3z = 5 \\ 3x + 4y + 5z = 5 \end{cases}.$$

²Acá estamos sobreentendiendo que si uno de los sistemas es incompatible, entonces el otro también lo es.

Vamos a ir aplicando las propiedades de la proposición anterior hasta llevarlo a un sistema escalerizado.

$$\begin{aligned} & \begin{cases} 2x + 3y + 5z = 9 \\ x + 2y + 3z = 5 \\ 3x + 4y + 5z = 5 \end{cases} \sim \begin{cases} x + 2y + 3z = 5 \\ 2x + 3y + 5z = 9 \\ 3x + 4y + 5z = 5 \end{cases} \sim \begin{cases} x + 2y + 3z = 5 \\ -y - z = -1 \\ 3x + 4y + 5z = 5 \end{cases} \\ & \sim \begin{cases} x + 2y + 3z = 5 \\ -y - z = -1 \\ -2y - 4z = -10 \end{cases} \sim \begin{cases} x + 2y + 3z = 5 \\ y + z = 1 \\ y + 2z = 5 \end{cases} \sim \begin{cases} x + 2y + 3z = 5 \\ y + z = 1 \\ z = 4 \end{cases} . \end{aligned}$$

En el primer paso, intercambiamos la primer ecuación con la segunda; en el segundo, a la segunda ecuación le sumamos la primera multiplicada por -2 ; en el tercero, a la tercera le sumamos la primera multiplicada por -3 ; en el cuarto, multiplicamos la segunda por -1 y la tercera por $-1/2$; en el quinto, a la tercera ecuación le sumamos la segunda multiplicada por -1 . Luego este sistema es compatible determinado, y su solución es $x = -1$, $y = -3$ y $z = 4$.

Ejemplo 1.4.

$$\begin{cases} x + 2y + 3z = -1 \\ 3x - y + 2z = 1 \\ -2x + 3y + z = 2 \end{cases} \sim \begin{cases} x + 2y + 3z = -1 \\ -7y - 7z = 4 \\ 7y + 7z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = -1 \\ -7y - 7z = 4 \\ 0 = 4 \end{cases} .$$

En el primer paso, a la segunda ecuación le sumamos la primera multiplicada por -3 y a la tercera le sumamos la primera multiplicada por 2 ; en el segundo paso, sumamos la segunda ecuación a la tercera. Este sistema resultó ser incompatible, porque es claro que la última ecuación no tiene solución.

Ejemplo 1.5.

$$\begin{cases} x + 2y + 3z = -1 \\ 3x - y + 2z = 1 \\ -2x + 3y + z = -2 \end{cases} \sim \begin{cases} x + 2y + 3z = -1 \\ -7y - 7z = 4 \\ 7y + 7z = -4 \end{cases} \sim \begin{cases} x + 2y + 3z = -1 \\ -7y - 7z = 4 \\ 0 = 0 \end{cases} .$$

Observar que los coeficientes de este sistema son los mismos que los del sistema del ejemplo anterior, así que los pasos de resolución son los mismos (solo cambian los términos independientes). Las soluciones se pueden escribir de la forma $x = -z + \frac{1}{7}$, $y = -z - \frac{4}{7}$ y la variable z está libre. Luego es un sistema compatible indeterminado con un grado de libertad.

Ejemplo 1.6.

$$\begin{cases} x + y - z = 1 \\ -x - y + z = -1 \\ 2x + 2y - 2z = 2 \end{cases} \sim \begin{cases} x + y - z = 1 \\ 0 = 0 \\ 0 = 0 \end{cases} .$$

Acá a la segunda ecuación le sumamos la primera y a la tercera le sumamos la primera multiplicada por -2 . Las soluciones se pueden escribir de la forma $z = x + y + 1$ y las variables x e y están libres. Este es un ejemplo de un sistema compatible indeterminado con *dos grados de libertad*. En sistemas con más variables, pueden aparecer soluciones que tengan más grados de libertad.

A continuación veremos un par de ejemplos de aplicación del método de escalerización para resolver sistemas que no son cuadrados.

Ejemplo 1.7. Un sistema 2×3 :

$$\begin{cases} x + y + z = 2 \\ x + 2y - z = 3 \end{cases} \sim \begin{cases} x + y + z = 2 \\ y - 2z = 1 \end{cases} .$$

Acá restamos la primer ecuación a la segunda. Este sistema es compatible indeterminado y las soluciones se pueden escribir de la forma $x = 1 - 3z$, $y = 1 + 2z$ y la variable z está libre, luego tiene un grado de libertad.

Ejemplo 1.8. Un sistema 3×2 :

$$\begin{cases} x + 2y = 1 \\ 2x - y = 2 \\ -x + 3y = -1 \end{cases} \sim \begin{cases} x + 2y = 1 \\ -5y = 0 \\ 5y = 0 \end{cases} \sim \begin{cases} x + 2y = 1 \\ -5y = 0 \\ 0 = 0 \end{cases} .$$

Acá primero a la segunda ecuación le sumamos la primera multiplicada por -2 , y a la tercera le sumamos la primera; luego a la tercera le sumamos la segunda. El sistema es compatible determinado y la solución es $x = 1$ e $y = 0$.

1.3. Sistemas homogéneos

Un sistema de ecuaciones se dice *homogéneo* si todas sus ecuaciones aparecen igualadas a cero, es decir si es de la forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{cases}$$

Un sistema homogéneo siempre admite la solución *trivial* $x_1 = \cdots = x_n = 0$, por lo que siempre es compatible. Luego es indeterminado en caso de admitir alguna solución no trivial, y es determinado si la única solución que admite es la trivial.

En los siguientes ejemplos estudiamos los sistemas homogéneos asociados a los sistemas de los ejemplos anteriores (es decir obtenidos cambiando la columna de términos independientes por una columna de ceros). Como las operaciones para resolverlos son las mismas que para los no homogéneos, simplemente describiremos el camino que lleva a su resolución.

Ejemplo 1.9.

$$\begin{cases} 2x + 3y + 5z = 0 \\ x + 2y + 3z = 0 \\ 3x + 4y + 5z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = 0 \\ 2x + 3y + 5z = 0 \\ 3x + 4y + 5z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = 0 \\ -y - z = 0 \\ 3x + 4y + 5z = 0 \end{cases} \\ \sim \begin{cases} x + 2y + 3z = 0 \\ -y - z = 0 \\ -2y - 4z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = 0 \\ y + z = 0 \\ 2y + 4z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = 0 \\ y + z = 0 \\ 2z = 0 \end{cases} .$$

El sistema es compatible determinado y su única solución es la trivial $x = y = z = 0$.

Ejemplo 1.10.

$$\begin{cases} x + 2y + 3z = 0 \\ 3x - y + 2z = 0 \\ -2x + 3y + z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = 0 \\ -7y - 7z = 0 \\ 7y + 7z = 0 \end{cases} \sim \begin{cases} x + 2y + 3z = 0 \\ -7y - 7z = 0 \\ 0 = 0 \end{cases} .$$

El sistema es compatible indeterminado. Una forma de escribir sus soluciones es $x = -z$, $y = -z$, con z libre; luego tiene un grado de libertad. Observar que dándole a z valores no nulos obtenemos soluciones no triviales del sistema. Por ejemplo, usando $z = 1$ obtenemos que $x = y = -1$ y $z = 1$ es solución del sistema.

Ejemplo 1.11.

$$\begin{cases} x + y - z = 0 \\ -x - y + z = 0 \\ 2x + 2y - 2z = 0 \end{cases} \sim \begin{cases} x + y - z = 0 \\ 0 = 0 \\ 0 = 0 \end{cases} .$$

El sistema es compatible indeterminado con dos grados de libertad. Las soluciones se pueden escribir de la forma $z = x + y$, con x e y libres.

Ejemplo 1.12. Un sistema 2×3 :

$$\begin{cases} x + y + z = 0 \\ x + 2y - z = 0 \end{cases} \sim \begin{cases} x + y + z = 0 \\ y - 2z = 0 \end{cases}.$$

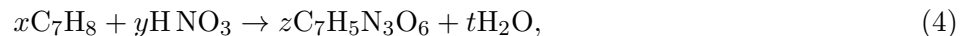
El sistema es compatible indeterminado con un grado de libertad. Las soluciones se pueden escribir de la forma $x = -3z$, $y = 2z$, con z libre.

Ejemplo 1.13. Un sistema 3×2 :

$$\begin{cases} x + 2y = 0 \\ 2x - y = 0 \\ -x + 3y = 0 \end{cases} \sim \begin{cases} x + 2y = 0 \\ -5y = 0 \\ 5y = 0 \end{cases} \sim \begin{cases} x + 2y = 0 \\ -5y = 0 \\ 0 = 0 \end{cases}.$$

El sistema es compatible determinado y su única solución es la trivial $x = y = 0$.

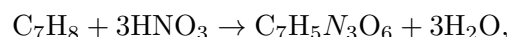
Ejemplo 1.14. Para producir TNT (trinitrotolueno) que tiene fórmula $C_7H_5N_3O_6$, se hace reaccionar tolueno C_7H_8 con ácido nítrico HNO_3 ; esto produce TNT y agua. Por lo tanto la fórmula que describe la reacción es



siendo x, y, z, t las cantidades de moléculas de reactivos y productos. Para que la fórmula esté equilibrada, la cantidad total de átomos de cada tipo que hay en los reactivos, debe de coincidir con la que hay en los productos. Por ejemplo, en la fórmula anterior tenemos 7 átomos de carbono en cada molécula de tolueno, y como tenemos x moléculas de tolueno, entonces inicialmente tenemos $7x$ átomos de carbono. Por otro lado luego de la reacción vamos a tener $7z$ átomos de carbono (que están en el TNT), luego tiene que ser $7x = 7z$. Razonando análogamente con el hidrógeno, nitrógeno y oxígeno, obtenemos el siguiente sistema de ecuaciones.

$$\begin{array}{l} \text{C)} \quad 7x = 7z \\ \text{H)} \quad 8x + y = 5z + 2t \\ \text{N)} \quad y = 3z \\ \text{O)} \quad 3y = 6z + t \end{array} \Leftrightarrow \begin{cases} 7x - 7z = 0 \\ 8x + y - 5z - 2t = 0 \\ y - 3z = 0 \\ 3y - 6z - t = 0 \end{cases}$$

Este es un sistema homogéneo. Para resolverlo, en vez de escalerizar resulta más simple observar que de las ecuaciones primera y tercera obtenemos $x = z$ e $y = 3z$. Sustituyendo estos valores en las dos ecuaciones restantes y despejando obtenemos $t = 3z$. Luego el sistema es compatible indeterminado con un grado de libertad: $x = z$, $y = t = 3z$ y la variable z está libre. Notar que si tomamos $z = 1$, hallamos los valores restantes y sustituimos en (4), obtenemos



esta es la ecuación *ajustada* de la reacción.

Como aplicación, supongamos que queremos obtener 10 moléculas de TNT. Con nuestra notación es $z = 10$, lo cual implica $x = 10$ e $y = t = 30$. Luego necesitamos 10 moléculas de tolueno y 30 de ácido nítrico, las cuales van a producir 10 de TNT y 30 de agua.

2. Matrices

2.1. Definiciones

Una *matriz* es una tabla de números, por ejemplo $\begin{pmatrix} 2 & -3 & 1/5 \\ \sqrt{2} & 0 & -1 \end{pmatrix}$ es una matriz 2×3 (tiene 2 filas y 3 columnas). En general una matriz 2×3 tiene la forma $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$ con a, b, \dots, f números reales arbitrarios.

Para matrices de tamaño grande es preferible usar subíndices en vez de letras distintas. En ese sentido una matriz 2×3 se escribe $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$, donde los a_{ij} son números reales. Observar que en la notación a_{ij} , el primer índice i indica la fila y el segundo índice j indica la columna, a las que pertenece el elemento. Generalizando lo anterior, una *matriz* $m \times n$ (m y n enteros positivos) es una tabla de números de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

en que los a_{ij} son números reales, para todo $i = 1, \dots, m$ y $j = 1, \dots, n$. Observar que una matriz $m \times n$ tiene m filas y n columnas. Los elementos a_{ij} se llaman los *coeficientes* o *entradas* de la matriz. Para las matrices usaremos la notación $A = (a_{ij})_{m \times n}$ y cuando no dé lugar a confusión escribiremos simplemente $A = (a_{ij})$. Al conjunto de todas las matrices $m \times n$ lo escribiremos $M_{m \times n}$. Una *matriz columna* es una matriz $m \times 1$ (para algún m) y una *matriz fila* es una matriz $1 \times n$ (para algún n). En general en esos casos se usan subíndices simples para las coeficientes:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \text{ matriz columna; } (a_1 \ a_2 \ \cdots \ a_n) \text{ matriz fila.}$$

La idea de $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ se puede generalizar a $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\}$ el conjunto de todas las n -uplas de números reales, siendo $n = 2, 3, \dots$. Como una n -upla (x_1, x_2, \dots, x_n) es lo mismo que una matriz fila $1 \times n$, entonces \mathbb{R}^n se identifica naturalmente con $M_{1 \times n}$. También se puede identificar \mathbb{R}^n con el conjunto de las matrices columna $M_{n \times 1}$, escribiendo una n -upla en forma vertical

$$(x_1, x_2, \dots, x_n) \leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Por la razón anterior a las matrices fila o columna se les suele llamar *vectores* (fila o columna), que es el nombre que se usa para los elementos de \mathbb{R}^n .

Decimos que dos matrices A y B son *iguales* y escribimos $A = B$ si tienen la misma cantidad de filas y de columnas, y coinciden coeficiente a coeficiente. Es decir, si $A = (a_{ij})_{m \times n}$ y $B = (b_{ij})_{r \times s}$, entonces $A = B$ si y solo si $m = r$, $n = s$ y $a_{ij} = b_{ij}$, para todo $i = 1, \dots, m$ y $j = 1, \dots, n$.

Una matriz se dice *cuadrada* si tiene el mismo número de filas que de columnas, es decir si es de la forma

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

Al conjunto de las matrices cuadradas $n \times n$ en general lo escribiremos simplemente M_n , en vez de $M_{n \times n}$. Si $A = (a_{ij})$ es una matriz cuadrada $n \times n$, entonces la *diagonal principal* de A es la n -upla $(a_{11}, a_{22}, \dots, a_{nn})$.

Matrices cuadradas particulares. Sea $A = (a_{ij})_{n \times n}$ una matriz cuadrada.

- A es *simétrica* si $a_{ij} = a_{ji}$, para todo $i, j = 1, \dots, n$ (es simétrica respecto a la diagonal principal).
- A es *antisimétrica* si $a_{ij} = -a_{ji}$, para todo $i, j = 1, \dots, n$.
- A es *triangular superior* si $a_{ij} = 0$, para todo $i > j$ (las entradas abajo de la diagonal principal son nulas).
- A es *triangular inferior* si $a_{ij} = 0$, para todo $i < j$ (las entradas arriba de la diagonal principal son nulas).
- A es *diagonal* si $a_{ij} = 0$, para todo $i \neq j$ (las entradas fuera de la diagonal principal son nulas).

Un caso particular de matriz diagonal es la matriz *identidad* I_n , que es la matriz $n \times n$ que tiene el valor 1 en la diagonal principal y 0 en el resto, es decir

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Si introducimos el símbolo δ_{ij} , llamado *delta de Kronecker*, definido por

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j \end{cases},$$

entonces $I_n = (\delta_{ij})_{n \times n}$. Cuando no de lugar a confusión, escribiremos I en vez de I_n .

2.2. Operaciones con matrices.

Trasposición. Si $A = (a_{ij})$ es una matriz $m \times n$, entonces su *traspuesta* A^t es la matriz obtenida intercambiando las filas con las columnas de A , es decir $A^t = (b_{ij})$, siendo $b_{ij} = a_{ji}$, para todo i, j . Notar que A^t es una matriz $n \times m$. La trasposición verifica las siguientes propiedades.

Proposición 2.1. ▪ Para toda matriz A es $(A^t)^t = A$.

- Una matriz A es *simétrica* si y solo si $A^t = A$.
- Una matriz A es *antisimétrica* si y solo si $A^t = -A$ (la matriz $-A$ se define en el párrafo siguiente).
- Una matriz A es *triangular superior* si y solo si A^t es *triangular inferior*. □

Suma de matrices. La *suma* de matrices en $M_{m \times n}$ se define sumando coeficiente a coeficiente,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

La matriz $O = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in M_{m \times n}$ es la *matriz nula*. Si $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, entonces su opuesta es $-A = \begin{pmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{m1} & -a_{m2} & \cdots & -a_{mn} \end{pmatrix}$. Por ejemplo, en matrices 2×3 es

$$O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad - \begin{pmatrix} 2 & -3 & 0 \\ -5 & 1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 3 & 0 \\ 5 & -1 & -2 \end{pmatrix}.$$

La suma de matrices verifica las siguientes propiedades.

Proposición 2.2. Sean $A, B, C \in M_{m \times n}$, arbitrarias. Entonces

1. $A + B = B + A$, (*conmutativa*);
2. $A + (B + C) = (A + B) + C$ (*asociativa*);
3. $A + O = O + A = A$ (*existencia de neutro*);
4. $A + (-A) = (-A) + A = O$ (*existencia de opuesto*). □

La propiedad asociativa permite sacar paréntesis escribiendo $A + B + C = (A + B) + C = A + (B + C)$. Lo mismo sucede si hay más sumandos.

Producto de un escalar por una matriz. Al trabajar con matrices, a los números reales se les suele llamar *escalares*. El producto de un escalar $c \in \mathbb{R}$ por una matriz $A \in M_{m \times n}$ es la matriz $cA \in M_{m \times n}$ definida multiplicando cada entrada de A por el escalar c , es decir

$$c \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}.$$

El producto de un escalar por una matriz verifica las siguientes propiedades.

Proposición 2.3. Sean $A, B \in M_{m \times n}$ y $c, d \in \mathbb{R}$ arbitrarios. Entonces

$$1A = A, \quad c(A + B) = cA + cB, \quad (c + d)A = cA + dA, \quad c(dA) = (cd)A. \quad \square$$

Notar que las operaciones anteriores también se pueden definir en $\mathbb{R}^n = M_{1 \times n}$, mediante

$$a(x_1, \dots, x_n) = (ax_1, \dots, ax_n), \quad (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

y verifican las mismas propiedades.

Producto de matrices. Si $A = (a_{ij}) \in M_{m \times n}$ y $B = (b_{ij}) \in M_{n \times p}$, entonces su *producto* es la matriz $AB = (c_{ij}) \in M_{m \times p}$, en que c_{ij} está definido por

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}, \quad (5)$$

para todo $i = 1, \dots, m$ y $j = 1, \dots, p$. Para recordar esta fórmula, es útil armar el siguiente diagrama

$$\begin{pmatrix} \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \cdots & b_{ij} & \cdots \\ \vdots & & \vdots \\ \cdots & b_{nj} & \cdots \\ \vdots & & \vdots \\ \cdots & c_{ij} & \cdots \\ \vdots & & \vdots \end{pmatrix}.$$

El elemento c_{ij} se obtiene multiplicando lugar a lugar los elementos de la fila (a_{i1}, \dots, a_{in}) con los de la columna (b_{1j}, \dots, b_{nj}) y luego sumándolos. Notar que para poder multiplicar una matriz A por otra B , es necesario que el número de columnas de A coincida con el de filas de B .

Ejemplo 2.4. Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$, entonces $AB = \begin{pmatrix} 5 & 4 & 5 & 0 \\ 11 & 10 & 14 & 0 \end{pmatrix}$.

Observación 2.5. Si A y B son matrices cuadradas de la misma dimensión, entonces podemos realizar los productos AB y BA . Por ejemplo, si

$$A = \begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix} \text{ y } B = \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} \Rightarrow AB = \begin{pmatrix} 8 & 4 \\ -16 & -8 \end{pmatrix} \text{ y } BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Notar que este ejemplo muestra varias particularidades del producto de matrices:

- $AB \neq BA$, luego el producto de matrices no es conmutativo;
- $BA = O$, pero $AB \neq O$;
- $BA = O$, siendo $A \neq O$ y $B \neq O$, luego el producto de dos matrices no nulas puede dar la matriz nula.

Además si $A = \begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix}$ y consideramos $E = \begin{pmatrix} 3 & 2 \\ 7 & 4 \end{pmatrix}$ y $F = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, entonces operando obtenemos

$$EA = FA = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}.$$

Esto muestra que para el producto de matrices no vale la propiedad cancelativa, es decir $EA = FA$ con $A \neq O$ no implica $E = F$ (lo mismo sucede si $AE = AF$).

La siguiente proposición resume las propiedades básicas del producto de matrices.

Proposición 2.6. Valen las siguientes propiedades.

1. $A(BC) = (AB)C$ (*asociativa*).
2. $A(B + C) = AB + AC$ (*distributiva*).
3. $(A + B)C = AC + BC$ (*distributiva*).

4. $AI_n = I_m A = A$, si $A \in M_{m \times n}$.
5. $AO = OA = O$, siendo O la matriz nula (notar que si $A \in M_{m \times n}$ con $m \neq n$, entonces en la igualdad anterior las tres matrices O tienen distintas dimensiones).
6. $A(cB) = (cA)B = c(AB)$, para todo $c \in \mathbb{R}$.

En lo anterior se asume que A , B y C son matrices de dimensiones adecuadas para poder multiplicarse y/o sumarse. \square

Observaciones 2.7. 1. La propiedad asociativa permite sacar paréntesis escribiendo $ABC = (AB)C = A(BC)$. Lo mismo sucede si hay más factores.

2. El único caso en que dadas dos matrices A y B , podemos calcular $A + B$ y AB , es cuando A y B son dos matrices cuadradas de la misma dimensión.
3. El producto de matrices diagonales es particularmente simple

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & 0 & \cdots & 0 \\ 0 & a_2 b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n b_n \end{pmatrix}.$$

4. Si $A \in M_n$ y $k \in \mathbb{N}$, entonces definimos la *potencia* k -ésima A^k de A mediante

$$A^0 = I \text{ (matriz identidad)}, \quad A^1 = A, \quad A^2 = AA, \quad A^3 = AAA, \quad \dots$$

Calcular A^k para valores grandes de k en general es difícil, pero para matrices diagonales es fácil

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}^k = \begin{pmatrix} a_1^k & 0 & \cdots & 0 \\ 0 & a_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^k \end{pmatrix}, \quad \forall k = 0, 1, 2, \dots$$

Finalizamos con la siguiente proposición, que describe cómo se relaciona la trasposición con las otras operaciones.

Proposición 2.8. *Valen las siguientes propiedades*

$$(A + B)^t = A^t + B^t, \quad (cA)^t = cA^t, \quad (AB)^t = B^t A^t,$$

siendo, en cada caso, A y B matrices de dimensiones adecuadas para realizar las operaciones correspondientes y c una constante. \square

2.3. Modelo de Leslie (primera parte)

El *modelo de Leslie* es un modelo que se usa para el estudio del crecimiento de una población. Este modelo describe el crecimiento del grupo de las hembras de una población animal (o humana). Partimos de una población dividida en n clases de edad, las cuales tienen todas la misma duración de l años³. Lo que nos interesa es saber cómo evoluciona la cantidad de hembras de cada clase, a medida que pasa el tiempo. Empezamos a contar el tiempo a partir del instante $t_0 = 0$. A los l años tendremos el instante t_1 , luego a los $l + l = 2l$ años tendremos el t_2 y así seguimos. En general el instante t_k es a los kl años, para $k = 0, 1, 2, \dots$

³Estamos asumiendo que el tiempo lo medimos en años, pero podría ser en cualquier otra unidad de medida: meses, días, horas, etc.

Mostraremos primero cómo funciona el modelo en un ejemplo y luego lo generalizaremos. Supongamos que la edad máxima que pueden alcanzar las hembras de una población animal es de 20 años y que la dividimos en 4 clases de edad, con intervalos de 5 años. En la primera clase están las que tienen edades de 0 a 5 años, en la segunda de 5 a 10, en la tercera de 10 a 15 y en la cuarta de 15 a 20. Supongamos que tenemos los siguientes datos.

1. En promedio, en 5 años las hembras de la primera clase no tienen hijas, las de la segunda tienen 1, las de la tercera 3 y las de la cuarta 2.
2. El promedio de supervivencia a los 5 años de las hembras de la clase 1 es de $1/4$, el de las de clase 2 es $1/2$ y el de las de clase 3 es $1/10$ (las de clase 4 se mueren, ya que pasan la edad máxima de sobrevivencia).

Llamaremos x_0 a la cantidad inicial de hembras de la primera clase e y_0 , z_0 y w_0 , a las cantidades iniciales de las clases dos, tres y cuatro, respectivamente. Pasado un año estamos en el instante t_1 y tendremos nuevas cantidades en cada clase, que escribiremos x_1 , y_1 , z_1 y w_1 , respectivamente, y así seguimos. Para operar con estos datos los agruparemos en vectores según la clase de edad

$$X_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ w_0 \end{pmatrix}, \quad X_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix}, \quad \dots \quad X_k = \begin{pmatrix} x_k \\ y_k \\ z_k \\ w_k \end{pmatrix}, \quad \dots \quad (6)$$

A X_k se le llama el k -ésimo *vector de distribución* de edades, ya que contiene la cantidad de hembras de cada clase, en el instante t_k ; a X_0 se le llama el *vector de distribución inicial* de edades. Notar que la cantidad total de hembras de la población en el instante t_k es $x_k + y_k + z_k + w_k$.

Lo que nos interesa es saber cómo pasar de un vector de distribución de edades X_{k-1} al siguiente X_k . Sean entonces

$$X_{k-1} = \begin{pmatrix} x_{k-1} \\ y_{k-1} \\ z_{k-1} \\ w_{k-1} \end{pmatrix}, \quad X_k = \begin{pmatrix} x_k \\ y_k \\ z_k \\ w_k \end{pmatrix}.$$

En el instante t_k , la cantidad de hembras que están en la clase 1 viene dado únicamente por las nacidas entre t_{k-1} y t_k (si nacieron antes están en la clase 2 o superior). Recordar que las que están en la primera clase no generan hijas, las de la segunda generan una hija cada una y las de la tercera generan 3 y las de la cuarta generan 2. Luego

$$x_k = 0 \times x_{k-1} + 1 \times y_{k-1} + 3 \times z_{k-1} + 2 \times w_{k-1} = y_{k-1} + 3z_{k-1} + 2w_{k-1}.$$

Ahora vamos a calcular las otras entradas de X_k . El número de hembras que están en la segunda clase en el instante t_k está dado por el número de hembras sobrevivientes de las que estaban en la clase 1 en el instante t_{k-1} . Como sobreviven $1/4$ del total de hembras de la primera clase, obtenemos $y_k = \frac{1}{4}x_{k-1}$. Razonando análogamente con las otras dos clases obtenemos $z_k = \frac{1}{2}y_{k-1}$ y $w_k = \frac{1}{10}z_{k-1}$, luego

$$\begin{aligned} x_k &= y_{k-1} + 3z_{k-1} + 2w_{k-1} \\ y_k &= \frac{1}{4}x_{k-1} \\ z_k &= \frac{1}{2}y_{k-1} \\ w_k &= \frac{1}{10}z_{k-1} \end{aligned}$$

Estas igualdades equivalen a la siguiente igualdad matricial

$$\begin{pmatrix} x_k \\ y_k \\ z_k \\ w_k \end{pmatrix} = \begin{pmatrix} 0 & 1 & 3 & 2 \\ 1/4 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/10 & 0 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ y_{k-1} \\ z_{k-1} \\ w_{k-1} \end{pmatrix}. \quad (7)$$

La matriz que aparece en la fórmula anterior es la *matriz de Leslie* de la población. Escribimos

$$L = \begin{pmatrix} 0 & 1 & 3 & 2 \\ 1/4 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/10 & 0 \end{pmatrix}.$$

Con esa notación, la fórmula (7) –que nos permite pasar de un vector de distribución de edades al siguiente– queda en

$$X_k = LX_{k-1}, \quad \forall k = 1, 2, \dots \quad (8)$$

Veremos cómo aplicar lo anterior cuando tenemos datos concretos. Supongamos que inicialmente hay 190 hembras, de las cuales 100 están en la primera clase, 60 en la segunda, 20 en la tercera y 10 en la cuarta. Con esos datos el vector de distribución inicial es

$$X_0 = \begin{pmatrix} 100 \\ 60 \\ 20 \\ 10 \end{pmatrix}.$$

Multiplicando este vector por la matriz de Leslie L obtenemos el vector de distribución de edades siguiente

$$X_1 = LX_0 = \begin{pmatrix} 0 & 1 & 3 & 2 \\ 1/4 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/10 & 0 \end{pmatrix} \begin{pmatrix} 100 \\ 60 \\ 20 \\ 10 \end{pmatrix} = \begin{pmatrix} 140 \\ 25 \\ 30 \\ 2 \end{pmatrix}.$$

Repitiendo el proceso obtenemos

$$X_2 = LX_1 = \begin{pmatrix} 119 \\ 35 \\ 12,5 \\ 3 \end{pmatrix}, \quad X_3 = LX_2 = \begin{pmatrix} 78,5 \\ 29,75 \\ 17,5 \\ 1,25 \end{pmatrix}, \quad X_4 = LX_3 = \begin{pmatrix} 84,75 \\ 19,625 \\ 14,875 \\ 1,75 \end{pmatrix}.$$

Por lo tanto, luego de 20 años (4 períodos de tiempo), tendremos aproximadamente 85 hembras en la primera clase, 20 en la segunda, 15 en la tercera y 2 en la cuarta. En total tendremos 122 hembras.

En lo que sigue generalizamos el ejemplo anterior. Supongamos que tenemos una población dividida en 4 clases de edad de igual duración cada una. Llamamos x_k a la cantidad de hembras de la primera clase e y_k , z_k y w_k , a las cantidades de hembras de las clases dos, tres y cuatro, respectivamente, en el instante t_k . Con estos valores nos construimos los vectores de distribuciones de edades X_0, X_1, \dots , como en (6).

Para saber cómo pasar de un vector de distribución de edades X_{k-1} al siguiente X_k , nos construimos la matriz de Leslie

$$L = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & 0 & 0 & 0 \\ 0 & b_2 & 0 & 0 \\ 0 & 0 & b_3 & 0 \end{pmatrix},$$

en la cual, para cada $i = 1, 2, 3, 4$, el valor a_i es la cantidad de hijas promedio que nacen de una sola hembra durante el tiempo que está en la clase i , y b_i a la fracción de la cantidad de hembras en la clase i , que se espera que sobrevivan y pasen a la clase $i + 1$. Notar que es $a_i \geq 0$ y $1 \geq b_i > 0$, para todo i . Con estas notaciones, razonado como antes obtenemos

$$\begin{pmatrix} x_k \\ y_k \\ z_k \\ w_k \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & 0 & 0 & 0 \\ 0 & b_2 & 0 & 0 \\ 0 & 0 & b_3 & 0 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ y_{k-1} \\ z_{k-1} \\ w_{k-1} \end{pmatrix}. \quad (9)$$

Lo cual escrito en forma más compacta es

$$X_k = LX_{k-1}, \quad \forall k = 1, 2, \dots \quad (10)$$

Notar que de la fórmula (10) se deduce

$$X_1 = LX_0, \quad X_2 = LX_1 = L(LX_0) = L^2X_0, \quad X_3 = LX_2 = L(L^2X_0) = L^3X_0, \quad \dots$$

Es decir, vale

$$X_k = L^k X_0, \quad \forall k = 1, 2, \dots$$

Luego para poder conocer la distribución de edades X_k en el instante t_k , nos alcanza con saber la distribución inicial de edades X_0 y calcular L^k (la potencia k -ésima de la matriz de Leslie L). El problema que tiene esta fórmula es que es engorroso calcular las potencias L^k , para valores de k grandes. Más adelante estudiaremos diagonalización de matrices, lo cual nos permitirá simplificar el cálculo anterior y poder predecir el comportamiento futuro de la población.

Observar que en lo anterior trabajamos con una población dividida en cuatro clases de edad, pero es claro que el razonamiento se adapta a cualquier cantidad de clases. Por ejemplo, si tenemos solo dos clases de edad, entonces la fórmula (9) queda en

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ b_1 & 0 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ y_{k-1} \end{pmatrix}.$$

3. Determinantes

En este tema las matrices son cuadradas.

3.1. Definición

El *determinante* de una matriz $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ es un número que escribimos

$$\det(A) = |A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

y que para matrices 2×2 y 3×3 se define por lo siguiente.

$$\begin{aligned} \begin{vmatrix} a & b \\ c & d \end{vmatrix} &:= ad - bc. \\ \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &:= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \quad (11)$$

Método de Sarrus. Es un método que solo sirve para calcular determinantes de matrices 3×3 . Consiste en construir una tabla repitiendo las dos primeras filas de la matriz debajo de ella misma

$$\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}$$

El determinante se obtiene sumando los productos de las diagonales que van de izquierda a derecha y de arriba hacia abajo, y luego restando los productos de las diagonales que van de derecha a izquierda y de arriba hacia abajo. El resultado queda

$$\det(A) = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}.$$

que coincide con la fórmula (11).

Observación 3.1. El método de Sarrus no es el mejor método para calcular determinantes. En la gran mayoría de los casos hay formas más rápidas y eficientes (ver la observación 3.7 y los ejemplos siguientes).

La fórmula general para el determinante de una matriz $n \times n$ se define recursivamente, generalizando el método que usamos para definir la fórmula en el caso $n = 3$. Para la misma necesitamos introducir los cofactores, que definimos a continuación.

Consideremos $A = (a_{ij})$ una matriz $n \times n$. Para cada $i, j = 1, \dots, n$, sea \tilde{A}_{ij} la matriz $(n-1) \times (n-1)$ obtenida suprimiendo la fila i y la columna j de A . El *cofactor* correspondiente al lugar (i, j) de la matriz A es el número $\Delta_{ij} := (-1)^{i+j} \det(\tilde{A}_{ij})$.

Por ejemplo, para una matriz $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, los cofactores Δ_{11} , Δ_{12} y Δ_{13} , son

$$\begin{aligned} \Delta_{11} &= \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}, & \Delta_{12} &= - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -a_{21}a_{33} + a_{23}a_{31}, \\ \Delta_{13} &= \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{21}a_{32} - a_{22}a_{31}. \end{aligned}$$

Notar que la fórmula (11) se puede escribir usando cofactores

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\Delta_{11} + a_{12}\Delta_{12} + a_{13}\Delta_{13}.$$

Esta es la fórmula que vamos a generalizar. Si $A = (a_{ij}) \in M_n$, entonces el *determinante* de A se define en forma recursiva mediante

$$\det(A) := a_{11}\Delta_{11} + a_{12}\Delta_{12} + \cdots + a_{1n}\Delta_{1n}. \quad (12)$$

Si A es una matriz $n \times n$, entonces se dice que $\det(A)$ es un determinante de *orden* n . La fórmula anterior nos dice que para calcular un determinante de orden n , tenemos que calcular n determinantes de orden $n - 1$. Luego, como sabemos calcular determinantes de orden 3, entonces podemos calcular determinantes de orden 4, sabiendo estos podemos calcular los de orden 5, y siguiendo así podemos calcular determinantes de orden n , para cualquier entero positivo n .

Ejemplo 3.2. El determinante de la matriz identidad I_n vale 1, para todo n . Esto se prueba por inducción en n , usando la fórmula (12).

3.2. Técnicas de cálculo

A continuación veremos algunas propiedades que nos permiten simplificar los cálculos.

Proposición 3.3. *Casos en que el determinante es nulo.*

1. Si una matriz tiene una fila o columna formada solo por ceros, entonces su determinante es nulo.
2. Si una matriz tiene dos filas o columnas iguales, entonces su determinante es nulo. □

Proposición 3.4. *Propiedades del determinante.*

1. Si en una matriz se intercambian dos filas o dos columnas, entonces su determinante se mantiene igual en valor absoluto, pero cambia de signo.
2. Si en una matriz multiplicamos una fila o columna por una constante, entonces su determinante también se multiplica por dicha constante. Es decir

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ ca_{i1} & \cdots & ca_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & \cdots & ca_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & ca_{nj} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

3. Si en la fila i o columna j de A se cumple que todos los elementos salvo el a_{ij} son nulos, entonces $\det(A) = a_{ij}\Delta_{ij}$. □

Corolario 3.5. Si A es una matriz triangular superior o inferior, entonces su determinante se obtiene multiplicando los elementos de la diagonal principal. En particular esto vale para las matrices diagonales. Es decir, vale

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

Dem. Aplicar la propiedad 3 de la proposición 3.4 al coeficiente a_{11} , luego al a_{22} y seguir así hasta llegar al a_{nn} . □

Proposición 3.6. Si en una matriz le sumamos a una fila un múltiplo de otra fila, o a una columna un múltiplo de otra columna, entonces su determinante no varía. \square

Observación 3.7. La técnica usual para calcular determinantes consiste en aplicar reiteradamente la proposición anterior, hasta llegar al determinante de una matriz que tiene solo un elemento no nulo en una fila o columna. Luego se reduce su orden mediante la propiedad 3 de la proposición 3.4 (usando también las otras propiedades para ir simplificando los cálculos). Aplicando reiteradamente este procedimiento terminamos reduciendo el cálculo de un determinante de orden n al cálculo de uno de orden 2.

Ejemplos 3.8. En lo que sigue, si por ejemplo escribimos $F_2 - F_1$ quiere decir que a la fila 2 le estamos restando la fila 1, análogamente $C_3 + 2C_4$ quiere decir que a la columna 3 le estamos sumando la columna 4 multiplicada por 2. Notar que en las sumas o restas anteriores siempre escribimos primero la fila o columna que es modificada. Solo vamos a aclarar este tipo de operaciones, identificar las otras queda a cargo del lector.

1.

$$\begin{vmatrix} 1 & 5 & 7 \\ 0 & 3 & 1 \\ 2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 5 & 7 \\ 0 & 3 & 1 \\ 0 & -8 & -11 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ -8 & -11 \end{vmatrix} = -33 + 8 = -25.$$

El primer paso fue $F_3 - 2F_1$.

2.

$$\begin{vmatrix} 2 & 3 & 5 \\ 6 & 10 & 4 \\ 50 & 20 & 30 \end{vmatrix} = 2 \begin{vmatrix} 2 & 3 & 5 \\ 3 & 5 & 2 \\ 50 & 20 & 30 \end{vmatrix} = 20 \begin{vmatrix} 2 & 3 & 5 \\ 3 & 5 & 2 \\ 5 & 2 & 3 \end{vmatrix} = 20 \begin{vmatrix} 2 & 3 & 5 \\ 1 & 2 & -3 \\ 5 & 2 & 3 \end{vmatrix} = 20 \begin{vmatrix} 0 & -1 & 11 \\ 1 & 2 & -3 \\ 5 & 2 & 3 \end{vmatrix} \\ = 20 \begin{vmatrix} 0 & -1 & 11 \\ 1 & 2 & -3 \\ 0 & -8 & 18 \end{vmatrix} = -20 \begin{vmatrix} -1 & 11 \\ -8 & 18 \end{vmatrix} = -40 \begin{vmatrix} -1 & 11 \\ -4 & 9 \end{vmatrix} = -1400.$$

El tercer paso es $F_2 - F_1$ (para obtener $a_{21} = 1$), el cuarto $F_1 - 2F_2$ y el quinto $F_3 - 5F_2$.

3.

$$\begin{vmatrix} 1 & -2 & 1 & 3 \\ 5 & 7 & 8 & 9 \\ 2 & 0 & -2 & 4 \\ 3 & 0 & -3 & 6 \end{vmatrix} = 6 \begin{vmatrix} 1 & -2 & 1 & 3 \\ 5 & 7 & 8 & 9 \\ 1 & 0 & -1 & 2 \\ 1 & 0 & -1 & 2 \end{vmatrix} = 0.$$

El determinante final es nulo por tener dos filas iguales.

4.

$$\begin{vmatrix} 2 & 1 & 0 & 1 \\ -1 & 2 & 1 & 2 \\ 0 & -1 & 2 & 4 \\ 1 & 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 0 & 5 & 2 & 5 \\ -1 & 2 & 1 & 2 \\ 0 & -1 & 2 & 4 \\ 0 & 5 & 2 & 4 \end{vmatrix} = -(-1) \begin{vmatrix} 5 & 2 & 5 \\ -1 & 2 & 4 \\ 5 & 2 & 4 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ -1 & 2 & 4 \\ 5 & 2 & 4 \end{vmatrix} = \begin{vmatrix} -1 & 2 \\ 5 & 2 \end{vmatrix} = -12.$$

El primer paso es $F_1 + 2F_2$ y $F_4 + F_2$ (dos cambios al mismo tiempo), luego $F_1 - F_3$.

5.

$$\begin{aligned} \begin{vmatrix} 2 & 9 & 5 & 7 \\ 4 & 2 & 3 & 2 \\ 7 & 4 & 2 & 3 \\ 9 & 6 & 3 & 2 \end{vmatrix} &= \begin{vmatrix} 2 & 9 & 5 & 7 \\ 4 & 2 & 3 & 2 \\ 7 & 4 & 2 & 3 \\ 5 & 4 & 0 & 0 \end{vmatrix} = \begin{vmatrix} -7 & 9 & 5 & 7 \\ 2 & 2 & 3 & 2 \\ 3 & 4 & 2 & 3 \\ 1 & 4 & 0 & 0 \end{vmatrix} = \begin{vmatrix} -7 & 37 & 5 & 7 \\ 2 & -6 & 3 & 2 \\ 3 & -8 & 2 & 3 \\ 1 & 0 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 37 & 5 & 7 \\ -6 & 3 & 2 \\ -8 & 2 & 3 \end{vmatrix} \\ &= - \begin{vmatrix} 37 & 5 & 2 \\ -6 & 3 & -1 \\ -8 & 2 & 1 \end{vmatrix} = - \begin{vmatrix} 25 & 11 & 0 \\ -6 & 3 & -1 \\ -14 & 5 & 0 \end{vmatrix} = - \begin{vmatrix} 25 & 11 \\ -14 & 5 \end{vmatrix} = 279. \end{aligned}$$

El primer paso es $F_4 - F_2$, luego $C_1 - C_2$ (para obtener $a_{14} = 1$), luego $C_2 - 4C_1$, reducimos a orden 3, $C_3 - C_2$, $F_1 + 2F_2$ y $F_3 + F_2$ (dos cambios al mismo tiempo).

Puede haber casos en que el método anterior sea difícil de aplicar, por ejemplo cuando no nos damos cuenta de cómo “hacer ceros” mediante la proposición 3.6. En esos casos se puede usar el resultado siguiente, que si bien requiere más cálculos, tiene la ventaja de que siempre es aplicable.

Proposición 3.9. Sea $A = (a_{ij}) \in M_n$.

1. Para cada $i = 1, \dots, n$, el determinante se puede calcular desarrollando a partir de la fila i ,

$$\det(A) = a_{i1}\Delta_{i1} + a_{i2}\Delta_{i2} + \dots + a_{in}\Delta_{in}.$$

2. Para cada $j = 1, \dots, n$, el determinante se puede calcular desarrollando a partir de la columna j ,

$$\det(A) = a_{1j}\Delta_{1j} + a_{2j}\Delta_{2j} + \dots + a_{nj}\Delta_{nj}. \quad \square$$

3.3. Propiedades

La siguiente proposición resume otras propiedades de los determinantes.

Proposición 3.10. 1. Para toda matriz cuadrada A , vale $\det(A^t) = \det(A)$.

2. El determinante es aditivo respecto a cada fila y a cada columna.

$$\begin{aligned} \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + a'_{i1} & \dots & a_{in} + a'_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} &= \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a'_{i1} & \dots & a'_{in} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}, \\ \begin{vmatrix} a_{11} & \dots & a_{1j} + a'_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} + a'_{nj} & \dots & a_{nn} \end{vmatrix} &= \begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & a'_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a'_{nj} & \dots & a_{nn} \end{vmatrix}. \quad \square \end{aligned}$$

Observación 3.11. Hay que tener cuidado porque el determinante no se comporta bien respecto a la suma de matrices, es decir en general es $\det(A + B) \neq \det(A) + \det(B)$. Respecto al producto de una matriz por un escalar, vale

$$\det(cA) = c^n \det(A), \quad \forall c \in \mathbb{R}, A \in M_n.$$

Sin embargo el determinante preserva el producto de matrices, es decir, vale el siguiente resultado.

Proposición 3.12. Para todo $A, B \in M_n$ vale $\det(AB) = \det(A)\det(B)$. □

3.4. Aplicación a sistemas $n \times n$

Consideremos un sistema cuadrado

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases} . \quad (13)$$

Sea Δ el determinante de la matriz formada por los coeficientes del sistema anterior:

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} .$$

El número Δ es el *determinante del sistema*. El interés en Δ viene dado por el siguiente resultado.

Teorema 3.13. *El sistema de ecuaciones (13) es compatible determinado si y solo si $\Delta \neq 0$.* \square

Este teorema se suele utilizar en aplicaciones teóricas o cuando el sistema tiene algún parámetro, dado que en los otros casos es más fácil escalarizar que calcular determinantes. Lo que nos va a ser útil es su aplicación a los sistemas homogéneos. Recordar que un sistema homogéneo siempre es compatible, dado que admite la solución trivial $x_1 = \cdots = x_n = 0$. El teorema anterior implica el siguiente resultado.

Corolario 3.14. *Consideremos un sistema homogéneo cuadrado*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = 0 \end{cases} . \quad (14)$$

Sea Δ el determinante del sistema. Entonces (14) admite soluciones no triviales si y solo si $\Delta = 0$. \square

En la sección siguiente veremos una aplicación importante de este corolario. A continuación mostramos un par de ejemplos de su aplicación.

Ejemplo 3.15. Consideremos el sistema

$$\begin{cases} 2x + 5y + 27z = 0 \\ 3x + 2y + 5z = 0 \\ 5x + 3y + 7z = 0 \end{cases} .$$

Su determinante es $\Delta = \begin{vmatrix} 2 & 5 & 27 \\ 3 & 2 & 5 \\ 5 & 3 & 7 \end{vmatrix} = 11 \neq 0$. Luego la única solución que admite es la trivial $x = y = z = 0$.

Ejemplo 3.16. Consideremos el sistema

$$\begin{cases} ax + y = 0 \\ x + ay = 0 \end{cases} .$$

en que a es un parámetro. Nos interesa saber para qué valores de a el sistema admite soluciones no triviales.

El determinante del sistema es $\Delta = \begin{vmatrix} a & 1 \\ 1 & a \end{vmatrix} = a^2 - 1$. Luego el sistema admite soluciones no triviales si y solo si $a^2 - 1 = 0$, lo cual equivale a $a = \pm 1$. Para hallar las soluciones tenemos que resolver los sistemas obtenidos sustituyendo a por 1 y por -1 .

$$a = 1 : \begin{cases} x + y = 0 \\ x + y = 0 \end{cases} ; \quad a = -1 : \begin{cases} -x + y = 0 \\ x - y = 0 \end{cases} .$$

Las soluciones para el caso $a = 1$ son $x = \lambda$, $y = -\lambda$ y para el caso $a = -1$ son $x = y = \lambda$, siendo λ libre.

3.5. Matriz inversa

Definición 3.17. Decimos que una matriz cuadrada $A \in M_n$ es *invertible* si existe una matriz $B \in M_n$ tal que $AB = BA = I$, siendo I la matriz identidad.

Proposición 3.18. Si $A \in M_n$ es invertible, entonces la matriz $B \in M_n$ que verifica $AB = BA = I$ es única.

Dem. Supongamos que existen dos matrices $B, C \in M_n$ tales que $AB = BA = I$ y $AC = CA = I$. Entonces

$$C = CI = C(AB) = (CA)B = IB = B, \quad \text{luego } C = B. \quad \square$$

Por la proposición anterior, si la matriz $A \in M_n$ es invertible entonces existe una única matriz B que verifica $AB = BA = I$; esta matriz se llama la *inversa* de A y se escribe $B = A^{-1}$. Luego la inversa A^{-1} de A (en caso de existir) queda caracterizada por ser la única matriz que verifica $AA^{-1} = A^{-1}A = I$.

En lo que sigue consideraremos las dos preguntas siguientes. ¿Cómo saber si una matriz es invertible? Si sabemos que una matriz es invertible, ¿cómo se halla su inversa? El siguiente resultado responde la primera.

Teorema 3.19. Una matriz $A \in M_n$ es invertible si y solo si $\det(A) \neq 0$. En caso de ser invertible, vale $\det(A^{-1}) = \frac{1}{\det A}$.

Dem. Si A es invertible, entonces

$$AA^{-1} = I \quad \Rightarrow \quad \det(AA^{-1}) = \det I \quad \Rightarrow \quad \det A \cdot \det(A^{-1}) = 1.$$

Esto implica $\det A \neq 0$ y $\det(A^{-1}) = \frac{1}{\det A}$. También vale que si $\det(A) \neq 0$, entonces A es invertible, pero la prueba es difícil y la omitiremos. \square

Ahora veremos de contestar la segunda pregunta. Para matrices 2×2 tenemos el siguiente resultado.

Proposición 3.20. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz arbitraria. Entonces A es invertible si y solo si $ad - bc \neq 0$. Si $ad - bc \neq 0$, entonces la inversa de A es

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Dem. La primera afirmación se deduce del teorema 3.19. Para ver que vale la segunda, solo hay que verificar que el producto de esas matrices da la matriz identidad, lo cual se deja como ejercicio. \square

Para matrices diagonales es fácil hallar la matriz inversa:

Proposición 3.21. Una matriz diagonal es invertible si y solo si todas las entradas de la diagonal principal son no nulas. En ese caso,

$$\text{si } A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}, \quad \text{entonces } A^{-1} = \begin{pmatrix} a_1^{-1} & 0 & \cdots & 0 \\ 0 & a_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n^{-1} \end{pmatrix}. \quad (15)$$

Dem. El determinante de A es $\det(A) = a_1 \cdots a_n$. Luego A es invertible si y solo si $\det A \neq 0$, lo cual equivale a que a_1, \dots, a_n sean no nulos. En ese caso, es fácil de probar que la matriz que aparece a la derecha en (15) es la inversa de A (es solo multiplicarlas). \square

Aplicación 3.1. La matriz identidad I es invertible y vale $I^{-1} = I$.

Método para obtener la matriz inversa. El teorema 3.19 nos permite saber si una matriz es invertible. En caso afirmativo, para hallar la matriz inversa existe un algoritmo que mostraremos a continuación. Para eso previamente necesitamos definir las operaciones elementales.

Definición 3.22. Las *operaciones elementales* en una matriz A son las siguientes:

1. Intercambiar dos filas o columnas de A .
2. Multiplicar una fila o columna de A por una constante no nula.
3. Sumarle a una fila o columna un múltiplo de otra fila o columna, respectivamente.

El algoritmo para hallar la inversa consiste en lo siguiente. Supongamos que tenemos una matriz $A \in M_n$ que sabemos es invertible y queremos hallar A^{-1} . Lo que hacemos es escribir la matriz A y a su derecha la matriz identidad $I \in M_n$. Luego, y esto es importante, elegimos si vamos a trabajar con columnas o con filas. Si por ejemplo decidimos trabajar con columnas, entonces solo vamos a poder seguir trabajando con columnas y no podemos trabajar con filas. De la misma forma, si empezamos trabajando con filas no podemos pasar a trabajar con columnas. Hecha la elección anterior, por ejemplo digamos que decidimos trabajar con columnas, entonces vamos a realizar operaciones elementales en las columnas de la matriz A , y cada vez que hagamos una operación en A la repetimos en la matriz I . La idea es realizar operaciones hasta transformar la matriz A en la identidad I , cuando se llegó a ese punto, la matriz I se transformó en A^{-1} .

A continuación veremos algunos ejemplos donde estudiamos la invertibilidad de matrices y calculamos las inversas. En lo que sigue usaremos las mismas notaciones de los ejemplos 3.8. Además, escribiremos $F_2 \leftrightarrow F_3$, para decir que intercambiamos la fila 2 con la fila 3, y análogamente para columnas.

Ejemplos 3.23. Sea $A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. Es un ejercicio el verificar que su determinante es $\det(A) = 1 \neq 0$, luego A es invertible. Para calcular su inversa tenemos que elegir primero si trabajar con columnas o con filas. En este caso trabajaremos con columnas.

$$\begin{array}{ccc|ccc|l}
 2 & 3 & 1 & 1 & 0 & 0 & \\
 -1 & 0 & -1 & 0 & 1 & 0 & C_1 - C_3 \\
 1 & 1 & 1 & 0 & 0 & 1 & \\
 \hline
 1 & 3 & 1 & 1 & 0 & 0 & \\
 0 & 0 & -1 & 0 & 1 & 0 & C_2 - C_3 \\
 0 & 1 & 1 & -1 & 0 & 1 & \\
 \hline
 1 & 2 & 1 & 1 & 0 & 0 & \\
 0 & 1 & -1 & 0 & 1 & 0 & C_3 - C_1 \\
 0 & 0 & 1 & -1 & -1 & 1 & \\
 \hline
 1 & 2 & 0 & 1 & 0 & -1 & \\
 0 & 1 & -1 & 0 & 1 & 0 & C_2 - 2C_1 \\
 0 & 0 & 1 & -1 & -1 & 2 & \\
 \hline
 1 & 0 & 0 & 1 & -2 & -1 & \\
 0 & 1 & -1 & 0 & 1 & 0 & C_3 + C_2 \\
 0 & 0 & 1 & -1 & 1 & 2 & \\
 \hline
 1 & 0 & 0 & 1 & -2 & -3 & \\
 0 & 1 & 0 & 0 & 1 & 1 & \\
 0 & 0 & 1 & -1 & 1 & 3 &
 \end{array}$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 1 \\ -1 & 1 & 3 \end{pmatrix}$.

Sea $A = \begin{pmatrix} 2 & 2 & 1 \\ -2 & -1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$. Dejamos como ejercicio el probar que vale $\det(A) = 1$, luego A es invertible.

Mostraremos cómo hallar su inversa operando con las filas.

2	2	1	1	0	0	$F_1 \leftrightarrow F_3$
-2	-1	-2	0	1	0	
1	0	1	0	0	1	
1	0	1	0	0	1	$F_2 + 2F_1$
-2	-1	-2	0	1	0	
2	2	1	1	0	0	
1	0	1	0	0	1	$F_3 - 2F_1$
0	-1	0	0	1	2	
2	2	1	1	0	0	
1	0	1	0	0	1	$F_3 + 2F_2$
0	-1	0	0	1	2	
0	2	-1	1	0	-2	
1	0	1	0	0	1	$(-1)F_2$
0	-1	0	0	1	2	
0	0	-1	1	2	2	
1	0	1	0	0	1	$(-1)F_3$
0	1	0	0	-1	-2	
0	0	-1	1	2	2	
1	0	1	0	0	1	$F_1 - F_3$
0	1	0	0	-1	-2	
0	0	1	-1	-2	-2	
1	0	0	1	2	3	$F_1 - F_3$
0	1	0	0	-1	-2	
0	0	1	-1	-2	-2	

Luego la inversa es $A^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ -1 & -2 & -2 \end{pmatrix}$.

Sea $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 3 \end{pmatrix}$. Vale $\det(A) = 6$, luego A es invertible. Trabajaremos con columnas.

$$\begin{array}{ccc|ccc|l}
 3 & 1 & 1 & 1 & 0 & 0 & \\
 2 & 3 & 2 & 0 & 1 & 0 & C_1 - C_3 \\
 3 & 3 & 3 & 0 & 0 & 1 & \\
 \hline
 2 & 1 & 1 & 1 & 0 & 0 & \\
 0 & 3 & 2 & 0 & 1 & 0 & C_2 - C_3 \\
 0 & 3 & 3 & -1 & 0 & 1 & \\
 \hline
 2 & 0 & 1 & 1 & 0 & 0 & \\
 0 & 1 & 2 & 0 & 1 & 0 & \frac{1}{2}C_1 \\
 0 & 0 & 3 & -1 & -1 & 1 & \\
 \hline
 1 & 0 & 1 & 1/2 & 0 & 0 & \\
 0 & 1 & 2 & 0 & 1 & 0 & C_3 - C_1 \\
 0 & 0 & 3 & -1/2 & -1 & 1 & \\
 \hline
 1 & 0 & 0 & 1/2 & 0 & -1/2 & \\
 0 & 1 & 2 & 0 & 1 & 0 & C_3 - 2C_2 \\
 0 & 0 & 3 & -1/2 & -1 & 3/2 & \\
 \hline
 1 & 0 & 0 & 1/2 & 0 & -1/6 & \\
 0 & 1 & 0 & 0 & 1 & -2/3 & \frac{1}{3}C_3 \\
 0 & 0 & 1 & -1/2 & -1 & 7/6 &
 \end{array}$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1/2 & 0 & -1/6 \\ 0 & 1 & -2/3 \\ -1/2 & -1 & 7/6 \end{pmatrix}$.

- Observaciones 3.24.*
1. Notar que en los ejemplos anteriores lo que hicimos primero fue transformar la matriz A en una matriz triangular (triangular superior en todos los casos, pero podría ser también triangular inferior) y luego llevarla a una forma diagonal. Es importante verificar siempre que la matriz hallada es realmente la matriz inversa (haciendo el producto), dado que es muy fácil cometer errores.
 2. Conviene remarcar que cuando aplicamos el algoritmo anterior, tenemos que elegir trabajar con filas o columnas, pero no podemos mezclar, porque si trabajamos con filas y columnas, cuando a la izquierda lleguemos a la matriz identidad, la matriz de la derecha no va a ser necesariamente la matriz inversa.

Finalizamos esta sección con algunas propiedades de las matrices invertibles. Recordar que una matriz A es invertible si existe una matriz B tal que $AB = I$ y $BA = I$, siendo I la matriz identidad. El primer resultado que veremos es que para que una matriz sea invertible alcanza con que se verifique una de esas dos igualdades⁴.

Proposición 3.25. *Sea A una matriz. Si existe una matriz B tal que $AB = I$, entonces A es invertible y $B = A^{-1}$. Lo mismo sucede si existe una matriz B tal que $BA = I$.*

Dem. Si $AB = I$, entonces $\det(AB) = \det(I)$, lo cual equivale a $\det(A)\det(B) = 1$. Esto implica que necesariamente es $\det(A) \neq 0$ y por lo tanto A es invertible. Luego multiplicando por la izquierda por A^{-1} obtenemos

$$AB = I \Rightarrow A^{-1}(AB) = A^{-1}I \Rightarrow (A^{-1}A)B = A^{-1} \Rightarrow IB = A^{-1} \Rightarrow B = A^{-1}.$$

Lo mismo sucede si vale $BA = I$ (pero en la cuenta anterior hay que multiplicar por la derecha por A^{-1}). \square

⁴Esto no es obvio, dado que el producto de matrices no verifica la propiedad conmutativa.

Proposición 3.26. Si A y B son matrices invertibles, entonces su producto AB es invertible y vale $(AB)^{-1} = B^{-1}A^{-1}$.

Dem. Multipliquemos AB por $B^{-1}A^{-1}$:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I,$$

luego aplicando la proposición anterior a la matriz AB obtenemos la tesis. \square

Proposición 3.27. Sea A una matriz invertible.

1. Si X y B son matrices tales que $AX = B$, entonces $X = A^{-1}B$.
2. Si X y B son matrices tales que $XA = B$, entonces $X = BA^{-1}$.

Dem. Las dos fórmulas se prueban igual, así que probaremos solo la primera.

$$AX = B \Rightarrow A^{-1}(AX) = A^{-1}B \Rightarrow (A^{-1}A)X = A^{-1}B \Rightarrow IX = A^{-1}B \Rightarrow X = A^{-1}B. \quad \square$$

El siguiente resultado muestra que vale la propiedad cancelativa del producto de matrices, con la condición de que la matriz que multiplica a ambos lados sea invertible.

Corolario 3.28. Sean A, B, C matrices tales que A es invertible.

1. Si $AB = O$ o $BA = O$, entonces $B = O$ (O es la matriz nula).
2. Si $AB = AC$ o $BA = CA$, entonces $B = C$.

Dem. Probaremos solo los primeros casos de cada ítem, los otros se prueban en forma análoga. Si $AB = O$, entonces $B = A^{-1}O = O$. Si $AB = AC$, entonces $B = A^{-1}(AC) = (A^{-1}A)C = IC = C$. \square

4. Diagonalización

En esta sección las matrices son cuadradas. Cuando estudiamos el modelo de Leslie, vimos que el vector de distribución de edades X_k se obtiene a partir del vector inicial X_0 mediante la fórmula $X_k = L^k X_0$, siendo L^k la potencia k -ésima de la matriz de Leslie L . El problema que tiene esta fórmula es la dificultad de calcular L^k para valores grandes de k . En lo que sigue estudiaremos la técnica de diagonalización, que en particular permite simplificar el cálculo anterior. Luego veremos algunas aplicaciones.

Una matriz cuadrada $A \in M_n$ se dice *diagonalizable* si existe una matriz diagonal $D \in M_n$ y una matriz invertible $P \in M_n$ tales que $A = PDP^{-1}$.

El resultado siguiente muestra que el cálculo de potencias de matrices diagonalizables se reduce al de potencias de matrices diagonales, el cual es muy fácil de hacer (observación 2.7.4).

Proposición 4.1. Si $A = PDP^{-1}$, entonces $A^k = PD^kP^{-1}$, para todo $k = 0, 1, 2, \dots$

Dem.

$$A^2 = AA = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PDDP^{-1} = PD^2P^{-1}.$$

Repetiendo ese razonamiento se llega a la tesis (la prueba formal es por inducción completa). \square

En lo que sigue intentaremos responder las siguientes preguntas.

1. ¿Cómo determinamos si una matriz es diagonalizable?
2. Supongamos que A es diagonalizable, ¿cómo encuentro las matrices D y P tales que $A = PDP^{-1}$?

Empezaremos con un caso concreto y luego lo generalizaremos.

Ejemplo 4.2. Consideremos la matriz

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 5 & 1 & -5 \\ 5 & 0 & -3 \end{pmatrix}.$$

Empezamos calculando su *polinomio característico*, que se define mediante el siguiente determinante

$$\chi_A(\lambda) = \begin{vmatrix} 2 - \lambda & 0 & 0 \\ 5 & 1 - \lambda & -5 \\ 5 & 0 & -3 - \lambda \end{vmatrix}.$$

Operando obtenemos

$$\chi_A(\lambda) = \begin{vmatrix} 2 - \lambda & 0 & 0 \\ 5 & 1 - \lambda & -5 \\ 5 & 0 & -3 - \lambda \end{vmatrix} = (2 - \lambda) \begin{vmatrix} 1 - \lambda & -5 \\ 0 & -3 - \lambda \end{vmatrix} = -(2 - \lambda)(1 - \lambda)(3 + \lambda).$$

Si se desarrollan los productos obtenemos $\chi_A(\lambda) = -\lambda^3 + 7\lambda - 6$, luego $\chi_A(\lambda)$ es un polinomio de grado 3 en la variable λ . Volviendo a la expresión $\chi_A(\lambda) = -(2 - \lambda)(1 - \lambda)(3 + \lambda)$, vemos que las raíces de $\chi_A(\lambda)$ son 1, 2 y -3 . Ahora para cada uno de esos valores de λ vamos a resolver el siguiente sistema homogéneo

$$\begin{pmatrix} 2 - \lambda & 0 & 0 \\ 5 & 1 - \lambda & -5 \\ 5 & 0 & -3 - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} (2 - \lambda)x = 0 \\ 5x + (1 - \lambda)y - 5z = 0 \\ 5x + (-3 - \lambda)z = 0 \end{cases}.$$

Caso $\lambda = 1$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 5 & 0 & -5 \\ 5 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x = 0 \\ 5x - 5z = 0 \\ 5x - 4z = 0 \end{cases} \Leftrightarrow x = z = 0.$$

Como la y quedó libre, las soluciones son los vectores $(0, y, 0)$, con $y \in \mathbb{R}$. Lo que nos interesa es obtener una solución no trivial; nos quedamos con la más simple que es $(0, 1, 0)$.

Caso $\lambda = 2$.

$$\begin{pmatrix} 0 & 0 & 0 \\ 5 & -1 & -5 \\ 5 & 0 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 0 = 0 \\ 5x - y - 5z = 0 \\ 5x - 5z = 0 \end{cases} \Leftrightarrow \begin{cases} x = z \\ y = 0 \end{cases}.$$

Las soluciones son los vectores $(x, 0, x)$, con $x \in \mathbb{R}$. Elegimos el $(1, 0, 1)$.

Caso $\lambda = -3$.

$$\begin{pmatrix} 5 & 0 & 0 \\ 5 & 4 & -5 \\ 5 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 5x = 0 \\ 5x + 4y - 5z = 0 \\ 5x = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ 4y = 5z \end{cases}.$$

Las soluciones son los vectores $(0, y, \frac{4}{5}y)$, con $y \in \mathbb{R}$. Elegimos el $(0, 5, 4)$.

Con estos valores formamos las matrices D y P siguientes

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 5 \\ 0 & 1 & 4 \end{pmatrix}.$$

Notar que D es una matriz diagonal cuyas entradas diagonales $(1, 2, -3)$ son las raíces de $\chi_A(\lambda)$. Por otro lado las tres columnas de P son las soluciones no triviales que elegimos anteriormente (escritas verticalmente). Es importante el orden de las columnas de P , que viene dado por las entradas diagonales de D : la primera columna es la solución correspondiente a $\lambda = 1$, la segunda corresponde a $\lambda = 2$ y la tercera a $\lambda = -3$.

El determinante de P vale -4 . Como es $\det P \neq 0$, entonces P es invertible. Operando obtenemos $P^{-1} = \begin{pmatrix} 5/4 & 1 & -5/4 \\ 1 & 0 & 0 \\ -1/4 & 0 & 1/4 \end{pmatrix}$. Las matrices así obtenidas verifican $A = PDP^{-1}$, es decir

$$\begin{pmatrix} 2 & 0 & 0 \\ 5 & 1 & -5 \\ 5 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 5 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 5/4 & 1 & -5/4 \\ 1 & 0 & 0 \\ -1/4 & 0 & 1/4 \end{pmatrix}.$$

Lo que vimos en el ejemplo anterior se puede aplicar en general. Lo veremos para matrices 3×3 , pero vale para matrices de cualquier tamaño. Consideremos una matriz 3×3 genérica

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

El determinante

$$\chi_A(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} \quad (16)$$

es un polinomio de grado 3 en la variable λ llamado el *polinomio característico* de A . Las raíces del polinomio característico son los *valores propios* de A . Supongamos que $\chi_A(\lambda)$ tiene tres raíces λ_1, λ_2 y λ_3 . Con estos valores construimos la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Consideremos λ_1 y formemos la ecuación

$$\begin{pmatrix} a_{11} - \lambda_1 & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda_1 & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda_1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Desarrollando el producto e igualando término a término, obtenemos el siguiente sistema

$$\begin{cases} (a_{11} - \lambda_1)x + a_{12}y + a_{13}z = 0 \\ a_{21}x + (a_{22} - \lambda_1)y + a_{23}z = 0 \\ a_{31}x + a_{32}y + (a_{33} - \lambda_1)z = 0 \end{cases} \quad (17)$$

Las soluciones no triviales de este sistema son los *vectores propios* correspondientes al valor propio λ_1 . Sea (u_1, u_2, u_3) uno de esos vectores propios (es decir, $x = u_1, y = u_2$ y $z = u_3$ es una solución no trivial de

(17)). Si ahora consideramos el sistema (17) cambiando λ_1 por λ_2 , entonces obtendremos otro vector propio (v_1, v_2, v_3) , y repitiendo el procedimiento cambiando λ_1 por λ_3 y obtendremos un tercer vector propio (w_1, w_2, w_3) . Con estos tres vectores propios formamos la matriz

$$P = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}. \quad (18)$$

Los cálculos anteriores implican que las matrices P y D verifican⁵ $AP = PD$. Si la matriz P es invertible, entonces deducimos que vale $A = PDP^{-1}$ y por lo tanto A es diagonalizable.

En realidad, en estas condiciones la matriz P siempre es invertible, es decir, vale el siguiente resultado.

Teorema 4.3. *Sea A una matriz 3×3 . Supongamos que su polinomio característico $\chi_A(\lambda)$ tiene 3 raíces distintas $\lambda_1, \lambda_2, \lambda_3$. Sean (u_1, u_2, u_3) , (v_1, v_2, v_3) y (w_1, w_2, w_3) tres vectores propios correspondientes a $\lambda_1, \lambda_2, \lambda_3$. Entonces la matriz P definida en (18) es invertible. \square*

En conclusión, si $A \in M_3$ es tal que su polinomio característico $\chi_A(\lambda)$ tiene 3 raíces distintas, entonces A es diagonalizable.

Observaciones 4.4. 1. En el procedimiento anterior es muy importante el orden de los elementos. Si es $A = PDP^{-1}$, con

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix},$$

entonces (u_1, u_2, u_3) tiene que ser un vector propio correspondiente a λ_1 , (v_1, v_2, v_3) un vector propio correspondiente a λ_2 y (w_1, w_2, w_3) un vector propio correspondiente a λ_3 .

2. Las entradas diagonales de D son λ_1, λ_2 y λ_3 , que son las raíces de $\chi_A(\lambda)$. Luego la matriz D es única a menos de permutar λ_1, λ_2 y λ_3 .
3. Para construir la matriz P tomamos soluciones no triviales de sistemas homogéneos, en lo cual estamos haciendo una elección. Así que no hay unicidad de la matriz P .
4. El teorema 4.3 vale para matrices $n \times n$ en general, en el sentido de que si el polinomio característico de A tiene n raíces distintas, entonces la matriz P formada por los vectores propios es invertible.
5. Si el polinomio característico $\chi_A(\lambda)$ tiene menos de 3 raíces, entonces hay que trabajar con más cuidado, dado que hay casos en que es diagonalizable y casos en que no. Ejemplos de polinomios de este tipo son los siguientes

$$\lambda^3 + 3\lambda^2 + 3\lambda + 1 = (\lambda + 1)^3; \quad \lambda^3 + \lambda^2 - \lambda - 1 = (\lambda + 1)^2(\lambda - 1); \quad \lambda^3 + \lambda^2 + \lambda + 1 = (\lambda + 1)(\lambda^2 + 1).$$

El primero solo tiene una raíz que es -1 , el segundo tiene solo dos raíces ± 1 y el tercero solo tiene una raíz que es -1 . En general en estas notas evitaremos meternos en ese tipo de situaciones, y si lo hacemos lo resolveremos para ese caso particular, como en el modelo estocástico de la Sección 5.3.

⁵Esto no es inmediato. Para verificarlo tienen que multiplicar las matrices y ver que coinciden; para eso se usa que los vectores propios son soluciones de los sistemas (17).

Como comentamos anteriormente, la técnica de diagonalización se puede aplicar a matrices cuadradas de cualquier tamaño. A continuación veremos un ejemplo con una matriz 2×2 .

Ejemplo 4.5. Queremos saber si $A = \begin{pmatrix} 23 & -50 \\ 10 & -22 \end{pmatrix}$ es diagonalizable. Empezamos hallando su polinomio característico

$$\chi_A(\lambda) = \begin{vmatrix} 23 - \lambda & -50 \\ 10 & -22 - \lambda \end{vmatrix} = \lambda^2 - \lambda - 6.$$

Las raíces de $\chi_A(\lambda) = \lambda^2 - \lambda - 6$ son $\lambda_1 = -2$ y $\lambda_2 = 3$. Luego podemos tomar $D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}$ como matriz diagonal. Ahora tenemos que obtener soluciones no triviales de

$$\begin{pmatrix} 23 - \lambda & -50 \\ 10 & -22 - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

para $\lambda_1 = -2$ y $\lambda_2 = 3$.

Para $\lambda_1 = -2$ obtenemos

$$\begin{pmatrix} 25 & -50 \\ 10 & -20 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 25x - 50y = 0 \\ 10x - 20y = 0 \end{cases} \Leftrightarrow x - 2y = 0.$$

Tomando $y = 1$ y $x = 2$, obtenemos $(u_1, u_2) = (2, 1)$.

Para $\lambda_2 = 3$ obtenemos

$$\begin{pmatrix} 20 & -50 \\ 10 & -25 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 20x - 50y = 0 \\ 10x - 25y = 0 \end{cases} \Leftrightarrow 2x - 5y = 0.$$

Luego tomando $x = 5$ e $y = 2$, obtenemos $(v_1, v_2) = (5, 2)$. Con estos valores formamos la matriz

$$P = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix}.$$

La inversa de P es

$$P^{-1} = \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix}$$

y por lo tanto obtenemos

$$A = PDP^{-1} \Leftrightarrow \begin{pmatrix} 23 & -50 \\ 10 & -22 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix}.$$

Esto simplifica el operar con la matriz A . Por ejemplo, si queremos calcular A^{10} es

$$\begin{aligned} \begin{pmatrix} 23 & -50 \\ 10 & -22 \end{pmatrix}^{10} &= \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}^{10} \begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} (-2)^{10} & 0 \\ 0 & 3^{10} \end{pmatrix} \begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1024 & 0 \\ 0 & 59049 \end{pmatrix} \begin{pmatrix} -2 & 5 \\ 1 & -2 \end{pmatrix}. \end{aligned}$$

Para terminar de hallar A^{10} solo resta hacer el último producto, lo cual queda como ejercicio.

Observación 4.6. En el ejemplo anterior también podríamos haber tomado $\lambda_1 = 3$ y $\lambda_2 = -2$. En ese caso sería $D = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$. Esto implica que ahora es $P = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$ y por lo tanto $P^{-1} = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}$. El resto sigue igual.

5. Aplicaciones

5.1. Genética, herencia autosómica.

Cada rasgo a heredar por un individuo (puede ser humano, animal o vegetal) está regido por dos genes, que se suelen representar por a y A . La herencia se dice *autosómica* si viene ligada a un gen que está en un cromosoma autosómico (es decir, en un cromosoma que no es sexual). En estos casos cada individuo de la población tiene un par de estos genes, independientemente de su sexo, los cuales pueden ser AA , Aa o aa . Cada una de las tres posibilidades anteriores determina el *genotipo* del individuo. Un ejemplo de rasgo controlado por el genotipo es el color de algunos tipos de flores, en las cuales el par AA da flores rojas, el Aa rosadas y el aa blancas. También el color de los ojos queda determinado por el genotipo.

En lo que sigue supondremos que estamos en un caso de herencia autosómica. Luego cada progenitor tiene los dos genes y la herencia se obtiene heredando al azar un gen de la madre y uno del padre. Como el padre y la madre pueden tener genotipos AA , Aa o aa , entonces la probabilidad de cada uno de los posibles genotipos de un descendiente está dada por la siguiente tabla

	$AA \times AA$	$AA \times Aa$	$AA \times aa$	$Aa \times Aa$	$Aa \times aa$	$aa \times aa$
AA	1	0,5	0	0,25	0	0
Aa	0	0,5	1	0,5	0,5	0
aa	0	0	0	0,25	0,5	1

(19)

Por ejemplo, cuando se cruzan Aa con Aa (cuarta columna de la tabla), el resultado puede ser AA , Aa , aA o aa . Notar que Aa y aA es lo mismo. Luego de los 4 casos posibles, tenemos uno con AA , dos con Aa y uno con aa . Entonces la probabilidad de que salga AA es $1/4 = 0,25$, de Aa es $2/4 = 0,5$ y de aa es $1/4 = 0,25$.

Supongamos que tenemos una población de plantas con una distribución variada de los genotipos AA , Aa y aa . Decidimos empezar a cruzar esta población con plantas que solo sean AA . Vamos a estudiar cómo evoluciona la distribución de los genotipos a medida que pasa el tiempo.

Sean

- p_0 la fracción inicial de plantas que tienen genotipo AA ,
- q_0 la fracción inicial de plantas que tienen genotipo Aa ,
- r_0 la fracción inicial de plantas que tienen genotipo aa .

Con esos valores nos armamos el vector de distribución inicial de genotipos

$$Y_0 = \begin{pmatrix} p_0 \\ q_0 \\ r_0 \end{pmatrix}.$$

Los cruzamos con los de tipo AA y obtenemos una nueva generación, que tendrá un cierto vector de distribución de genotipos que llamamos Y_1 , cruzamos de nuevo y obtenemos otra generación que tendrá un cierto vector de distribución Y_2 y así seguimos. En la k -ésima generación obtendremos un vector de distribución

$$Y_k = \begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix}, \quad \forall k = 0, 1, 2, \dots$$

en el cual

- p_k es la fracción de plantas que tienen genotipo AA ,
- q_k es la fracción de plantas que tienen genotipo Aa ,
- r_k es la fracción de plantas que tienen genotipo aa .

Como cada planta tiene que ser de algún genotipo, es

$$p_k + q_k + r_k = 1, \quad \forall k = 0, 1, 2, \dots$$

Nos interesa saber cómo se relaciona el vector de distribución Y_{k-1} con el siguiente Y_k , siendo

$$Y_{k-1} = \begin{pmatrix} p_{k-1} \\ q_{k-1} \\ r_{k-1} \end{pmatrix}, \quad Y_k = \begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix}.$$

Supongamos que en la generación $(k-1)$ -ésima hay una cierta cantidad P de plantas, luego tendremos $p_{k-1}P$ plantas de genotipo AA , $q_{k-1}P$ plantas de genotipo Aa y $r_{k-1}P$ plantas de genotipo aa . Las cruzamos con P plantas de genotipo AA . La tabla (19) nos dice cuáles son las probabilidades del resultado, para cada genotipo. Con genotipo AA vamos a obtener 1 por cada AA que teníamos anteriormente más $1/2$ por cada Aa que teníamos anteriormente, luego obtenemos

$$p_{k-1}P + \frac{1}{2}q_{k-1}P = \left(p_{k-1} + \frac{1}{2}q_{k-1} \right) P$$

plantas de genotipo AA . Con genotipo Aa vamos a obtener $1/2$ por cada Aa y 1 por cada aa , luego obtenemos

$$\frac{1}{2}q_{k-1}P + r_{k-1}P = \left(\frac{1}{2}q_{k-1} + r_{k-1} \right) P$$

plantas de genotipo Aa . De genotipo aa no vamos a obtener ninguno (porque estamos cruzando con AA). Dividiendo la cantidad de plantas de cada genotipo por la cantidad total de plantas P , obtenemos las proporciones de cada genotipo. Luego de lo anterior deducimos

$$p_k = p_{k-1} + \frac{1}{2}q_{k-1}, \quad q_k = \frac{1}{2}q_{k-1} + r_{k-1}, \quad r_k = 0.$$

Escribiendo las fórmulas anteriores usando los vectores de distribución obtenemos

$$\begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix} = \begin{pmatrix} p_{k-1} + \frac{1}{2}q_{k-1} \\ \frac{1}{2}q_{k-1} + r_{k-1} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_{k-1} \\ q_{k-1} \\ r_{k-1} \end{pmatrix}.$$

Es decir, $Y_k = MY_{k-1}$, siendo

$$M = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Por la misma razón que con el modelo de Leslie, deducimos que vale $Y_k = M^k Y_0$, para todo $k = 1, 2, \dots$. Luego para cada valor de k , el vector de distribución Y_k queda determinado conociendo M^k e Y_0 .

Para obtener M^k vamos a aplicar diagonalización. El polinomio característico de M es

$$\chi_M(\lambda) = \begin{vmatrix} 1 - \lambda & 1/2 & 0 \\ 0 & 1/2 - \lambda & 1 \\ 0 & 0 & -\lambda \end{vmatrix} = (1 - \lambda)(1/2 - \lambda)(-\lambda) = -(\lambda - 1)(\lambda - 1/2)\lambda.$$

Luego los valores propios de M son $\lambda_1 = 1$, $\lambda_2 = 1/2$ y $\lambda_3 = 0$. Operando obtenemos los vectores propios respectivos

$$V_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad V_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad V_3 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Luego es $M = PDP^{-1}$, siendo

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P = P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Las fórmulas anteriores⁶ nos permiten calcular fácilmente las potencias de M .

$$M^k = PD^kP^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2^k & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 - \frac{1}{2^k} & 1 - \frac{1}{2^{k-1}} \\ 0 & \frac{1}{2^k} & \frac{1}{2^{k-1}} \\ 0 & 0 & 0 \end{pmatrix}.$$

Luego usando $Y_k = M^k Y_0$ obtenemos Y_k :

$$\begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix} = \begin{pmatrix} 1 & 1 - \frac{1}{2^k} & 1 - \frac{1}{2^{k-1}} \\ 0 & \frac{1}{2^k} & \frac{1}{2^{k-1}} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_0 \\ q_0 \\ r_0 \end{pmatrix} = \begin{pmatrix} p_0 + \left(1 - \frac{1}{2^k}\right)q_0 + \left(1 - \frac{1}{2^{k-1}}\right)r_0 \\ \frac{1}{2^k}q_0 + \frac{1}{2^{k-1}}r_0 \\ 0 \end{pmatrix}.$$

Como sabemos que vale $p_0 + q_0 + r_0 = 1$, de la igualdad anterior deducimos

$$\begin{aligned} p_k &= 1 - \frac{1}{2^k}q_0 - \frac{1}{2^{k-1}}r_0, \\ q_k &= \frac{1}{2^k}q_0 + \frac{1}{2^{k-1}}r_0, \\ r_k &= 0. \end{aligned}$$

Como $\lim_{k \rightarrow +\infty} \frac{1}{2^k} = 0$, deducimos que los genotipos de la población tienden a la distribución

$$p = 1, \quad q = r = 0.$$

Es decir, a medida que pasa el tiempo todas las plantas tienden a ser del genotipo AA .

5.2. Modelo de Leslie (segunda parte)

Recordar que el modelo de Leslie parte de una población dividida en n clases de edad, en la cual todas las clases tienen la misma duración de l años. Medimos la cantidad de hembras de la población cada $t_k = kl$ años ($k = 0, 1, \dots$). La cantidad de hembras de la primera clase en el instante t_k la escribimos x_k , la cantidad

⁶Notar que en este caso se dió una situación curiosa, que es que vale $P^{-1} = P$. Esto es muy atípico y casi nunca ocurre.

de las de la segunda clase y_k , etc. Con estos números armamos el vector de distribución de edades X_k y el vector de distribución inicial de edades X_0 ,

$$X_k = \begin{pmatrix} x_k \\ y_k \\ \vdots \end{pmatrix}, \quad k = 1, 2, \dots \quad X_0 = \begin{pmatrix} x_0 \\ y_0 \\ \vdots \end{pmatrix}.$$

Llamamos a_i a la cantidad de hijas promedio que nacen de una sola hembra durante el tiempo que está en la clase i , y b_i a la fracción de la cantidad de hembras en la clase i , que se espera que sobrevivan y pasen a la clase $i + 1$. Es $a_i \geq 0$ y $0 < b_i \leq 1$, para todo i . También asumiremos que para algún i es $a_i > 0$, dado que alguna clase tiene que ser fértil. Con estos números construimos la matriz de Leslie de la población

$$L = \begin{pmatrix} a_1 & a_2 & \dots & a_{n-1} & a_n \\ b_1 & 0 & \dots & 0 & 0 \\ 0 & b_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & b_{n-1} & 0 \end{pmatrix}. \quad (20)$$

El pasaje de X_{k-1} a X_k viene dado por $X_k = LX_{k-1}$, de donde se deduce

$$X_k = L^k X_0, \quad \forall k = 1, 2, \dots \quad (21)$$

Nos interesa saber cómo va a evolucionar la población con el paso del tiempo. Un primer indicador es la *tasa neta de reproducción*, que es el número R definido por

$$R = a_1 + a_2 b_1 + a_3 b_1 b_2 + \dots + a_n b_1 b_2 \dots b_{n-1}.$$

Notar que a_1 es el promedio de hijas que tiene una hembra cuando está en la primera clase, $a_2 b_1$ es el promedio que tiene cuando está en la segunda, $a_3 b_1 b_2$ es el promedio que tiene cuando está en la tercera, etc. Luego R puede interpretarse como el promedio de hijas que tiene una hembra durante toda su vida. Pensado de esa forma, el siguiente resultado es bastante natural.

Teorema 5.1. ■ *Si $R > 1$, entonces la población tiende a crecer exponencialmente.*

- *Si $R < 1$, entonces la población tiende a decrecer exponencialmente (se extingue).*
- *Si $R = 1$, entonces la población tiende a estabilizarse.* □

Más adelante veremos una justificación del resultado anterior. A continuación aplicaremos diagonalización para profundizar en el modelo de Leslie. Para que la lectura sea más ágil, en lo que sigue van a haber pocas demostraciones, pero algunas de las pruebas aparecen en la sección siguiente.

Nos interesa conocer los valores y vectores propios de L . El siguiente teorema nos simplifica los cálculos.

Teorema 5.2. 1. *El polinomio característico de L es*

$$\chi_L(\lambda) = (-1)^n (\lambda^n - a_1 \lambda^{n-1} - a_2 b_1 \lambda^{n-2} - a_3 b_1 b_2 \lambda^{n-3} - \dots - a_n b_1 b_2 \dots b_{n-1}).$$

2. *Toda matriz de Leslie L tiene un único valor propio positivo λ_1 .*

3. *Si $\lambda \neq 0$ es un valor propio de L , entonces $V = \begin{pmatrix} 1 \\ \frac{b_1}{\lambda} \\ \frac{b_1 b_2}{\lambda^2} \\ \vdots \\ \frac{b_1 b_2 \dots b_{n-1}}{\lambda^{n-1}} \end{pmatrix}$ es un vector propio correspondiente.* □

Vale el siguiente resultado.

Teorema 5.3. Si λ_1 es el valor propio positivo de la matriz de Leslie L y $\lambda_i \leq 0$ es otro valor propio cualquiera de L , entonces el valor absoluto de λ_i es siempre menor o igual que λ_1 . \square

Por el teorema anterior al valor propio positivo λ_1 se le llama el *valor propio dominante* de L . Si además vale $|\lambda_i| < \lambda_1$ para todo otro valor propio λ_i de L , entonces decimos que λ_1 es *estrictamente dominante*.

El siguiente resultado da una condición para que el valor propio positivo sea estrictamente dominante.

Teorema 5.4. Si en una población hay dos clases de edad consecutivas fértiles, es decir si en su matriz de Leslie L existe algún i tal que a_i y a_{i+1} son no nulos, entonces λ_1 es estrictamente dominante. \square

En lo que sigue veremos algunas consecuencias de las propiedades anteriores. De ahora en adelante supondremos que el valor propio λ_1 es estrictamente dominante y que L es diagonalizable (esto último no es necesario pero simplifica los cálculos). Además, para simplificar la escritura supondremos que es $n = 3$, aunque los resultados valen en general. La matriz de Leslie es

$$L = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & 0 & 0 \\ 0 & b_2 & 0 \end{pmatrix}. \quad (22)$$

Sean $\lambda_1 > 0$, $\lambda_2 \leq 0$ y $\lambda_3 \leq 0$ los vectores propios y V_1 , V_2 y V_3 los vectores propios correspondientes. Si escribimos

$$V_1 = \begin{pmatrix} v_{11} \\ v_{21} \\ v_{31} \end{pmatrix}, \quad V_2 = \begin{pmatrix} v_{12} \\ v_{22} \\ v_{32} \end{pmatrix}, \quad V_3 = \begin{pmatrix} v_{13} \\ v_{23} \\ v_{33} \end{pmatrix},$$

entonces sabemos que vale $L = PDP^{-1}$, siendo

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad \text{y} \quad P = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix}.$$

Lo que nos interesa es conocer el vector de distribución de edades X_k , para valores de k grandes. Este vector lo podemos calcular mediante $X_k = L^k X_0$. Además sabemos que $L = PDP^{-1}$ implica $L^k = PD^k P^{-1}$. Juntando las dos fórmulas anteriores deducimos

$$X_k = PD^k P^{-1} X_0, \quad \forall k = 1, 2, \dots$$

Consideremos un valor de k arbitrario fijo. Multiplicando la fórmula anterior por $\frac{1}{\lambda_1^k}$ obtenemos

$$\frac{1}{\lambda_1^k} X_k = \frac{1}{\lambda_1^k} PD^k P^{-1} X_0 = P \left(\frac{1}{\lambda_1^k} D^k \right) P^{-1} X_0.$$

Como es $D^k = \begin{pmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{pmatrix}$, obtenemos $\frac{1}{\lambda_1^k} X_k = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{\lambda_2}{\lambda_1}\right)^k & 0 \\ 0 & 0 & \left(\frac{\lambda_3}{\lambda_1}\right)^k \end{pmatrix} P^{-1} X_0$. Ahora usando que

valen $|\lambda_2| < \lambda_1$ y $|\lambda_3| < \lambda_1$, deducimos $\lim_{k \rightarrow +\infty} \left(\frac{\lambda_2}{\lambda_1}\right)^k = \lim_{k \rightarrow +\infty} \left(\frac{\lambda_3}{\lambda_1}\right)^k = 0$. Luego

$$\lim_{k \rightarrow +\infty} \frac{1}{\lambda_1^k} X_k = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1} X_0.$$

Si escribimos $P^{-1}X_0 = \begin{pmatrix} c \\ d \\ e \end{pmatrix}$, obtenemos

$$\lim_{k \rightarrow +\infty} \frac{1}{\lambda_1^k} X_k = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c \\ d \\ e \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix} = c \begin{pmatrix} v_{11} \\ v_{21} \\ v_{31} \end{pmatrix} = cV_1.$$

Luego hemos probado la fórmula siguiente

$$\lim_{k \rightarrow +\infty} \frac{1}{\lambda_1^k} X_k = cV_1. \quad (23)$$

Por lo tanto⁷ $\frac{1}{\lambda_1^k} X_k \simeq cV_1$, para valores de k grandes. Luego

$$X_k \simeq c\lambda_1^k V_1, \quad \text{para } k \text{ grande.}$$

Esto nos dice que a medida que pasa el tiempo, el vector de distribución de edades tiende a ser un múltiplo de V_1 . Usando la fórmula para V_1 dada por el teorema 5.2, la fórmula anterior queda

$$\begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} \simeq c\lambda_1^k \begin{pmatrix} 1 \\ \frac{b_1}{\lambda_1} \\ \frac{b_1 b_2}{\lambda_1^2} \end{pmatrix} \quad (24)$$

Notar que la constante c no la conocemos, pero de la fórmula (23) se deduce que es mayor o igual a cero; además siempre podemos suponer, y lo haremos, que c es positiva. Luego de (24) deducimos

$$\frac{y_k}{x_k} \simeq \frac{b_1}{\lambda_1}, \quad \frac{z_k}{x_k} \simeq \frac{b_1 b_2}{\lambda_1^2}.$$

Este resultado muestra que las proporciones entre las distintas clases de edad tienden a estabilizarse en constantes que dependen del valor propio dominante, pero no dependen de la distribución inicial de edades. Además, de $X_k \simeq c\lambda_1^k V_1$ y $X_{k-1} \simeq c\lambda_1^{k-1} V_1$, deducimos

$$X_k \simeq \lambda_1 X_{k-1}, \quad \text{para } k \text{ grande.} \quad (25)$$

Luego el vector de distribución de edades tiende a ser un múltiplo escalar del precedente, siendo este escalar el valor propio dominante de la matriz de Leslie.

Por otro lado, de $X_k \simeq c\lambda_1^k V_1$ obtenemos lo siguiente.

Teorema 5.5. *Sea λ_1 el valor propio dominante de L . Entonces*

- *si $\lambda_1 > 1$, entonces la población tenderá a crecer exponencialmente;*
- *si $\lambda_1 < 1$, entonces la población tenderá a decrecer exponencialmente (se extingue);*
- *si $\lambda_1 = 1$, entonces la población tenderá a estabilizarse.* □

El caso $\lambda = 1$ es particularmente interesante porque es $X_k \simeq cV_1$, para k grande. En este caso la población se aproxima a una distribución límite de edades, que es un múltiplo escalar del vector propio V_1 .

⁷El símbolo \simeq se lee “es aproximadamente igual a”.

El problema para aplicar lo anterior es que λ_1 es la raíz positiva del polinomio característico de L , que sabemos existe, pero no siempre es fácil de hallar. Este problema se resuelve recordando la tasa neta de reproducción $R = a_1 + a_2b_1 + a_3b_1b_2 + \dots + a_nb_1b_2\dots b_{n-1}$.

Proposición 5.6. *El número R y el valor propio positivo λ_1 están relacionados por lo siguiente*

$$R > 1 \Leftrightarrow \lambda_1 > 1, \quad R = 1 \Leftrightarrow \lambda_1 = 1, \quad R < 1 \Leftrightarrow \lambda_1 < 1.$$

Es decir, R y λ_1 en general son diferentes, pero tienen la misma relación con el número 1. □

Notar que del teorema y la proposición anteriores se deduce el teorema 5.1.

Ejemplo 5.7. Consideremos una población de la cual sabemos que su matriz de Leslie es

$$L = \begin{pmatrix} 0 & 4 & 3 \\ 1/2 & 0 & 0 \\ 0 & 1/4 & 0 \end{pmatrix}.$$

Supongamos que la duración de cada clase de edad es de 5 años. El valor de R es

$$R = 0 + 4(1/2) + 3(1/2)(1/4) = 19/8 > 1,$$

luego la población tiende a crecer exponencialmente. El polinomio característico de L es

$$\chi_L(\lambda) = -\lambda^3 + 2\lambda + 3/8.$$

Por tanteo se obtiene que $\lambda_1 = 3/2$ es el valor propio dominante. Notar que λ_1 y R son ambos mayores que 1 (en concordancia con la proposición anterior). Aplicando la fórmula (25) deducimos

$$X_k \simeq \frac{3}{2}X_{k-1}.$$

Por lo tanto en cada período de tiempo la población de hembras de cada clase, así como el total, tiende a incrementarse en un 50%. Calculando el vector propio V_1 correspondiente a $\lambda_1 = 3/2$ obtenemos

$$V_1 = \begin{pmatrix} 1 \\ 1/3 \\ 1/18 \end{pmatrix}.$$

Esto nos dice que las hembras de la población tenderán a distribuirse en las 3 clases de edad en proporciones 1, 1/3 y 1/18. Para traducir las proporciones anteriores en porcentajes, debemos resolver

$$100 = a + \frac{a}{3} + \frac{a}{18} \quad \Rightarrow \quad 100 = \left(1 + \frac{1}{3} + \frac{1}{18}\right)a = \frac{25}{18}a \quad \Rightarrow \quad a = 72.$$

Luego $a = 72$, $a/3 = 24$ y $a/18 = 4$. Entonces la población tiende a una distribución que consiste en

72% de hembras de 0 a 5 años; 24% de hembras de 5 a 10 años; 4% de hembras de 10 a 15 años.

En este caso se pueden obtener los otros valores propios, que son $\lambda_2 = \frac{-3+\sqrt{5}}{4} \simeq -0,19$ y $\lambda_3 = \frac{-3-\sqrt{5}}{4} \simeq -1,3$. Notar que vale $|\lambda_2| \simeq 0,19 < 1,5 = \lambda_1$ y $|\lambda_3| \simeq 1,3 < 1,5 = \lambda_1$ en concordancia con el teorema 5.4.

5.3. Un modelo estocástico

En este modelo la población se supone dividida en grupos que llamamos *estados* y lo que lo caracteriza es que la matriz que describe la transición de un estado al siguiente verifica que todos sus coeficientes son mayores o iguales que cero y las columnas suman 1. Cuando esto sucede, decimos que se trata de un modelo *estocástico*. Un ejemplo es el modelo de genética visto en la Sección 5.1. No vamos a detenernos en las características generales de este tipo de modelo y solo mostraremos cómo se aplica con el siguiente ejemplo.

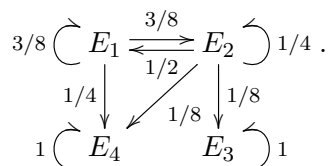
Consideremos una población que se distribuye, de acuerdo al padecimiento de una determinada enfermedad, en cuatro estados: el estado E_1 de salud, el estado E_2 de enfermedad, el estado E_3 de muerte a causa de dicha enfermedad y el estado E_4 de muerte por otra razón.

Luego de un relevamiento de datos, se obtiene la siguiente información de lo que sucede semana a semana:

- $3/8$ de las personas sanas permanecen sanas, $3/8$ se enferman y el resto muere (no por la enfermedad).
- la mitad de las personas enfermas se curan, mientras que $1/4$, sigue enferma, $1/8$ muere a causa de la enfermedad, y el resto muere por otra causa.

Nos concentramos en esta población, es decir que no le incorporamos los nuevos nacimientos. Lo que nos interesa es saber a largo plazo cuál es el porcentaje de muertes a causa de la enfermedad.

La información anterior puede resumirse en el siguiente diagrama, que llamaremos *grafo de transición* del modelo



Tomando como referencia una semana que llamamos semana 0, el vector

$$R_k = \begin{pmatrix} s_k \\ e_k \\ m_k \\ n_k \end{pmatrix}$$

reúne la información de cuál es la cantidad de individuos respectivamente sanos s_k , enfermos e_k , muertos por la enfermedad m_k y muertos por otra razón n_k , que hay en la k -ésima semana. Vamos a estudiar cómo evoluciona la población semana a semana, es decir, cómo es el pasaje de R_k a R_{k+1} . Esa información la obtenemos del grafo de transición.

La cantidad s_{k+1} de personas sanas en la semana $k + 1$, es la cantidad que se mantuvo sana que es $\frac{3}{8}s_k$, más la cantidad de personas que estaban enfermas y se curaron, que es $\frac{1}{2}e_k$. Luego

$$s_{k+1} = \frac{3}{8}s_k + \frac{1}{2}e_k.$$

La cantidad e_{k+1} de enfermos se obtiene sumando la cantidad de sanos que se enfermaron que es $\frac{3}{8}s_k$, más la cantidad de enfermos que siguieron enfermos, que es $\frac{1}{4}e_k$. Luego

$$e_{k+1} = \frac{3}{8}s_k + \frac{1}{4}e_k.$$

La cantidad de muertos por la enfermedad m_{k+1} es la suma de los que estaban enfermos y se murieron que es $\frac{3}{8}e_k$ más los que ya habían muerto por la enfermedad que es m_k . Luego

$$m_{k+1} = \frac{1}{8}e_k + m_k.$$

En forma análoga se obtiene

$$n_{k+1} = \frac{1}{4}s_k + \frac{1}{8}e_k + n_k.$$

Luego el vector R_{k+1} puede calcularse a partir de R_k , mediante $R_{k+1} = MR_k$, siendo

$$M = \begin{pmatrix} 3/8 & 1/2 & 0 & 0 \\ 3/8 & 1/4 & 0 & 0 \\ 0 & 1/8 & 1 & 0 \\ 1/4 & 1/8 & 0 & 1 \end{pmatrix}.$$

Como antes la condición $R_{k+1} = MR_k$ implica $R_k = M^k R_0$. Para calcular M^k veremos si podemos diagonalizar M , es decir, si existen una matriz diagonal D y una matriz invertible P tales que $M = PDP^{-1}$. El polinomio característico de M es

$$\chi_M(\lambda) = \begin{vmatrix} 3/8 - \lambda & 1/2 & 0 & 0 \\ 3/8 & 1/4 - \lambda & 0 & 0 \\ 0 & 1/8 & 1 - \lambda & 0 \\ 1/4 & 1/8 & 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 (\lambda^2 - 5/8\lambda - 3/32).$$

Este polinomio tiene raíces $\frac{3}{4}$, $\frac{-1}{8}$ y 1, de las cuales 1 es una raíz doble. Luego formamos la matriz diagonal

$$D = \begin{pmatrix} 3/4 & 0 & 0 & 0 \\ 0 & -1/8 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Los vectores propios correspondientes a cada uno de los valores propios son los siguientes:

$$\text{para } \lambda = \frac{3}{4}, \text{ son de la forma } \begin{pmatrix} -8/3z \\ -2z \\ z \\ 11/3z \end{pmatrix}; \quad \text{para } \lambda = -\frac{1}{8}, \text{ son de la forma } \begin{pmatrix} 9z \\ -9z \\ z \\ -z \end{pmatrix};$$

$$\text{para } \lambda = 1, \text{ son de la forma } \begin{pmatrix} 0 \\ 0 \\ z \\ t \end{pmatrix}.$$

Para formar la matriz P de vectores propios, vamos a aprovechar el hecho de que tenemos dos grados de libertad en los vectores propios correspondientes al valor propio (doble) 1 y tomar entonces dos vectores propios cualesquiera, teniendo cuidado de que no sean uno múltiplo del otro (para que la matriz P no tenga una columna múltiplo de la otra y pueda entonces ser invertible). Elegimos entonces la siguiente matriz

$$P = \begin{pmatrix} -8/3 & 9 & 0 & 0 \\ -2 & -9 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 11/3 & -1 & 0 & 1 \end{pmatrix}.$$

Esta matriz es invertible y su inversa es

$$P^{-1} = \begin{pmatrix} -3/14 & -3/14 & 0 & 0 \\ 1/21 & -4/63 & 0 & 0 \\ 1/6 & 5/18 & 1 & 0 \\ 5/6 & 13/18 & 0 & 1 \end{pmatrix}.$$

Al ser M invertible deducimos que M es diagonalizable y vale $M = PDP^{-1}$. Supongamos que en la semana 0, la población es de s personas sanas y e personas enfermas, luego el vector de distribución inicial es

$$R_0 = \begin{pmatrix} s \\ e \\ 0 \\ 0 \end{pmatrix}.$$

Entonces después de que pasan k semanas, la nueva distribución en los cuatro estados está dada por el siguiente vector

$$R_k = M^k R_0 = (PDP^{-1})^k R_0 = PD^k P^{-1} R_0 = PD^k P^{-1} \begin{pmatrix} s \\ e \\ 0 \\ 0 \end{pmatrix}.$$

Observar que si queremos saber lo que pasa a la larga, hacemos tender k a infinito, obteniendo

$$\begin{aligned} \lim_{k \rightarrow +\infty} R_k &= \lim_{k \rightarrow +\infty} PD^k P^{-1} \begin{pmatrix} s \\ e \\ 0 \\ 0 \end{pmatrix} = P \left(\lim_{k \rightarrow +\infty} D^k \right) P^{-1} \begin{pmatrix} s \\ e \\ 0 \\ 0 \end{pmatrix} = P \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} P^{-1} \begin{pmatrix} s \\ e \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1/6 & 5/8 & * & * \\ 5/6 & 13/18 & * & * \end{pmatrix} \begin{pmatrix} s \\ e \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ (1/6)s + (5/18)e \\ (5/6)s + (13/18)e \end{pmatrix}. \end{aligned}$$

Las estrellas (*) representan números que no afectan el resultado final, por lo cual no los calculamos. Notar que en lo anterior usamos

$$\lim_{k \rightarrow +\infty} D^k = \lim_{k \rightarrow +\infty} \begin{pmatrix} (3/4)^k & 0 & 0 & 0 \\ 0 & (-1/8)^k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \lim_{k \rightarrow +\infty} (3/4)^k & 0 & 0 & 0 \\ 0 & \lim_{k \rightarrow +\infty} (-1/8)^k & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

dado que los valores absolutos de $3/4$ y $-1/8$ son menores que uno.

Resumiendo hemos probado que vale

$$\lim_{k \rightarrow +\infty} R_k = \begin{pmatrix} 0 \\ 0 \\ (1/6)s + (5/18)e \\ (5/6)s + (13/18)e \end{pmatrix}.$$

Luego si al comienzo están todos sanos, se tiene $e = 0$ y la tendencia a largo plazo es que $1/6$ de los individuos (un 17% aproximadamente) morirán a causa de la enfermedad.

6. Apéndice 1. Diagonalización

En este apéndice mostramos de dónde sale lo que hicimos para que ver si una matriz es diagonalizable y exploramos los casos en que el polinomio característico tiene menos raíces que su grado.

Empezamos explicando la técnica de diagonalización. Por simplicidad trabajaremos con matrices 2×2 , pero el razonamiento vale en general. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz 2×2 . Supongamos que existen una matriz diagonal $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ y una matriz invertible $P = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$ tales que $A = PDP^{-1}$. Esta última relación equivale a $AP = PD$, es decir

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (26)$$

Desarrollando los productos e igualando coeficiente a coeficiente obtenemos que la igualdad anterior equivale a las cuatro igualdades siguientes

$$\begin{cases} au_1 + bu_2 = \lambda_1 u_1 \\ cu_1 + du_2 = \lambda_1 u_2 \end{cases}, \quad \begin{cases} av_1 + bv_2 = \lambda_2 v_1 \\ cv_1 + dv_2 = \lambda_2 v_2 \end{cases}. \quad (27)$$

Pasando todos los términos de las ecuaciones para un mismo lado obtenemos

$$\begin{cases} (a - \lambda_1)u_1 + bu_2 = 0 \\ cu_1 + (d - \lambda_1)u_2 = 0 \end{cases}, \quad \begin{cases} (a - \lambda_2)v_1 + bv_2 = 0 \\ cv_1 + (d - \lambda_2)v_2 = 0 \end{cases} \quad (28)$$

Como la matriz $P = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$ es invertible, entonces u_1, u_2 y v_1, v_2 , no son ambos nulos. Luego u_1, u_2 y v_1, v_2 son soluciones no triviales de los siguientes sistemas homogéneos

$$\begin{cases} (a - \lambda_1)x + by = 0 \\ cx + (d - \lambda_1)y = 0 \end{cases}, \quad \begin{cases} (a - \lambda_2)x + by = 0 \\ cx + (d - \lambda_2)y = 0 \end{cases} \quad (29)$$

Esto implica que los determinantes de los sistemas son nulos

$$\begin{vmatrix} a - \lambda_1 & b \\ c & d - \lambda_1 \end{vmatrix} = \begin{vmatrix} a - \lambda_2 & b \\ c & d - \lambda_2 \end{vmatrix} = 0.$$

Por lo tanto, si escribimos

$$\chi_A(\lambda) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = \lambda^2 - (a + d)\lambda + ac - bd,$$

entonces λ_1 y λ_2 tienen que ser raíces del *polinomio característico* $\chi_A(\lambda)$.

En resumen, dada una matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, si existen una matriz diagonal $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ y una matriz invertible $P = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$ tales que $A = PDP^{-1}$, entonces λ_1 y λ_2 tienen que ser raíces del polinomio característico $\chi_A(\lambda)$ y u_1, u_2 y v_1, v_2 tienen que ser soluciones de los sistemas (29). Recíprocamente, si el polinomio $\chi_A(\lambda)$ tiene raíces λ_1 y λ_2 , entonces esos valores hacen que los determinantes de los sistemas (29) sean nulos, por lo cual admiten soluciones no triviales u_1, u_2 y v_1, v_2 . Esas soluciones verifican entonces las ecuaciones (28), que equivalen a las ecuaciones (27), y estas equivalen a la ecuación matricial (26). Luego si escribimos $P = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$ y $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, entonces P y D verifican $AP = PD$. Finalmente, si la matriz P es invertible, entonces despejando en $AP = PD$, obtenemos $A = PDP^{-1}$. Notar que si el polinomio característico tiene dos raíces distintas, entonces la invertibilidad de la matriz P viene garantizada por el teorema 4.3 (en su versión 2×2). Luego los casos delicados son cuando el polinomio característico tiene raíces múltiples o tiene raíces complejas. Eso es lo que veremos a continuación. Lo mostraremos para matrices 3×3 .

En lo que sigue estudiaremos el caso en que el polinomio característico de una matriz $A \in M_3$ tiene menos de tres raíces. Los siguientes ejemplos muestran las distintas posibilidades.

Ejemplo 6.1. Sea

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}.$$

Su polinomio característico es

$$\chi_A(\lambda) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 0 & 3-\lambda & 1 \\ 0 & 1 & 3-\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 - 6\lambda + 8) = -(\lambda-2)^2(\lambda-4).$$

Luego los valores propios son 2 y 4, siendo 2 una raíz doble de $\chi_A(\lambda)$. Para hallar los vectores propios correspondientes al valor propio 4 tenemos que resolver

$$\begin{cases} -2x + y + z = 0 \\ -y + z = 0 \\ y - z = 0 \end{cases} \Rightarrow x = y = z.$$

Luego los vectores propios son de la forma (x, x, x) , con $x \neq 0$. Elegimos el vector $(1, 1, 1)$.

Para hallar los vectores propios correspondientes al valor propio 2 tenemos que hallar (x, y, z) tales que

$$\begin{cases} y + z = 0 \\ y + z = 0 \\ y + z = 0 \end{cases} \Rightarrow z = -y; \text{ la variable } x \text{ está libre.}$$

Luego las soluciones son de la forma $(x, y, -y)$, con x e y libres (el sistema es compatible indeterminado con dos grados de libertad). Como queremos que la matriz P sea invertible, nos tomamos dos vectores propios de forma tal que no sea uno múltiplo del otro, por ejemplo con $x = 1$ e $y = 0$ obtenemos el $(1, 0, 0)$ y con $x = 0$ e $y = 1$ obtenemos el $(0, 1, -1)$. Luego con esos valores y vectores propios construimos

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

Por la misma razón de antes sabemos que vale $AP = PD$. Observar que es $\det P = 2 \neq 0$, por lo cual P es invertible. Luego $AP = PD$ implica $A = PDP^{-1}$ y por lo tanto A es diagonalizable.

Ejemplo 6.2. Sea

$$A = \begin{pmatrix} 2 & 3 & 2 \\ 0 & 2 & 0 \\ 0 & 2 & 4 \end{pmatrix}.$$

Su polinomio característico es

$$\chi_A(\lambda) = \begin{vmatrix} 2-\lambda & 3 & 2 \\ 0 & 2-\lambda & 0 \\ 0 & 2 & 4-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} 2-\lambda & 0 \\ 2 & 4-\lambda \end{vmatrix} = -(\lambda-2)^2(\lambda-4).$$

Luego, como en el caso anterior, los valores propios son 2 y 4, siendo 2 una raíz doble de $\chi_A(\lambda)$. Para hallar los vectores propios correspondientes al valor propio 2 tenemos que resolver

$$\begin{cases} 3y + 2z = 0 \\ 0 = 0 \\ 2y + 2z = 0 \end{cases} \Rightarrow y = z = 0.$$

Esto implica que los vectores propios correspondientes al valor propio 2 son los vectores de la forma $(x, 0, 0)$, con $x \neq 0$ libre (el sistema es compatible indeterminado con un grado de libertad). El problema acá es que $\lambda = 2$ es una raíz doble de $\chi_A(\lambda)$, pero no hay forma de encontrar dos vectores propios correspondientes que no sean colineales, dado que si $x_1 \neq 0$ y $x_2 \neq 0$, entonces $(x_1, 0, 0) = \frac{x_1}{x_2}(x_2, 0, 0)$. Entonces no vamos a poder formar una matriz invertible P cuyas columnas sean vectores propios. Así que en este caso A no es diagonalizable.

Ejemplo 6.3. Sea

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Su polinomio característico es

$$\chi_A(\lambda) = \begin{vmatrix} -\lambda & -1 & 1 \\ 1 & -1-\lambda & -1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 + \lambda + 1).$$

Este polinomio tiene una raíz $\lambda = 2$. Si tratamos de hallar las raíces de $\lambda^2 + \lambda + 1$, nos encontramos con

$$\lambda^2 + \lambda + 1 = 0 \quad \Rightarrow \quad \lambda = \frac{-1 \pm \sqrt{-3}}{2}.$$

Como el polinomio $\lambda^2 + \lambda + 1$ no tiene raíces reales, deducimos que la única raíz de $\chi_A(\lambda)$ es $\lambda = 2$. Luego caemos en una situación similar a la del ejemplo anterior, en que no podemos armar una matriz invertible cuyas columnas sean vectores propios de A y por lo tanto la matriz A no es diagonalizable.

En resumen, mirando los ejemplos anteriores encontramos tres tipos de situaciones:

- El polinomio característico es $\chi_A(\lambda) = (2-\lambda)(\lambda^2 + \lambda + 1)$, en que $\lambda^2 + \lambda + 1$ no tiene raíces reales. En este caso A no es diagonalizable.
- El polinomio característico es $\chi_A(\lambda) = -(\lambda-2)^2(\lambda-4)$. Considerando la raíz doble $\lambda = 2$, tenemos dos posibilidades respecto al sistema de ecuaciones que nos da los vectores propios correspondientes (recordar que este sistema siempre es compatible indeterminado)
 - El sistema tiene dos grados de libertad, en ese caso A es diagonalizable.
 - El sistema tiene un grado de libertad, en ese caso A no es diagonalizable.

Para matrices 4×4 sucede algo similar. Por ejemplo, si $\chi_A(\lambda) = (\lambda-2)(\lambda-3)(\lambda^2 + \lambda + 1)$, entonces A no es diagonalizable (porque $\lambda^2 + \lambda + 1$ no tiene raíces reales). En cambio si $\chi_A(\lambda) = (\lambda-2)(\lambda-3)(\lambda-1)^2$, entonces hay que estudiar los grados de libertad del sistema asociado al valor propio $\lambda = 1$: si tiene dos grados de libertad, entonces A es diagonalizable y si tiene un grado de libertad entonces A no es diagonalizable.

7. Apéndice 2. Modelo de Leslie

Este apéndice es para quienes quieran entender un poco más la teoría que sustenta el modelo de Leslie. Para simplificar la notación asumiremos que la población está dividida en solo tres clases de edad. La matriz de Leslie es

$$L = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & 0 & 0 \\ 0 & b_2 & 0 \end{pmatrix}.$$

Como siempre es $a_i \geq 0$ y $0 < b_i \leq 1$, para todo i , y para algún i es $a_i > 0$.

Teorema 7.1. *El polinomio característico de L es*

$$\chi_L(\lambda) = (-1)^3 (\lambda^3 - a_1\lambda^2 - a_2b_1\lambda - a_3b_1b_2).$$

Dem. Desarrollando el determinante por la tercer columna obtenemos

$$\begin{aligned} \chi_L(\lambda) &= \begin{vmatrix} a_1 - \lambda & a_2 & a_3 \\ b_1 & -\lambda & 0 \\ 0 & b_2 & -\lambda \end{vmatrix} = a_3 \begin{vmatrix} b_1 & -\lambda \\ 0 & b_2 \end{vmatrix} + (-\lambda) \begin{vmatrix} a_1 - \lambda & a_2 \\ b_1 & -\lambda \end{vmatrix} = a_3b_1b_2 - \lambda(-a_1\lambda + \lambda^2 - a_2b_1) \\ &= -\lambda^3 + a_1\lambda^2 + a_2b_1\lambda + a_3b_1b_2. \end{aligned}$$

Esta es la prueba para $n = 3$. La prueba general para n arbitrario se hace por inducción completa. \square

Teorema 7.2. 1. *La matriz L tiene un único valor propio positivo λ_1 .*

2. *Si $\lambda \neq 0$ es un valor propio de L , entonces $V = \begin{pmatrix} 1 \\ \frac{b_1}{\lambda} \\ \frac{b_1b_2}{\lambda^2} \end{pmatrix}$ es un vector propio de L correspondiente a λ .*

Dem. El polinomio característico de L es

$$\chi_L(\lambda) = -\lambda^3 + a_1\lambda^2 + a_2b_1\lambda + a_3b_1b_2.$$

Estamos interesados en las raíces positivas de $\chi_L(\lambda)$. Notar que $\lambda_1 \neq 0$ es raíz de $\chi_L(\lambda)$ si y solo si verifica

$$-\lambda_1^3 + a_1\lambda_1^2 + a_2b_1\lambda_1 + a_3b_1b_2 = 0 \Leftrightarrow a_1\lambda_1^2 + a_2b_1\lambda_1 + a_3b_1b_2 = \lambda_1^3 \Leftrightarrow \frac{a_1}{\lambda_1} + \frac{a_2b_1}{\lambda_1^2} + \frac{a_3b_1b_2}{\lambda_1^3} = 1. \quad (30)$$

Luego si \mathbb{R}^+ son los reales positivos y definimos una función $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ mediante

$$f(\lambda) = \frac{a_1}{\lambda} + \frac{a_2b_1}{\lambda^2} + \frac{a_3b_1b_2}{\lambda^3}, \quad \forall \lambda > 0,$$

entonces $\lambda_1 > 0$ es raíz de $\chi_L(\lambda)$ si y solo si verifica $f(\lambda_1) = 1$.

La derivada de f es

$$f'(\lambda) = -\frac{a_1}{\lambda^2} - \frac{2a_2b_1}{\lambda^3} - \frac{2a_3b_1b_2}{\lambda^4}, \quad \forall \lambda > 0.$$

Como la derivada es negativa, deducimos que f es estrictamente decreciente. Además observar que valen $\lim_{\lambda \rightarrow 0^+} f(\lambda) = +\infty$ y $\lim_{\lambda \rightarrow +\infty} f(\lambda) = 0$. Esto implica que existe un único $\lambda_1 > 0$ tal que $f(\lambda_1) = 1$, que es lo que queríamos probar.

Para probar la segunda parte, directamente calculamos el producto de la matriz L por el vector V

$$LV = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & 0 & 0 \\ 0 & b_2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{b_1}{\lambda} \\ \frac{b_1b_2}{\lambda^2} \end{pmatrix} = \begin{pmatrix} a_1 + a_2\frac{b_1}{\lambda} + a_3\frac{b_1b_2}{\lambda^2} \\ b_1 \\ \frac{b_1b_2}{\lambda} \end{pmatrix} = \lambda_1 \begin{pmatrix} \frac{a_1}{\lambda} + \frac{a_2b_1}{\lambda^2} + \frac{a_3b_1b_2}{\lambda^3} \\ \frac{b_1}{\lambda} \\ \frac{b_1b_2}{\lambda^2} \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \frac{b_1}{\lambda} \\ \frac{b_1b_2}{\lambda^2} \end{pmatrix} = \lambda V.$$

Notar que en la penúltima igualdad usamos que vale $\frac{a_1}{\lambda} + \frac{a_2b_1}{\lambda^2} + \frac{a_3b_1b_2}{\lambda^3} = 1$, lo cual se deduce de (30). \square

Proposición 7.3. *El número R y el valor propio positivo λ_1 están relacionados por lo siguiente*

$$R > 1 \Leftrightarrow \lambda_1 > 1, \quad R = 1 \Leftrightarrow \lambda_1 = 1, \quad R < 1 \Leftrightarrow \lambda_1 < 1.$$

Es decir, R y λ_1 en general son diferentes, pero tienen la misma relación con el número 1.

Dem. El polinomio característico de L es

$$\chi_L(\lambda) = -\lambda^3 + a_1\lambda^2 + a_2b_1\lambda + a_3b_1b_2$$

y sabemos que $\lambda_1 > 0$ es raíz de $\chi_L(\lambda)$ si y solo si verifica $f(\lambda_1) = 1$, siendo

$$f(\lambda) = \frac{a_1}{\lambda} + \frac{a_2b_1}{\lambda^2} + \frac{a_3b_1b_2}{\lambda^3}.$$

Observar que vale

$$f(1) = a_1 + a_2b_1 + a_3b_1b_2 = R.$$

Nosotros ya vimos que la función $f(\lambda)$ es estrictamente decreciente. Luego

$$\begin{aligned}\lambda_1 > 1 &\Leftrightarrow f(\lambda_1) < f(1) &\Leftrightarrow 1 < R, \\ \lambda_1 = 1 &\Leftrightarrow f(\lambda_1) = f(1) &\Leftrightarrow 1 = R, \\ \lambda_1 < 1 &\Leftrightarrow f(\lambda_1) > f(1) &\Leftrightarrow 1 > R. \quad \square\end{aligned}$$