

Fundamentos matematicos de la prediccion de valores de cria^{*}

Francisco Javier Mendizábal Múgica

1994

Resumen

Notas de la asignatura Mejora Genética Animal. Escuela Técnica Superior de Ingenieros Agrónomos, UPNA. Pamplona-Iruña, 1994.¹

^{*}Creative Commons CC-BY License, derecho de copiar, distribuir, exhibir y representar la obra y hacer obras derivadas siempre y cuando reconozca y cite esta obra.

¹Reescritas y corregidas por Carolina García Baccino, Fernando Macedo y Andrés Legarra, 2020.

Índice

1. Introduccion	4
2. Algebra	4
2.1. Algebra matricial de los números reales y álgebra de matrices . . .	4
2.2. Definición y dimensión de una matriz	4
2.2.1. Ejemplos de matrices:	5
2.3. Cálculo matricial	6
2.3.1. Definición de las operaciones de cálculo	6
2.3.2. Reglas del álgebra matricial	11
2.4. Matrices especiales	17
2.4.1. MATRIZ CUADRADA	17
2.4.2. MATRIZ SIMETRICA	17
2.4.3. MATRIZ TRIANGULAR	17
2.4.4. MATRIZ DIAGONAL	18
2.4.5. MATRIZ ESCALAR	18
2.4.6. MATRIZ IDENTIDAD O UNIDAD	18
2.4.7. INVERSA DE UNA MATRIZ CUADRADA	19
2.4.8. MATRIZ NULA	19
2.4.9. MATRICES PARTICIONADAS	19
3. Matrices cuadradas	20
3.1. Determinante de una matriz cuadrada	20
3.2. Deteminante de una matriz cuadrada de orden 2:	21
3.3. Determinante de una matriz cuadrada de orden 3	21
3.4. Desarrollo del determinante de una matriz segun una fila o una columna	22
3.4.1. Subdeterminante:	22
3.4.2. Adjunto (cofactor)	23
3.4.3. Desarrollo del determinante de \mathbf{A} según la i -ésima fila: . .	23
3.5. Dificultad del cálculo del determinante de una matriz de orden n	25
3.6. Propiedades de los determinantes	26
3.7. Matriz adjunta e inversa de una matriz	30
3.7.1. Definición de matriz adjunta de \mathbf{A} : $\text{adj}(\mathbf{A})$	30
3.7.2. Multiplicación de una matriz cuadrada \mathbf{A} por su matriz adjunta $\text{adj}(\mathbf{A})$	31
3.7.3. Inversa de una matriz cuadrada \mathbf{A}	32
3.7.4. Inversa de una matriz diagonal	33
3.7.5. Inversa del producto de dos matrices cuadradas	33
4. Sistemas de ecuaciones	33
4.1. Ecuaciones en el álgebra de las matrices	33
4.1.1. Definición de un sistema de ecuaciones y su solución . . .	33
4.1.2. Un sistemas de ecuaciones en notación matricial	34
4.1.3. Solución de ecuaciones en el álgebra de las matrices. . . .	35

4.2.	Rango de la matriz de coeficientes de un sistema ecuaciones	36
4.2.1.	Sistemas de ecuaciones con un numero infinito de soluciones	36
4.2.2.	Rango de una matriz	38
4.3.	Operadores elementales y sistemas de ecuaciones	39
4.3.1.	Operaciones elementales en filas y columnas	39
4.3.2.	Forma canónica de una matriz cuadrada. Cálculo del rango de una matriz.	43
4.3.3.	Eigenvalues and eigenvectors	46
4.3.4.	Operaciones elementales en sistemas de ecuaciones	46
4.4.	Solución de un sistema de ecuaciones con rango incompleto	48
4.4.1.	Inversa generalizada de una matriz	48
4.4.2.	Cálculo de g-inversas de una matriz	49
4.4.3.	Solución de sistemas de ecuaciones con rango incompleto por medio de g-inversas. Funciones estimables.	53
4.5.	Absorción de ecuaciones	55
5.	Modelos lineales en Producción Animal	57
5.1.	Definición de un modelo lineal	57
5.2.	Método de los mínimos cuadrados (LSQ)	60
5.2.1.	Ecuaciones normales con rango completo	60
5.2.2.	Ecuaciones normales con rango incompleto. Restricciones en el vector β	65
5.3.	Comparación entre dos modelos. Sesgo.	66
5.4.	Vector de esperanzas matemáticas y matriz de varianzas-covarianzas de un vector de variables aleatorias.	70
5.4.1.	Definición	70
5.4.2.	Matriz de varianzas-covarinzas de un vector con esperanza matemática nula.	71
5.5.	Método de los mínimos cuadrados generalizado (GLS)	72
5.6.	Efectos fijos y al azar. Un modelo mixto	77
5.7.	BLUP (Best Linear Unbiased Prediction)	79
6.	Ejemplo de utilización práctica de un modelo lineal	80
6.1.	Datos	80
6.2.	Factores de influencia	80
6.2.1.	Influencia genética sobre el rendimiento de una vaca	80
6.2.2.	Factores de influencia ambientales	81
6.2.3.	Modelo	82
6.2.4.	Asunciones del modelo	83
6.3.	Apéndice: script Julia para el ejemplo presentado	91

1. Introduccion

Tanto la estimación de parametros poblacionales, como los métodos modernos de predicción del valor de cria de los individuos de una población, tienen como fundamento de sus cálculos el método de los mínimos cuadrados. Las demostraciones teóricas y la aplicación de estos métodos son sencillas en poblaciones pequeñas y con una estructura equilibrada, pero se vuelven muy complicados en caso contrario. Sin embargo se simplifican notablemente, si se abandona el álgebra de números reales para trabajar en el álgebra de matrices. Por ello es especialmente interesante para las personas relacionadas con la predicción y utilización de valores de cría, dominar los fundamentos del álgebra de matrices.

Con estos apuntes se pretende mostrar estos fundamentos del modo mas sencillo posible y dar unos primeros pasos en su aplicacion practica.

2. Algebra

2.1. Algebra matricial de los números reales y álgebra de matrices

Partimos de la base de que conocemos el algebra de los números reales. En ella trabajamos con ecuaciones, cuyos elementos son “incógnitas”. Las incógnitas se simbolizan con letras. Números y letras se unen por medio de operaciones matemáticas. Estas operaciones tienen determinadas propiedades. Siempre que se tenga en cuenta estas propiedades, se puede introducir cambios en las ecuaciones. El objetivo final de estos cambios suele ser encontrar valores reales para las incógnitas.

Entre el álgebra de los números reales y el álgebra de matrices se pueden destacar las siguientes diferencias:

1. Los elementos de las ecuaciones del álgebra de matrices no son números reales ni simples incógnitas, sino matrices de números reales y matrices de incógnitas.
2. Las operaciones con matrices se definen de forma diferente a las operaciones con números reales y poseen propiedades distintas.

Nuestro primer paso sera la definición de matriz y de sus operaciones, así como el repaso de las reglas del álgebra de matrices.

2.2. Definición y dimensión de una matriz

Una matriz es una ordenación regular de elementos en filas y columnas. Los elementos de una matriz pueden ser números reales o incógnitas (letras minúsculas). Todas las filas tienen la misma longitud, es decir el mismo número de elementos y lo mismo hay que decir de las columnas.

Generalmente una matriz se representa por una letra mayúscula (en negrita, sin cursiva) y sus elementos individuales se escriben como letras minúsculas con dos subíndices. Por ejemplo:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = a_{ij} \begin{cases} i = 1, 2 \\ j = 1, 3 \end{cases}$$

El primer subíndice de un elemento informa sobre la fila en la que este se encuentra, mientras que el segundo indica la columna correspondiente. A veces se separan por una coma:

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}$$

La DIMENSION de una matriz se compone de dos números enteros, de los cuales el primero es igual al número de filas y el segundo al número de columnas de la matriz. La matriz \mathbf{A} de nuestro ejemplo tiene una dimensión de 2 x 3, porque tiene dos filas y tres columnas.

Para llevar a cabo operaciones con matrices es imprescindible tener en cuenta su dimensión.

Un vector es una matriz con la dimension $m \times 1$ (vector columna) o $1 \times n$ (vector fila). Los vectores se representan por medio de letras minúsculas (negrita, no cursiva: por ejemplo \mathbf{x} o \mathbf{x}) y los elementos de un vector, por la misma letra minúscula (cursiva) acompañada de un único subíndice (por ejemplo x_1). Un escalar es un número real, o lo que es lo mismo, “una matriz de dimensión 1×1 , cuyo único elemento es un número real”. Se representan por una letra minúscula (cursiva), por ejemplo g o g .

2.2.1. Ejemplos de matrices:

$$\mathbf{B} = \begin{pmatrix} 4 & 7 & 5 & 6 & 7 \\ 9 & 3 & 2 & 1 & 2 \\ 2 & 0 & 4 & 2 & 3 \\ 6 & 6 & 5 & 3 & 2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \end{pmatrix}$$

\mathbf{B} es una matriz de dimensión 4 x 5.

$$\mathbf{c} = \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

c es un vector columna de dimensión 3, o una matriz de dimensión 3 x 1.

$$d = (5 \quad 4 \quad 3 \quad 0)$$

d es un vector fila de dimensión 4, o una matriz de dimensión 1 x 4

2.3. Cálculo matricial

2.3.1. Definición de las operaciones de cálculo

Para definir una operación es necesario determinar lo siguiente:

- dimensión de las matrices con las que se pretende operar
- dimensión de la matriz resultante
- modo de operar para calcular cada uno de los elementos de la matriz resultante

Las operaciones que se van a explicar a continuación son: transposición, adición, sustracción, producto de dos matrices, producto de un escalar por una matriz y producto directo de dos matrices.

Observación: el producto de una escalar por una matriz y el producto directo son dos operaciones en si, distintas del producto de dos matrices.

2.3.1.1. Transposicion Matriz de partida: una matriz \mathbf{A} , de dimensión $n \times m$. Matriz resultante: una matriz $\mathbf{B} = \mathbf{A}'$, de dimension $m \times n$. El indicador de transposicion es el apóstrofe ', por ejemplo en \mathbf{A}' , y a veces un superíndice T como en \mathbf{A}^T .

El elemento situado en la posición (i, j) de $\mathbf{B} = \mathbf{A}'$ es igual al elemento (j, i) de \mathbf{A} : $b_{i,j} = a_{j,i}$ dicho de otra forma, la matriz transpuesta de \mathbf{A} es la que resulta de escribir las filas de \mathbf{A} como columnas.

Por ejemplo, matriz $\mathbf{A} = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 \end{pmatrix}$ de dimensión 2×4 y matriz $\mathbf{B} = \mathbf{A}' = \begin{pmatrix} 2 & 6 \\ 3 & 7 \\ 4 & 8 \\ 5 & 9 \end{pmatrix}$ de dimension 4×2 .

El elemento $a_{1,3}$ de \mathbf{A} es igual al elemento $b_{3,1}$ de \mathbf{B} :

$$a_{1,3} = b_{3,1} = 4$$

Los elementos de la primera fila de \mathbf{A} (2 3 4 5), son los elementos de la primera columna de \mathbf{B} .

2.3.1.2. Adición

- Matrices de partida: dos matrices **A** y **B**, ambas con la misma dimensión $m \times n$. La suma de dos matrices de dimensiones distintas es imposible.
- Matriz resultante: una matriz **C**, con la misma dimensión que **A** y **B**, es decir $m \times n$.
- El elemento situado en la posición (i, j) de **C** es igual a la suma de los elementos correspondientes de **A** y **B**: $c_{i,j} = a_{i,j} + b_{i,j}$.

Ejemplo:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 2 \\ 1 & 0 & 1 \\ 3 & 2 & 9 \end{pmatrix}$$

dimensión 4×3

$$\mathbf{B} = \begin{pmatrix} 7 & 6 & 9 \\ 3 & 3 & 8 \\ 6 & 7 & 4 \\ 9 & 0 & 2 \end{pmatrix}$$

A y **B** tienen la misma dimensión, luego se puede llevar a cabo la adición. La matriz resultado **C** tendrá también la dimensión 4×3

$$\mathbf{C} = \begin{pmatrix} 1+7 & 2+6 & 3+9 \\ 6+3 & 5+3 & 2+8 \\ 1+6 & 0+7 & 1+4 \\ 3+9 & 2+0 & 9+2 \end{pmatrix} = \begin{pmatrix} 8 & 8 & 12 \\ 9 & 8 & 10 \\ 7 & 7 & 5 \\ 12 & 2 & 11 \end{pmatrix}$$

El elemento $c_{2,3}$ de **C** es $c_{2,3} = a_{2,3} + b_{2,3} = 2 + 8 = 10$

2.3.1.3. Substracción La substracción de dos matrices es igual a la adición, pero cambiando de signo todos los elementos del segundo sumando. La substracción de las matrices del ejemplo anterior sería:

$$\mathbf{D} = \mathbf{A} - \mathbf{B} = \begin{pmatrix} 1-7 & 2-6 & 3-9 \\ 6-3 & 5-3 & 2-8 \\ 1-6 & 0-7 & 1-4 \\ 3-9 & 2-0 & 9-2 \end{pmatrix} = \begin{pmatrix} -6 & -4 & -6 \\ 3 & 2 & -6 \\ -5 & -7 & -3 \\ -6 & 2 & 7 \end{pmatrix}$$

2.3.1.4. Multiplicación

- Matrices de partida:
 - una matriz **A** de dimensión $m \times n$

- una matriz **B** de dimensión $n \times p$
- Matriz resultante: una matriz de dimensión $m \times p$

La multiplicación se indica como $\mathbf{A} \times \mathbf{B}$ o bien directamente como \mathbf{AB} . Obsérvese que, para que la multiplicación sea posible, el número de columnas de la primera matriz (n) tiene que ser igual al número de filas de la segunda matriz.

dimensión de la primera matriz	dimension de la segunda matriz	dimension de la matriz resultado
$m \times n$	$n \times p$	$m \times p$
+-----+		

Solo se puede llevar a cabo la multiplicación si estos dos números son iguales

- Para calcular el elemento $c_{i,j}$ de **C** se toma la i -ésima fila de **A** :

$$(a_{i,1} a_{i,2} a_{i,n}) \text{ y la } j\text{-ésima columna de } \mathbf{B} : \begin{pmatrix} b_{1,j} \\ b_{2,j} \\ b_{n,j} \end{pmatrix}$$

El elemento $c_{i,j}$ se calcula como:

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j} = a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \dots + a_{i,n} b_{n,j}$$

Esto se puede ver mas facilmente con un ejemplo:

Matriz **A** = $\begin{pmatrix} 4 & 5 & 0 & 1 \\ 1 & 6 & 1 & 2 \end{pmatrix}$ de dimension 2×4 .

Matriz **B** = $\begin{pmatrix} 2 & 6 & 3 \\ 3 & 5 & 0 \\ 2 & 4 & 2 \\ 1 & 2 & 7 \end{pmatrix}$ de dimension 4×3

dimensión de la primera matriz	dimension de la segunda matriz	dimension de la matriz resultado
2×4	4×3	2×3
+-----+		

4=4 luego se puede llevar a cabo la multiplicación

Para calcular el elemento $c_{1,3}$ de **C**, tomamos la primera fila de **A** y la tercera columna de **B**, multiplicamos los elementos de las mismas dos a dos y sumamos los productos:

Primera fila de **A**

$$(4 \quad 5 \quad 0 \quad 1)$$

Tercera columna de \mathbf{B}

$$\begin{pmatrix} 3 \\ 0 \\ 2 \\ 7 \end{pmatrix}$$

$$c_{1,3} = 4 \times 3 + 5 \times 0 + 0 \times 2 + 1 \times 7 = 18$$

Para calcular ahora el elemento $c_{2,1}$ de \mathbf{C} , tomamos la segunda fila de \mathbf{A} y la primera columna de \mathbf{B} , y repetimos la operación anterior.

Segunda fila de \mathbf{A}

$$(1 \quad 6 \quad 1 \quad 2)$$

Primera columna de \mathbf{B}

$$\begin{pmatrix} 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

$$c_{2,1} = 1 \times 2 + 6 \times 3 + 1 \times 2 + 2 \times 1 = 24$$

Del mismo modo se calculan los 2 x 3 elementos de \mathbf{C} .

2.3.1.5. Multiplicación de un escalar por una matriz La multiplicación de un escalar por una matriz se define como una operación especial, ya que según la definición de multiplicación de dos matrices, no sería posible multiplicar una matriz de dimension 4 x 1 por una matriz de dimensión $n \times m$, mas que en el caso de que $n=1$. Esta operación es un caso especial del producto directo, el cual se explicará mas adelante.

- Matrices de partida: un escalar k y una matriz \mathbf{A} , ésta última con dimension $n \times m$
- Matriz resultante: una matriz \mathbf{B} , de dimensión $n \times m$ (la misma dimensión de \mathbf{A})
- Para calcular el elemento $b_{i,j}$ de \mathbf{B} , se multiplica el elemento correspondiente de \mathbf{A} por el escalar k : $b_{i,j} = k \times a_{i,j}$

Ejemplo: escalar $k = 5$; matriz

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

dimension 2 x 3

$$\mathbf{B} = k\mathbf{A} = \begin{pmatrix} 5 \times 3 & 5 \times 2 & 5 \times 0 \\ 5 \times 4 & 5 \times 1 & 5 \times 2 \end{pmatrix} = \begin{pmatrix} 15 & 10 & 0 \\ 20 & 5 & 10 \end{pmatrix}$$

dimension 2 x 3

2.3.1.6. Producto de Kronecker A veces² se le llama también *producto directo* pero vamos a evitar ese nombre, ya que se utiliza también con el llamado *producto de Hadamard* (o elemento-por-elemento). Vamos a indicar el producto de Kronecker por el simbolo \otimes ([como en wikipedia](#)).

No hay que confundir esta operacion con la multiplicacion.

- Matrices de partida:
 - una matriz \mathbf{A} de dimensión $m \times n$
 - una matriz \mathbf{B} de dimensión $p \times q$
- Matriz resultante del producto de Kronecker: una matriz \mathbf{C} de dimensión $(m \times p) \times (n \times q)$
- Modo de cálculo:

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \dots & a_{1,n}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \dots & a_{2,n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}\mathbf{B} & a_{m,2}\mathbf{B} & \dots & a_{m,n}\mathbf{B} \end{pmatrix}$$

donde $a_{i,j}\mathbf{B}$ representa el producto del escalar $a_{i,j}$ por la matriz \mathbf{B} .

La mejor forma de comprender cómo se lleva a cabo esta operación es utilizando un ejemplo:

matriz $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$ de dimension 2 x 2

matriz $\mathbf{B} = \begin{pmatrix} 1 & 4 \\ 6 & 8 \\ 4 & 1 \end{pmatrix}$ de dimension 3 x 2.

²por ejemplo en la version original de estas notas

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \left(\begin{array}{cc|cc} 2\mathbf{B} & 3\mathbf{B} \\ 5\mathbf{B} & 7\mathbf{B} \end{array} \right) = \left(\begin{array}{cc|cc} 2 \times 1 & 2 \times 4 & 3 \times 1 & 3 \times 4 \\ 2 \times 6 & 2 \times 8 & 3 \times 6 & 3 \times 8 \\ 2 \times 4 & 2 \times 1 & 3 \times 4 & 3 \times 1 \\ \hline 5 \times 1 & 5 \times 4 & 7 \times 1 & 7 \times 4 \\ 5 \times 6 & 5 \times 8 & 7 \times 6 & 7 \times 8 \\ 5 \times 4 & 5 \times 1 & 7 \times 4 & 7 \times 1 \end{array} \right)$$

$$\mathbf{C} = \begin{pmatrix} 2 & 8 & 3 & 12 \\ 12 & 16 & 18 & 24 \\ 8 & 2 & 12 & 13 \\ 5 & 20 & 7 & 28 \\ 30 & 40 & 42 & 56 \\ 20 & 5 & 28 & 7 \end{pmatrix}$$

La dimension de \mathbf{C} es 6 x 4, donde:

6 = 2 x 3 número de filas de \mathbf{A} por número de filas de \mathbf{B}

4 = 2 x 2 número de columnas de \mathbf{A} por número de columnas de \mathbf{B}

2.3.2. Reglas del álgebra matricial

Nuestro objetivo no es escribir un libro de matemáticas, sino introducir al lector de la manera mas sencilla posible en el álgebra de matrices. Por ello no se harán aquí demostraciones, sino simples comprobaciones de las reglas, por medio de matrices sencillas. Estas comprobaciones pueden servir también de ejercicios fáciles. Para quien quiera profundizar en este tema, existen libros de consulta suficientes y [wikipedia](https://es.wikipedia.org) es un recurso excelente.

2.3.2.1. TRANSPOSICION

1. La transpuesta de la transpuesta de una matriz \mathbf{A} , es la matriz de partida $\mathbf{A} : (\mathbf{A}')' = \mathbf{A}$

2.3.2.2. ADICION

2. La adición de matrices es asociativa: si queremos sumar tres matrices, obtendremos el mismo resultado sumando primero las dos primeras y sumando luego a la matriz resultante la tercera, que sumando primero la segunda y la tercera y sumando luego la primera a la matriz resultante:

$$\mathbf{A} + \mathbf{B} + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$$

Ejemplo:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 5 & 0 & 4 \\ 3 & 1 & 7 \\ 6 & 2 & 1 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 0 & 4 & 3 \\ 6 & 5 & 4 \\ 1 & 2 & 3 \end{pmatrix}$$

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 4 \\ 3 & 1 & 7 \\ 6 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 4 & 3 \\ 6 & 5 & 4 \\ 1 & 2 & 3 \end{pmatrix} = \\ \begin{pmatrix} 6 & 2 & 7 \\ 7 & 6 & 13 \\ 13 & 10 & 10 \end{pmatrix} + \begin{pmatrix} 0 & 4 & 3 \\ 6 & 5 & 4 \\ 1 & 2 & 3 \end{pmatrix} = \\ \begin{pmatrix} 6 & 6 & 10 \\ 13 & 11 & 17 \\ 14 & 12 & 13 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 4 \\ 3 & 1 & 7 \\ 6 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 4 & 3 \\ 6 & 5 & 4 \\ 1 & 2 & 3 \end{pmatrix} = \\ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 5 & 4 & 7 \\ 9 & 6 & 11 \\ 7 & 4 & 4 \end{pmatrix} = \\ \begin{pmatrix} 6 & 6 & 10 \\ 13 & 11 & 17 \\ 14 & 12 & 13 \end{pmatrix} \end{aligned}$$

3. La adición de matrices es conmutativa:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 5 & 0 & 4 \\ 3 & 1 & 7 \\ 6 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+5 & 2+0 & 3+4 \\ 4+3 & 4+1 & 6+7 \\ 7+6 & 8+2 & 9+1 \end{pmatrix} = \begin{pmatrix} 6 & 2 & 7 \\ 7 & 6 & 13 \\ 13 & 10 & 10 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 0 & 4 \\ 3 & 1 & 7 \\ 6 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 5+1 & 0+2 & 4+3 \\ 3+4 & 1+4 & 7+6 \\ 6+7 & 2+8 & 1+9 \end{pmatrix} = \begin{pmatrix} 6 & 2 & 7 \\ 7 & 6 & 13 \\ 13 & 10 & 10 \end{pmatrix}$$

2.3.2.3. ADICION Y TRANSPOSICION

4. La traspuesta de la suma de dos matrices es igual a la suma de las traspuestas de las mismas: $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$

$$(\mathbf{A} + \mathbf{B})' = \begin{pmatrix} 6 & 7 & 13 \\ 2 & 6 & 10 \\ 7 & 13 & 10 \end{pmatrix}$$

$$(\mathbf{A}' + \mathbf{B})' = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} + \begin{pmatrix} 5 & 3 & 6 \\ 0 & 1 & 2 \\ 4 & 7 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 7 & 13 \\ 2 & 6 & 10 \\ 7 & 13 & 10 \end{pmatrix}$$

2.3.2.4. MULTIPLICACION

5. La multiplicación es asociativa: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}; \mathbf{B} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 4 & 1 \end{pmatrix}; \mathbf{C} = \begin{pmatrix} 5 & 6 \\ 3 & 4 \end{pmatrix}$$

Dimensiones: \mathbf{A} : 2×3 \mathbf{B} : 3×2 \mathbf{C} : 2×2 .

Comprobación de que las multiplicaciones son posibles:

$$\mathbf{AB} : (2 \times \underline{3}) \Leftrightarrow (\underline{3} \times 2) \rightarrow 2 \times 2$$

$$(\mathbf{AB})\mathbf{C} : (2 \times \underline{2}) \Leftrightarrow (\underline{2} \times 2) \rightarrow 2 \times 2$$

$$\mathbf{BC} : (3 \times \underline{2}) \Leftrightarrow (\underline{2} \times 2) \rightarrow 2 \times 2$$

$$\mathbf{A}(\mathbf{BC}) : (2 \times \underline{3}) \Leftrightarrow (\underline{3} \times 2) \rightarrow 2 \times 2$$

De las dos formas son posibles las multiplicaciones y el resultado final es una matriz 2×2 .

Cálculo:

- Primer procedimiento:

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 20 & 10 \\ 47 & 28 \end{pmatrix}$$

$$(\mathbf{AB})\mathbf{C} = \begin{pmatrix} 20 & 10 \\ 47 & 28 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 130 & 160 \\ 319 & 364 \end{pmatrix}$$

- Segundo procedimiento:

$$\mathbf{BC} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 19 & 24 \\ 21 & 26 \\ 23 & 28 \end{pmatrix}$$

$$\mathbf{A}(\mathbf{BC}) = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 19 & 24 \\ 21 & 26 \\ 23 & 28 \end{pmatrix} = \begin{pmatrix} 130 & 160 \\ 319 & 364 \end{pmatrix}$$

6. Una característica muy importante de la multiplicación de dos matrices, es la siguiente:

La multiplicación de dos matrices no es conmutativa

Esto quiere decir que $\mathbf{A} \times \mathbf{B}$ (también escrito \mathbf{AB}) no es necesariamente igual a $\mathbf{B} \times \mathbf{A}$ (también escrito \mathbf{BA}).

Que se pueda llevar a cabo la multiplicación $\mathbf{A} \times \mathbf{B}$, no implica ni siquiera que se pueda llevar a cabo la multiplicación $\mathbf{B} \times \mathbf{A}$. Aplicando esto a las matrices \mathbf{B} y \mathbf{C} del ejemplo anterior, la multiplicación $\mathbf{B} \times \mathbf{C}$ es posible:

dimensión de B	dimensión de C	dimensión de B x C
3 x 2	2 x 2	3 x 2
+---- == ----+		

pero la multiplicación $\mathbf{C} \times \mathbf{B}$ no lo es:

dimensión de C	dimensión de B	dimensión de C x B
2 x 2	3 x 2	imposible
+---- != ----+		

- Puede ser que las dos multiplicaciones sean posibles, como en el caso de las matrices \mathbf{A} y \mathbf{B} del ejemplo anterior:

dimensión de A	dimensión de B	dimensión de A x B
2 x 3	3 x 2	2 x 2
+---- == ----+		

dimensión de B	dimensión de A	dimensión de B x A
3 x 2	2 x 3	3 x 3
+---- == ----+		

Pero la matriz $(\mathbf{A} \times \mathbf{B})$, dimensión 2 x 2, no puede ser en ningún caso igual a la matriz $(\mathbf{B} \times \mathbf{A})$, que tiene la dimensión 3 x 3.

- Las matrices $\mathbf{D} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y $\mathbf{E} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$ se pueden multiplicar en los dos sentidos, y en los dos casos el producto es una matriz de dimensión 2×2 . Pero tampoco en este caso es $\mathbf{D} \times \mathbf{E} = \mathbf{E} \times \mathbf{D}$

$$\mathbf{D} \times \mathbf{E} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

$$\mathbf{E} \times \mathbf{D} = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

La multiplicación de dos matrices llevada a cabo en las dos direcciones da lugar al mismo resultado solamente en casos muy especiales. Por ello, al decir que se van a multiplicar dos matrices \mathbf{L} y \mathbf{M} , hay que especificar en qué dirección se va a llevar a cabo la dirección:

- cuando la operación a realizar es $\mathbf{L} \times \mathbf{M}$, se dice que la matriz \mathbf{M} se va a *premultiplicar* por \mathbf{L} , o que se va a multiplicar por \mathbf{L} por la izquierda
- cuando la operación a realizar es $\mathbf{M} \times \mathbf{L}$, se dice que la matriz \mathbf{M} se va a *postmultiplicar* por \mathbf{L} , o que se va a multiplicar por \mathbf{L} por la derecha.

2.3.2.5. MULTIPLICACION Y TRASPOSICION

7. La traspuesta del producto de dos matrices \mathbf{A} y \mathbf{B} es igual al producto de las traspuestas de \mathbf{A} y \mathbf{B} en la dirección opuesta:

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Tomando las matrices \mathbf{B} y \mathbf{C} del ejemplo anterior:

Comprobación de que las dimensiones concuerdan:

$$\begin{array}{ccc} \text{dimensión de B} & \text{dimensión de C} & \text{dimensión de B x C} \\ 3 \times 2 & 2 \times 2 & \rightarrow 3 \times 2 \\ +--- == ----+ \end{array}$$

$$\begin{array}{ccc} \text{dimensión de C'} & \text{dimensión de B'} & \text{dimensión de C' x B'} \\ 2 \times 2 & 2 \times 3 & \rightarrow 2 \times 3 \\ +--- == ----+ \end{array}$$

Comprobación del resultado:

$$\mathbf{BC} = \begin{pmatrix} 19 & 24 \\ 21 & 26 \\ 23 & 28 \end{pmatrix}; (\mathbf{BC})' = \begin{pmatrix} 19 & 21 & 23 \\ 24 & 26 & 28 \end{pmatrix}$$

$$\mathbf{C}'\mathbf{B}' = \begin{pmatrix} 5 & 3 \\ 6 & 4 \end{pmatrix} \begin{pmatrix} 2 & 3 & 4 \\ 3 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 19 & 21 & 23 \\ 24 & 26 & 28 \end{pmatrix}$$

Esta propiedad se puede generalizar del siguiente modo: la traspuesta del producto de varias matrices es igual al producto de las correspondientes traspuestas en la dirección contraria

$$(\mathbf{A} \times \mathbf{B} \times \mathbf{C} \times \mathbf{D})' = \mathbf{D}' \times \mathbf{C}' \times \mathbf{B}' \times \mathbf{A}'$$

o de manera mas concisa:

$$(\mathbf{ABCD})' = \mathbf{D}'\mathbf{C}'\mathbf{B}'\mathbf{A}'$$

2.3.2.6. MULTIPLICACION Y ADICION

8. La multiplicación y la adición de matrices son distributivas:

$$\mathbf{L} \times (\mathbf{M} + \mathbf{N}) = \mathbf{L} \times \mathbf{M} + \mathbf{L} \times \mathbf{N}$$

$$(\mathbf{M} + \mathbf{N}) \times \mathbf{P} = \mathbf{M} \times \mathbf{P} + \mathbf{N} \times \mathbf{P}$$

or

$$\mathbf{L}(\mathbf{M} + \mathbf{N}) = \mathbf{LM} + \mathbf{LN}$$

$$(\mathbf{M} + \mathbf{N})\mathbf{P} = \mathbf{MP} + \mathbf{NP}$$

2.3.2.7. RESUMEN DE LAS REGLAS DEL ALGEBRA DE MATRICES

1. $(\mathbf{A}')' = \mathbf{A}$
2. $\mathbf{A} + \mathbf{B} + \mathbf{C} = (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
3. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
4. $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
5. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
6. \mathbf{AB} no es necesariamente igual a \mathbf{BA}
7. $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$
8. $\mathbf{L}(\mathbf{M} + \mathbf{N}) = \mathbf{LM} + \mathbf{LN}$ y también $(\mathbf{M} + \mathbf{N})\mathbf{P} = \mathbf{MP} + \mathbf{NP}$

2.4. Matrices especiales

Hay matrices que, debido a su estructura, tienen características especiales. Estas características pueden ser muy interesantes o incluso condición necesaria para la resolución de los problemas. En este capítulo se muestran algunas de estas matrices especiales. Su valor o aplicación se verá más adelante.

2.4.1. MATRIZ CUADRADA

Es una matriz con dimensión $n \times n$, es decir que tiene el mismo número de filas que de columnas. La dimensión de una matriz cuadrada se puede representar por un solo número, que se llama “orden” de la matriz.

La *diagonal* (frecuentemente $diag()$) de una matriz cuadrada es el vector de los elementos de la matriz que se encuentran sobre la diagonal que cruza la matriz desde arriba a la izquierda hasta abajo a la derecha.

La *traza* (frecuentemente $tr()$) de una matriz cuadrada es la suma de los elementos de la diagonal.

Ejemplo: una matriz cuadrada \mathbf{A} , de orden 3: $\mathbf{A} = \begin{pmatrix} 3 & 2 & 7 \\ 6 & 5 & 4 \\ 2 & 1 & 6 \end{pmatrix}$

- diagonal de \mathbf{A} : $diag(\mathbf{A}) = (3 \ 5 \ 6)$
- traza de \mathbf{A} : $tr(\mathbf{A}) = 3 + 5 + 6 = 14$

2.4.2. MATRIZ SIMETRICA

Es una matriz cuadrada \mathbf{A} tal que $\mathbf{A} = \mathbf{A}'$. Entonces es $a_{i,j} = a_{j,i}$.

\mathbf{A} tiene que ser cuadrada. Si \mathbf{A} tuviera la dimensión $n \times m$, siendo n distinto de m , la dimensión de \mathbf{A}' sería $m \times n$, por lo que no podría ser $\mathbf{A} = \mathbf{A}'$.

En una matriz simétrica, la mitad que se encuentra a la derecha y arriba de la diagonal es la imagen especular de la parte que se encuentra a la izquierda y abajo de la diagonal.

Ejemplo:

$$\mathbf{A} = \begin{pmatrix} 8 & 1 & 2 & 3 \\ 1 & 4 & 4 & 5 \\ 2 & 4 & 3 & 6 \\ 3 & 5 & 6 & 1 \end{pmatrix} = \mathbf{A}' \quad \begin{aligned} a_{1,3} &= a_{3,1} = 2 \\ a_{2,4} &= a_{4,2} = 5 \end{aligned}$$

2.4.3. MATRIZ TRIANGULAR

Es una matriz cuadrada en la que, o bien los elementos sobre la diagonal, o bien los elementos bajo la diagonal, son todos cero. Ejemplos:

- $\begin{pmatrix} 5 & 0 & 0 & 0 \\ 7 & 4 & 0 & 0 \\ 6 & 5 & 3 & 0 \\ 2 & 1 & 4 & 2 \end{pmatrix}$, que es una *triangular inferior* y
- $\begin{pmatrix} 3 & 6 & 9 \\ 0 & 8 & 10 \\ 0 & 0 & 7 \end{pmatrix}$, que es una *triangular superior*.

2.4.4. MATRIZ DIAGONAL

Es una matriz cuadrada, en la que todos los elementos fuera de la diagonal son ceros. Ejemplos:

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 26 \end{pmatrix}$$

2.4.5. MATRIZ ESCALAR

Es una matriz diagonal, en la que todos los elementos de la diagonal son iguales.

Ejemplo: $\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

La multiplicación de una matriz escalar por otra matriz es igual a la multiplicación del escalar correspondiente por esa matriz. Comprobar que:

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} =$$

$$5 \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 15 & 20 \\ 25 & 30 \end{pmatrix}$$

2.4.6. MATRIZ IDENTIDAD O UNIDAD

Es una matriz escalar representada por \mathbf{I} (y a veces por \mathbf{I}_n para indicar que es de orden n), cuyos elementos sobre la diagonal son unos. En el álgebra de matrices no hay solamente una unidad, sino una unidad para cada orden:

$$\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ etc.}$$

La multiplicación de una matriz \mathbf{A} por la correspondiente matriz unidad \mathbf{I} , es la misma matriz \mathbf{A} :

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

2.4.7. INVERSA DE UNA MATRIZ CUADRADA

La inversa de una matrix cuadrada \mathbf{A} , de orden n , es otra matrix cuadrada \mathbf{A}^{-1} , de la misma dimensión, tal que la multiplicación de \mathbf{A} por su inversa \mathbf{A}^{-1} sea igual a la matriz unidad de dimensión n : $\mathbf{A} \times \mathbf{A}^{-1} = \mathbf{I}_n$.

Ejemplo con $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y su inversa $\mathbf{A}^{-1} = \begin{pmatrix} -20 & 10 \\ 15 & -05 \end{pmatrix}$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -20 & 10 \\ 15 & -05 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

No todas las matrices cuadradas tienen una inversa.

2.4.8. MATRIZ NULA

De dimensión $n \times m$, es una matrix cuyos elementos son todos ceros y se indica como $\mathbf{0}_{n,m}$. Hay tantas matrices nulas como dimensiones son posibles. La suma de una matrix \mathbf{A} y la matrix nula de su misma dimensión, es igual a la matrix \mathbf{A} : $\mathbf{A}_{n,m} + \mathbf{0}_{n,m} = \mathbf{A}_{n,m}$.

2.4.9. MATRICES PARTICIONADAS

Hasta ahora se ha supuesto que los elementos de una matrix eran números reales, o incógnitas. Pero esos elementos también pueden ser matrices. Si agrupamos los elementos de una matrix en submatrices, estamos considerando la matrix original como una matrix particionada.

Ejemplo: Podemos construir la matrix \mathbf{E} , a partir de las matrices \mathbf{A} , \mathbf{B} , \mathbf{C} y \mathbf{D} siguientes:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} \quad \mathbf{C} = (7 \quad 8) \quad \mathbf{D} = (9)$$

$$\mathbf{E} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \left(\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \\ \hline 7 & 8 & 9 \end{array} \right)$$

Podemos considerar a \mathbf{E} como la mariz de los números 1,2,5 etc. o como la matrix de las submatrices \mathbf{A} , \mathbf{B} etc. Para ello es condicion necesaria que:

- número de filas de \mathbf{A} = número de filas de \mathbf{B}
- número de filas de \mathbf{C} = número de filas de \mathbf{D}

- número de columnas de \mathbf{A} = número de columnas de \mathbf{C}
- número de columnas de \mathbf{B} = número de columnas de \mathbf{D}

Con las matrices:

$$\mathbf{F} = \begin{pmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{pmatrix} \text{ y } \mathbf{G} = \begin{pmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{pmatrix}$$

no se puede construir ninguna matriz particionada. Por ejemplo:

$$\left(\begin{array}{cc} \mathbf{F} & \mathbf{G} \end{array} \right) = \begin{pmatrix} 5 & 6 & 7 & 4 & 5 \\ 8 & 9 & 10 & 6 & 7 \\ & & & 8 & 9 \end{pmatrix}$$

no es una ordenación regular de elementos en filas y columnas, donde todas las filas tengan la misma longitud, es decir el mismo número de elementos, es decir, NO ES UNA MATRIZ.

Con matrices particionadas se puede operar exactamente igual que con matrices de elementos simples, pero bajo la condición adicional de que las operaciones con las submatrices cumplan a su vez con las reglas del álgebra de matrices.

Por ejemplo, sean las matrices particionadas \mathbf{E} y \mathbf{L}

$$\mathbf{E} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \quad ; \quad \mathbf{L} = \begin{pmatrix} \mathbf{M} \\ \mathbf{N} \end{pmatrix}$$

se pueden multiplicar de la siguiente forma:

$$\mathbf{E} \times \mathbf{L} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{M} \\ \mathbf{N} \end{pmatrix} = \begin{pmatrix} \mathbf{AM} & \mathbf{BN} \\ \mathbf{CM} & \mathbf{DN} \end{pmatrix}$$

pero solo con la condición de que:

- los productos \mathbf{AM} , \mathbf{BN} , \mathbf{CM} y \mathbf{DN} sean posibles
- la dimensión de \mathbf{AM} sea igual a la dimensión de \mathbf{BN}
- la dimensión de \mathbf{CM} sea igual a la dimensión de \mathbf{DN}

3. Matrices cuadradas

3.1. Determinante de una matriz cuadrada

Sea \mathbf{A} una matriz cuadrada de orden n . El determinante de esa matriz es un número, que corresponde a esa matriz y la caracteriza, dando una idea de algunas de las propiedades más importantes \mathbf{A} . Una teoría rigurosa del determinante es bastante farragosa y no la presentaremos aquí, pero se puede consultar, por ejemplo, en [wikipedia](https://es.wikipedia.org/wiki/Determinante).

El determinante de la matriz cuadrada \mathbf{A} se representa por $|\mathbf{A}|$ y se calcula de la siguiente forma:

- el determinante de \mathbf{A} es la suma de una serie determinada de productos de los elementos de \mathbf{A}
- cada producto es la multiplicación de n elementos de \mathbf{A} , donde n es el orden de \mathbf{A} .
- si de uno de estos productos forma parte un elemento de la fila i de \mathbf{A} , no puede haber ningún otro elemento de esa fila i formando parte de ese mismo producto.
- si de uno de estos productos forma parte un elemento de la columna j de \mathbf{A} , no puede haber ningún otro elemento de esa columna j formando parte de ese mismo producto.
- Se suman todos los productos posibles que cumplen las dos condiciones anteriores.
- antes de sumarse, cada producto recibe un signo positivo o negativo. Más adelante se explicará qué signo recibe cada producto.

Leyendo de esta definición, no es fácil hacerse una idea clara de qué es el determinante de una matriz. Por ello, se va a empezar por calcular determinantes de matrices de dimensión pequeña, que sirvan como ejemplo.

3.2. Deteminante de una matriz cuadrada de orden 2:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$|\mathbf{A}| = a_{11}a_{12} - a_{21}a_{22}$$

Los productos que se suman son todos productos de dos elementos ($2 = \text{orden de } \mathbf{A}$). El primer elemento del primer producto es a_{11} (fila 1, columna 1) y por eso no puede ser el segundo elemento ni de la fila 1 ni de la columna 1. El único elemento de \mathbf{A} que cumple esta condición es a_{22} .

El primer elemento del segundo producto es a_{12} (fila 1, columna 2) y por eso no puede ser el segundo elemento ni de la fila 1 ni de la columna 2. El único elemento de \mathbf{A} que cumple esta condición es a_{21} .

No hay ningún otro producto de dos elementos que cumpla las condiciones.

3.3. Determinante de una matriz cuadrada de orden 3

Sea

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

su determinante aplicando las reglas anteriores ³ es:

$$|\mathbf{A}| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

3.4. Desarrollo del determinante de una matriz segun una fila o una columna

Hemos visto ya unas reglas fáciles para calcular el determinante de matrices de orden 2 o 3. Sin embargo, el cálculo del determinante de una matriz de mayor dimensión puede ser muy complicado.

Hay un procedimiento que facilita el cálculo de determinantes de matrices de orden $n > 3$: determinante según una fila o una columna. No pretendemos aquí demostrar la validez del método, sino simplemente mostrarlo por medio de una matriz de orden 3. Para ello es necesario definir primero los *subdeterminantes* y los *menores* (os *cofactores* o *adjuntos*).

3.4.1. Subdeterminante:

sea \mathbf{B} una matriz de dimensión $m \times n = 4 \times 5$:

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \end{pmatrix}$$

Un subdeterminante de \mathbf{B} es el determinante de una matriz \mathbf{C} , cuadrada de orden p , que está constituida por elementos extraídos de p filas y p columnas de \mathbf{B} por ejemplo:

$$\mathbf{C} = \begin{pmatrix} b_{11} & b_{14} \\ b_{31} & b_{34} \end{pmatrix}$$

que es el determinante de la matriz \mathbf{C} , de los elementos de \mathbf{B} que están en las filas 1 ó 3 y en las columnas 1 o 4.

En capítulos posteriores se vuelven a utilizar subdeterminantes. Para el desarrollo de un determinante hace falta un caso especial de subdeterminante, que se conoce como adjunto o cofactor de un elemento de la matriz.

³que no detallaremos aquí

3.4.2. Adjunto (cofactor)

Sea una matrix cuadrada, por ejemplo la matriz \mathbf{A} , de dimension 3 x 3:

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

El adjunto o cofactor de un elemento $a_{i,j}$ de \mathbf{A} es el determinante de la submatriz $\mathbf{M}_{i,j}$, que resulta de suprimir la fila i y la columna j de \mathbf{A} , es decir, la fila y la columna donde se encuentra elemento $a_{i,j}$. El adjunto del elemento $a_{i,j}$ se representa por $|\mathbf{M}_{i,j}|$ ⁴.

Por ejemplo el adjunto del elemento $a_{2,1}$ es el determinante de la matriz $\mathbf{M}_{2,1}$:

$$\mathbf{M}_{2,1} = \begin{pmatrix} - & a_{12} & a_{13} \\ - & - & - \\ - & a_{32} & a_{33} \end{pmatrix}$$

$$|\mathbf{M}_{2,1}| = a_{12}a_{33} - a_{13}a_{32}$$

El adjunto del elemento $a_{2,2}$ es el determinante de la matriz $\mathbf{M}_{2,2}$:

$$\mathbf{M}_{2,2} = \begin{pmatrix} a_{11} & - & a_{13} \\ - & - & - \\ a_{31} & - & a_{33} \end{pmatrix}$$

$$|\mathbf{M}_{2,2}| = a_{11}a_{33} - a_{13}a_{31}$$

De este modo se pueden hallar los adjuntos de todos los elementos de \mathbf{A} . Por ejemplo si \mathbf{A} es de orden 3, las diferentes submatrices obtenidas borrando una de las filas y una columnas de \mathbf{A} son 9, y los menores, cofactores o adjuntos son los determinantes de esas submatrices.

3.4.3. Desarrollo del determinante de \mathbf{A} según la i -ésima fila:

Es posible calcular el determinante de una matriz como una expansión en la que tomamos una columna de referencia y vamos construyendo un producto en el que se utilizan los adjuntos que hemos presentado previamente. Estos adjuntos son determinantes de orden inferior. Sea la matriz \mathbf{A} y la fila i . La regla es así:

$$|\mathbf{A}| = \sum_{j=1}^n (-1)^{(i+j)} a_{ij} |\mathbf{M}_{ij}|$$

⁴No hay una notación consensual para el adjunto de una matriz. Para evitar confusiones, aquí hemos cambiado la notación original y seguido el ejemplo de SR Searle.

Primero presentaremos la regla considerando la primera fila. Después generalizaremos la regla a cualquier fila o columna.

Consideremos la primera fila de \mathbf{A} .

El determinante de \mathbf{A} es igual a la suma de todos los elementos de la primera fila, multiplicados previamente por su adjunto y provistos de un signo más o menos.

Tomemos la primera fila de \mathbf{A} : $(a_{11} a_{12} a_{13})$.

El determinante de \mathbf{A} será igual a:

$$|\mathbf{A}| = \sum_{j=1}^n (-1)^{(1+j)} a_{1j} |\mathbf{M}_{1j}|$$

La expresión $i + j$ previamente ($i = 1$ en este caso, ya que consideramos la primera columna) da lugar al signo positivo siendo i la fila y j la columna del elemento que, por su adjunto, da lugar a ese sumando. En la primera columna, $i=1$ por ejemplo:

- para el elemento a_{11} : $(-1)^{1+1} = 1$
- para el elemento a_{12} : $(-1)^{1+2} = -1$ etc.

Otra regla más sencilla todavía para encontrar el signo del sumando correspondiente a un elemento es la siguiente:

- el primer elemento arriba a la izquierda de la matriz (1,1) recibe siempre un signo más.
- en la fila 1, el segundo elemento recibe un signo menos, el tercero un signo más, el cuarto signo menos y así sucesivamente, de modo que al avanzar horizontalmente se va cambiando cada vez el signo.
- lo mismo para las columnas verticalmente.

Así en $\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{pmatrix}$ los signos son: $\begin{pmatrix} + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \end{pmatrix}$

Observación : el producto $a_{ij} |\mathbf{M}_{ij}|$ recibe un signo positivo o negativo, pero los valores de a_{ij} y $|\mathbf{M}_{ij}|$ también pueden ser positivos o negativos, de modo que el signo definitivo del sumando se compone del producto de los tres signos.

Vamos a llevar a cabo el desarrollo del determinante de una matriz de orden 3 según la segunda fila.

La matriz es:

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

La segunda fila de \mathbf{A} es $(a_{11}a_{12}a_{13})$.

- los signos para los sumandos son $(-, +, -)$ de acuerdo a lo anterior.
- los adjuntos de las elementos de la segunda fila son:

$$|\mathbf{M}_{21}| = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = a_{12}a_{33} - a_{13}a_{32}$$

$$|\mathbf{M}_{22}| = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}a_{33} - a_{13}a_{31}$$

$$|\mathbf{M}_{23}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = a_{11}a_{32} - a_{12}a_{31}$$

y el determinante de \mathbf{A} es:

$$\begin{aligned} |\mathbf{A}| &= \\ &- a_{21}|\mathbf{M}_{21}| + a_{22}|\mathbf{M}_{22}| - a_{23}|\mathbf{M}_{23}| = \\ &- a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) - a_{23}(a_{11}a_{32} - a_{12}a_{31}) = \\ &- a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} + a_{22}a_{11}a_{33} - a_{22}a_{13}a_{31} - a_{23}a_{11}a_{32} + a_{12}a_{23}a_{31} \end{aligned}$$

Se puede comprobar que los sumandos aquí y en el apartado 2.1 son los mismos y que los signos coinciden.

Otrs posibilidades de calcular el determinante seria desarrollarlo según:

- la primera fila: $|\mathbf{A}| = a_{11}|\mathbf{M}_{11}| - a_{12}|\mathbf{M}_{12}| + a_{13}|\mathbf{M}_{13}|$
- la tercera fila: $|\mathbf{A}| = a_{31}|\mathbf{M}_{31}| - a_{32}|\mathbf{M}_{32}| + a_{33}|\mathbf{M}_{33}|$
- la primera columna: $|\mathbf{A}| = a_{11}|\mathbf{M}_{11}| - a_{21}|\mathbf{M}_{21}| + a_{31}|\mathbf{M}_{31}|$
- la segunda columna: $|\mathbf{A}| = -a_{12}|\mathbf{M}_{12}| + a_{22}|\mathbf{M}_{22}| - a_{32}|\mathbf{M}_{32}|$
- la tercera columna: $|\mathbf{A}| = +a_{13}|\mathbf{M}_{13}| - a_{23}|\mathbf{M}_{23}| + a_{33}|\mathbf{M}_{33}|$

3.5. Dificultad del cálculo del determinante de una matriz de orden n

El cálculo de determinantes de matrices de orden superior a tres es difícil de hacer si no es por medio del desarrollo según una fila o columna.

Para hallar por medio de un desarrollo el determinante de una matriz de orden 4, es necesario calcular cuatro determinantes de orden 3 (los cuatro adjuntos de los elementos de la fila o columna).

Para hallar por medio de un desarrollo el determinante de una matriz de orden 5, es necesario calcular cinco determinantes de orden 4 (los cinco adjuntos), es decir 5 x 4 determinantes de orden 3.

Para hallar por medio de un desarrollo el determinante de una matriz de orden 6, es necesario calcular seis determinantes de orden 5 (los seis adjuntos), es decir 6 x 5 x 4 determinantes de orden 3.

Para hallar por medio de un desarrollo el determinante de una matriz de orden n , es necesario un número de : $n \times (n-1) \times (n-2) \times \dots \times 4 \times 3 \times 2 \times 1 = n!$ determinantes de orden 3. Con la dimensión de 1 a matriz, aumenta la dificultad y el costo del calculo del determinante de manera mas que exponencial.

Para poder hallar determinantes de matrices de orden grande son necesarios otros algoritmos, que no se van a explicar aqui. En la práctica existen rutinas para el cálculo por ordenador. Casos especiales son las matrices triangulares y las matrices diagonales. Cuando la matriz cuyo determinante se quiere calcular tiene alguna fila o columna en la que se encuentre un numero elevado de elementos iguales a cero, el desarrollo del determinante según esa fila o columna se simplifica notablemente, porque al tenerse que multiplicar por cero, no es necesario hallar los adjuntos de los elementos nulos. El caso más simple es el de una matriz triangular:

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}|\mathbf{M}_{11}| - 0|\mathbf{M}_{12}| + 0|\mathbf{M}_{13}| = a_{11}|\mathbf{M}_{11}|$$

$$a_{11}|\mathbf{M}_{11}| = a_{11} \begin{vmatrix} a_{22} & 0 \\ a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}$$

EL DETERMINANTE DE UNA MATRIZ TRIANGULAR ES IGUAL AL PRODUCTO DE LOS ELEMENTOS DE LA DAGONAL

Como las matrices diagonales son un caso especial de las triangulares, lo mismo se puede aplicar a las matrices diagonales

3.6. Propiedades de los determinantes

Los determinantes tienen las siguientes propiedades:

1. el determinante de la traspuesta de una matriz \mathbf{A} , es igual al determinante de \mathbf{A} : $|\mathbf{A}'| = |\mathbf{A}|$.

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

Ojo, el elemento $(1, 2)$ de \mathbf{A} es a_{12} , pero el elemento $(1, 2)$ de \mathbf{A}' es a_{21} , dado que es la traspuesta de \mathbf{A} . Utilizando la regla del determinante de una matriz de orden 3:

$$\begin{aligned}
 |\mathbf{A}| &= \\
 &a_{11}a_{22}a_{33} \quad (1) \\
 &+a_{12}a_{23}a_{31} \quad (2) \\
 &+a_{13}a_{21}a_{32} \quad (3) \\
 &-a_{13}a_{22}a_{31} \quad (4) \\
 &-a_{12}a_{21}a_{33} \quad (5) \\
 &-a_{11}a_{23}a_{32} \quad (6)
 \end{aligned}$$

y el determinante de la traspuesta es:

$$\begin{aligned}
 |\mathbf{A}'| &= \\
 &a_{11}a_{22}a_{33} \quad (1) \\
 &+a_{21}a_{32}a_{13} \quad (3) \\
 &+a_{31}a_{12}a_{23} \quad (2) \\
 &-a_{31}a_{22}a_{13} \quad (4) \\
 &-a_{21}a_{12}a_{33} \quad (5) \\
 &-a_{11}a_{32}a_{23} \quad (6)
 \end{aligned}$$

En los dos casos se encuentran los mismos productos con los mismos signos.

Una vez comprobada esta propiedad, ya no es necesario diferenciar entre filas y columnas. Al describir las propiedades de los determinantes se hablará siempre de filas, pero exactamente lo mismo valdrá para las columnas.

2. El determinante de una matriz que tiene dos filas iguales, es cero.

En efecto, sea $\mathbf{A} = \begin{pmatrix} a & b & c \\ a & b & c \\ d & e & f \end{pmatrix}$

El determinante de \mathbf{A} será:

$$|\mathbf{A}| = abf + bcd + cae - cbd - baf - ace = 0$$

Los mismos producto aparecen dos veces, pero cada vez con signo distinto y el resultado es cero.

3. Si se multiplican todos los elementos de una fila de \mathbf{A} por un escalar k , el determinante de la nueva matrix \mathbf{B} es igual al producto de k por el determinante de \mathbf{A} : $|k\mathbf{A}| = k|\mathbf{A}|$

$$\mathbf{B} = \begin{pmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\begin{aligned} |\mathbf{B}| = & \\ & ka_{11}a_{22}a_{33} \\ & +ka_{12}a_{23}a_{31} \\ & +ka_{13}a_{21}a_{32} \\ & -ka_{13}a_{22}a_{31} \\ & -ka_{12}a_{21}a_{33} \\ & -ka_{11}a_{23}a_{32} \end{aligned}$$

En $|\mathbf{B}|$ se encuentran los mismos sumandos que en $|\mathbf{A}|$ pero multiplicados todos ellos por el escalar k .

4. Si se intercambian dos filas contiguas de \mathbf{A} , el determinante de la matriz resultante \mathbf{B} , es igual al determinante de \mathbf{A} , cambiado de signo: $|\mathbf{B}| = -|\mathbf{A}|$

En efecto, sean \mathbf{A} y \mathbf{B}

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

$$\begin{aligned} |\mathbf{A}| = & \\ & a_{11}a_{22}a_{33} \quad (1) \\ & +a_{12}a_{23}a_{31} \quad (2) \\ & +a_{13}a_{21}a_{32} \quad (3) \\ & -a_{13}a_{22}a_{31} \quad (4) \\ & -a_{12}a_{21}a_{33} \quad (5) \\ & -a_{11}a_{23}a_{32} \quad (6) \end{aligned}$$

$$\begin{aligned} |\mathbf{B}| = & \\ & a_{21}a_{12}a_{33} \quad (5) \\ & +a_{22}a_{13}a_{31} \quad (4) \\ & +a_{23}a_{11}a_{32} \quad (6) \\ & -a_{23}a_{12}a_{31} \quad (2) \\ & -a_{22}a_{11}a_{33} \quad (1) \\ & -a_{21}a_{13}a_{32} \quad (3) \end{aligned}$$

Los productos en los dos casos son los mismos, pero si empre con signo contrario.

5. Sea \mathbf{A} , \mathbf{B} y \mathbf{C} tres matrices del mismo orden, que son iguales en todo menos en una fila y tales que los elementos de \mathbf{C} para esa fila son además iguales a la suma de los elementos homólogos de \mathbf{A} y \mathbf{B} . En este caso, el determinante de \mathbf{C} es igual a la suma de los determinantes de \mathbf{A} y \mathbf{B} :

Por ejemplo:

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

$$\text{y } \mathbf{C} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

El determinante de \mathbf{C} será:

$$\begin{aligned} |\mathbf{C}| &= c_{11}a_{22}a_{33} \\ &\quad + c_{12}a_{23}a_{31} \\ &\quad + c_{13}a_{21}a_{32} \\ &\quad - c_{13}a_{22}a_{31} \\ &\quad - c_{12}a_{21}a_{33} \\ &\quad - c_{11}a_{23}a_{32} \end{aligned}$$

Si sustituimos $c_{11} = a_{11} + b_{11}$, $c_{12} = a_{12} + b_{12}$, $c_{13} = a_{13} + b_{13}$ queda:

$$\begin{aligned} |\mathbf{C}| &= a_{11}a_{22}a_{33} \\ &\quad + a_{12}a_{23}a_{31} \\ &\quad + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} \\ &\quad - a_{12}a_{21}a_{33} \\ &\quad - a_{11}a_{23}a_{32} \\ &\quad + b_{11}a_{22}a_{33} \\ &\quad + b_{12}a_{23}a_{31} \\ &\quad + b_{13}a_{21}a_{32} \\ &\quad - b_{13}a_{22}a_{31} \\ &\quad - b_{12}a_{21}a_{33} \\ &\quad - b_{11}a_{23}a_{32} \end{aligned} = |\mathbf{A}| + |\mathbf{B}|$$

6. Transformemos una matriz \mathbf{A} en otra matriz \mathbf{B} , sumando a cada uno de los elementos de una fila de \mathbf{A} , el elemento que se encuentra en su misma columna pero de otra fila determinada de \mathbf{A} y multiplicado por un escalar k , dejando el resto de filas igual. Entonces el determinante de la nueva matriz \mathbf{B} es igual al determinante de \mathbf{A} : $|\mathbf{A}| = |\mathbf{B}|$

Por ejemplo:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \mathbf{B} = \begin{pmatrix} a_{11} + ka_{21} & a_{12} + ka_{22} & a_{13} + ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Según la quinta propiedad de los determinantes, el determinante de \mathbf{B} es igual al determinante de \mathbf{A} , más el determinante de una matriz \mathbf{C} , de la siguiente forma:
 $|\mathbf{B}| = |\mathbf{A}| + |\mathbf{C}|$

$$\mathbf{C} = \begin{pmatrix} ka_{21} & ka_{22} & ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Según la tercera propiedad de los determinantes, el determinante de \mathbf{C} es igual a

$$|\mathbf{C}| = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Y según la segunda propiedad de los determinantes: $|\mathbf{C}| = k0 = 0$ luego:

$$|\mathbf{B}| = |\mathbf{A}| + |\mathbf{C}| = |\mathbf{A}| + 0 = |\mathbf{A}|$$

3.7. Matriz adjunta e inversa de una matriz

3.7.1. Definición de matriz adjunta de \mathbf{A} : $\text{adj}(\mathbf{A})$

Se llama matriz adjunta de \mathbf{A} , a la traspuesta de la matriz de los adjuntos de los elementos de \mathbf{A} , $|\mathbf{M}_{ij}|$, multiplicados por $(-1)^{(i+j)}$. Para escribir la matriz adjunta de \mathbf{A} , seguimos los siguientes pasos :

1. Escribimos la matriz de los adjuntos:

$$\begin{pmatrix} |\mathbf{M}_{11}| & |\mathbf{M}_{12}| & |\mathbf{M}_{13}| \\ |\mathbf{M}_{21}| & |\mathbf{M}_{22}| & |\mathbf{M}_{23}| \\ |\mathbf{M}_{31}| & |\mathbf{M}_{32}| & |\mathbf{M}_{33}| \end{pmatrix}$$

2. Multiplicamos los adjuntos por $(-1)^{(i+j)}$ es decir, les adjudicamos un signo del mismo modo que hemos hecho para el cálculo del determinante:

$$\begin{pmatrix} |\mathbf{M}_{11}| & -|\mathbf{M}_{12}| & |\mathbf{M}_{13}| \\ -|\mathbf{M}_{21}| & |\mathbf{M}_{22}| & -|\mathbf{M}_{23}| \\ |\mathbf{M}_{31}| & -|\mathbf{M}_{32}| & |\mathbf{M}_{33}| \end{pmatrix}$$

3. Trasponemos esta última matriz:

$$\text{adj}(\mathbf{A}) = \begin{pmatrix} |\mathbf{M}_{11}| & -|\mathbf{M}_{21}| & |\mathbf{M}_{31}| \\ -|\mathbf{M}_{12}| & |\mathbf{M}_{22}| & -|\mathbf{M}_{32}| \\ |\mathbf{M}_{13}| & -|\mathbf{M}_{23}| & |\mathbf{M}_{33}| \end{pmatrix}$$

$\text{adj}(\mathbf{A})$ es la matriz adjunta de \mathbf{A} .

Observación : un adjunto, lo mismo que un determinante, es un número real. Los elementos de la matriz adjunta por lo tanto son números reales.

3.7.2. Multiplicación de una matriz cuadrada \mathbf{A} por su matriz adjunta $\text{adj}(\mathbf{A})$

Vamos a ver cual es el resultado de postmultiplicar una matriz cuadrada \mathbf{A} por su adjunta $\text{adj}(\mathbf{A})$:

$$\mathbf{A} \text{adj}(\mathbf{A}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} |\mathbf{M}_{11}| & -|\mathbf{M}_{21}| & |\mathbf{M}_{31}| \\ -|\mathbf{M}_{12}| & |\mathbf{M}_{22}| & -|\mathbf{M}_{32}| \\ |\mathbf{M}_{13}| & -|\mathbf{M}_{23}| & |\mathbf{M}_{33}| \end{pmatrix}$$

La matriz resultante es una matriz \mathbf{B} , cuadrada de orden 3. Los elementos de la diagonal de \mathbf{B} son los siguientes:

$$\begin{aligned} b_{11} &= a_{11}|\mathbf{M}_{11}| + a_{12}|\mathbf{M}_{12}| + a_{13}|\mathbf{M}_{13}| \\ b_{22} &= -a_{21}|\mathbf{M}_{21}| + a_{22}|\mathbf{M}_{22}| + a_{23}|\mathbf{M}_{23}| \\ b_{33} &= a_{31}|\mathbf{M}_{31}| + a_{32}|\mathbf{M}_{32}| + a_{33}|\mathbf{M}_{33}| \end{aligned}$$

expresiones correspondientes al desarrollo del determinante de \mathbf{A} según la primera, segunda y tercera filas respectivamente. Es decir, los elementos de la diagonal de \mathbf{B} son todos iguales e iguales al determinante de \mathbf{A} .

Vamos a ver ahora cuánto vale un elemento fuera de la diagonal, por ejemplo el elemento b_{12} :

$$\begin{aligned} b_{12} &= -a_{11}|\mathbf{M}_{21}| + a_{12}|\mathbf{M}_{22}| - a_{13}|\mathbf{M}_{23}| = \\ &= -a_{11} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{12} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{13} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} = \\ &= -a_{11}(a_{12}a_{33} - a_{13}a_{32}) \\ &+ a_{12}(a_{11}a_{33} - a_{13}a_{31}) \\ &- a_{13}(a_{11}a_{31} - a_{12}a_{32}) \\ &= -a_{11}a_{12}a_{33} + a_{11}a_{13}a_{32} + a_{12}a_{11}a_{33} \\ &- a_{12}a_{13}a_{31} - a_{13}a_{11}a_{31} + a_{13}a_{12}a_{32} = 0 \end{aligned}$$

O de otro modo, se comprueba que b_{12} es igual al determinante de la siguiente matriz:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

que por la segunda propiedad de los determinantes, es igual a cero.

Se puede comprobar como todos los demás elementos no diagonales de \mathbf{B} son iguales al determinante de una matriz que tiene dos filas iguales, es decir, son todos cero.

De modo que el producto de \mathbf{A} por su adjunta $adj(\mathbf{A})$ es una matriz en la que los elementos de la diagonal son iguales al escalar determinante de \mathbf{A} :

$$\mathbf{A} \, adj(\mathbf{A}) = \begin{pmatrix} |\mathbf{A}| & 0 & 0 \\ 0 & |\mathbf{A}| & 0 \\ 0 & 0 & |\mathbf{A}| \end{pmatrix} = |\mathbf{A}| \, \mathbf{I}_3$$

3.7.3. Inversa de una matriz cuadrada \mathbf{A}

El producto de la matriz cuadrada \mathbf{A} , de orden n por su matriz adjunta, es igual a la matriz unidad de orden n , multiplicada por el escalar $|\mathbf{A}|$ (determinante de \mathbf{A}).

$$\mathbf{A} \, adj(\mathbf{A}) = |\mathbf{A}| \, \mathbf{I}_n$$

Como el determinante de \mathbf{A} es un escalar, se puede transformar la expresión anterior, de la siguiente forma:

$$\mathbf{A} \frac{1}{|\mathbf{A}|} adj(\mathbf{A}) = \mathbf{I}_n$$

Según la definición de inversa de \mathbf{A} (ver **INVERSA DE UNA MATRIZ CUADRADA**), la expresión

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} adj(\mathbf{A})$$

es la inversa de \mathbf{A} porque el producto de \mathbf{A} por ella es igual a la matriz unidad de orden n .

3.7.4. Inversa de una matriz diagonal

Una matriz diagonal es especialmente fácil de invertir. Sea

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

Se puede comprobar fácilmente que su inversa es:

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & 0 & 0 \\ 0 & \frac{1}{a_{22}} & 0 \\ 0 & 0 & \frac{1}{a_{33}} \end{pmatrix}$$

La inversa de una matriz diagonal \mathbf{A} es otra matriz diagonal \mathbf{B} cuyos elementos no nulos son iguales a los inversos de los elementos homólogos de \mathbf{A} : $b_{ii} = 1/a_{ii}$.

3.7.5. Inversa del producto de dos matrices cuadradas

Sea la matriz cuadrada \mathbf{C} igual al producto de dos matrices cuadradas \mathbf{A} y \mathbf{B} : $\mathbf{C} = \mathbf{AB}$. La matriz $\mathbf{D} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ es la matriz inversa de \mathbf{C} , ya que el producto de \mathbf{D} por \mathbf{C} es igual a la matriz unidad: $\mathbf{DC} = \mathbf{I}_n$.

En efecto:

$$\mathbf{DC} = \mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{AB} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}_n$$

La inversa del producto de dos matrices cuadradas es igual al producto de las inversas de las dos matrices en dirección contraria.

4. Sistemas de ecuaciones

4.1. Ecuaciones en el álgebra de las matrices

4.1.1. Definición de un sistema de ecuaciones y su solución

Un sistema de ecuaciones en el álgebra de los números reales es un conjunto de ecuaciones formadas por números reales e incógnitas, como por ejemplo el siguiente:

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$2 \times x_1 + 2 \times x_2 + 2 \times x_3 = 12$$

Una solución del sistema de ecuaciones es un conjunto de números reales tales que, sustituyendo las incógnitas por esos números, todas las ecuaciones sean verdaderas al mismo tiempo.

En nuestro ejemplo, se puede sustituir la incógnita x_1 por el número 1, x_2 por 2 y x_3 por 3, siendo verdadero que:

$$4 \times 1 + 5 \times 2 - 3 \times 3 = 5$$

$$1 - 2 + 2 \times 3 = 5$$

$$2 \times 1 + 2 \times 2 + 2 \times 3 = 12$$

Por lo tanto $(x_1 = 1; x_2 = 2; x_3 = 3)$ es una solución del sistema de ecuaciones.

Si el sistema de ecuaciones tiene menos ecuaciones que incógnitas, se puede encontrar un número infinito de combinaciones de números que cumplen las ecuaciones. Por ejemplo si el sistema es:

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$(x_1 = 1; x_2 = 2; x_3 = 3)$ sería una solución del sistema, pero también lo sería $(x_1 = 23/9; x_2 = -4/9; x_3 = 1)$, lo mismo que $(x_1 = 18/11; x_2 = 1; x_3 = 24/11)$ y otras muchas combinaciones de números.

En la valoración genética de los animales también se trabaja con sistemas de ecuaciones y éstas se pueden resolver más fácilmente en el álgebra de matrices.

4.1.2. Un sistemas de ecuaciones en notación matricial

Con ayuda de la operación “multiplicación de matrices”, se puede escribir un sistema de ecuaciones en notación matricial .

Tomando los primeros miembros de las ecuaciones (obviando las incógnitas), se obtiene la matriz de los coeficientes **A**:

$$\mathbf{A} = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

Los valores a la derecha de los signos de igualdad así como las incógnitas se pueden escribir como vectores columna:

$$\mathbf{y} = \begin{bmatrix} 5 \\ 5 \\ 12 \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Entonces, el sistema de ecuaciones en notación matricial será:

$$\mathbf{A} \times \mathbf{b} = \mathbf{y}$$

o lo que es lo mismo:

$$\begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 12 \end{bmatrix}$$

Efectivamente, llevando a cabo la multiplicación de \mathbf{A} por \mathbf{b} , resulta:

$$\begin{bmatrix} 4 \times x_1 + 5 \times x_2 - 3 \times x_3 \\ x_1 - x_2 + 2 \times x_3 \\ 2 \times x_1 + 2 \times x_2 + 2 \times x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 12 \end{bmatrix}$$

Un sistema de ecuaciones en notación matricial queda reducido a una sola ecuación.

4.1.3. Solución de ecuaciones en el álgebra de las matrices.

Sea el sistema de ecuaciones en notación matricial:

$$\mathbf{A}\mathbf{b} = \mathbf{y}$$

Para hallar una solución al sistema, es decir valores para \mathbf{b} , hay que despejar \mathbf{b} en la ecuación. Esto se puede hacer mediante la inversa de \mathbf{A} .

Multiplicando por la izquierda a ambos lados del signo igualdad por la matriz \mathbf{A}^{-1} resulta:

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{b} = \mathbf{A}^{-1}\mathbf{y}$$

Como la multiplicación es asociativa:

$$(\mathbf{A}^{-1}\mathbf{A})\mathbf{b} = \mathbf{A}^{-1}\mathbf{y}$$

Por la definición de inversa, $\mathbf{A}^{-1}\mathbf{A}$ es la matriz identidad o unidad:

$$\mathbf{I}\mathbf{b} = \mathbf{A}^{-1}\mathbf{y}$$

Y una propiedad de la matriz identidad es que $\mathbf{I}\mathbf{b} = \mathbf{b}$, luego la solución es:

$$\mathbf{b} = \mathbf{A}^{-1}\mathbf{y}$$

En el ejemplo:

- El determinante de \mathbf{A} es:

$$|\mathbf{A}| = 4 \times (-1) \times 2 + 5 \times 2 \times 2 + (-3) \times 1 \times 2 - (-3) \times (-1) \times 2 - 5 \times 1 \times 2 - 4 \times 2 \times 2 = -26$$

- La inversa de \mathbf{A} es:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{adj}(\mathbf{A}) = -1/26 \begin{bmatrix} -6 & -16 & 7 \\ 2 & 14 & -11 \\ 4 & 2 & -9 \end{bmatrix}$$

- Y la solución:

$$\mathbf{b} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -1/26 \begin{bmatrix} -6 & -16 & 7 \\ 2 & 14 & -11 \\ 4 & 2 & -9 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \\ 12 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

coincide con la que ya conocíamos.

4.2. Rango de la matriz de coeficientes de un sistema ecuaciones

4.2.1. Sistemas de ecuaciones con un numero infinito de soluciones

4.2.1.1. Algebras de los números reales Ya se ha visto que un sistema de ecuaciones con más incógnitas que ecuaciones, tiene un número infinito de soluciones.

También puede ocurrir que en un sistema de ecuaciones con el mismo número de incógnitas que de ecuaciones, una de éstas sea una combinación lineal de las otras. En este caso, esta ecuación no aporta nueva información al sistema y por lo tanto es innecesaria. El número de ecuaciones del sistema es realmente inferior al número de incógnitas.

Partiendo del segundo sistema de ecuaciones presentado, podemos construir los siguientes nuevos sistemas:

(1)

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

en el que la tercera ecuación es igual a una vez la primera más cero veces la segunda

(2)

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$6 \times x_1 + 3 \times x_2 - x_3 = 5$$

en el que la tercera ecuación es igual a la primera más dos veces la segunda y

(3)

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$7 \times x_1 + 11 \times x_2 - 8x_3 = -15$$

en el que la tercera ecuación es igual a dos veces la primera menos la segunda.

Aunque estos sistemas de ecuaciones tienen tres ecuaciones cada uno, la tercera ecuación no aporta en ningún caso nueva información, porque es linealmente dependiente de las otras. Los tres sistemas tienen un número infinito de soluciones. Si queremos resolverlos por los métodos clásicos del álgebra de los números reales, llegaremos siempre a la expresión $0 = 0$. Sólo encontraremos una solución, si previamente damos un valor arbitrario a x_1 , x_2 o x_3 , o si añadimos arbitrariamente una tercera ecuación independiente. Pero el número de soluciones posibles es infinito.

4.2.1.2. Álgebra de las matrices Las matrices de coeficientes de los tres sistemas de ecuaciones anteriores son las siguientes:

$$\mathbf{A}_1 = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 4 & 5 & -3 \end{bmatrix}; \mathbf{A}_2 = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 6 & 3 & 1 \end{bmatrix}; \mathbf{A}_3 = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 7 & 11 & -6 \end{bmatrix}$$

En las tres matrices \mathbf{A}_1 , \mathbf{A}_2 y \mathbf{A}_3 se puede comprobar que hay una dependencia lineal entre las filas. Por ejemplo:

$$\mathbf{A}_2 = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 4 + 2 \times 1 & 5 + 2 \times (-1) & (-3) + 2 \times 2 \end{bmatrix}$$

Así, según la quinta propiedad de los determinantes, el determinante de \mathbf{A}_2 es:

$$|\mathbf{A}_2| = \begin{vmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 4 & 5 & -3 \end{vmatrix} + \begin{vmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ (2 \times 1) & (2 \times (-1)) & (2 \times 2) \end{vmatrix}$$

Y según la tercera propiedad de los determinantes:

$$|\mathbf{A}_2| = \begin{vmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 4 & 5 & -3 \end{vmatrix} + 2 \begin{vmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 1 & -1 & 2 \end{vmatrix}$$

De acuerdo con la segunda propiedad de los determinantes, estos dos últimos determinantes son cero, de modo que:

$$|\mathbf{A}_2| = 0$$

Del mismo modo son $|\mathbf{A}_1| = |\mathbf{A}_3| = |\mathbf{A}_2| = 0$

Para resolver los tres sistemas de ecuaciones necesitamos hallar las inversas de \mathbf{A}_1 , \mathbf{A}_3 y \mathbf{A}_2 .

$$\mathbf{A}_1^{-1} = \frac{1}{|\mathbf{A}_1|} \text{adj}(\mathbf{A}_1)$$

Como $|\mathbf{A}_1| = 0$, $(1/|\mathbf{A}_1|) = \infty$ y no se puede encontrar ninguna inversa de \mathbf{A}_1 . Una matriz con determinante nulo, no tiene inversa y por lo tanto no se puede resolver el sistema de ecuaciones correspondiente. Una matriz de este tipo se llama MATRIZ SINGULAR. En el álgebra de matrices, del mismo modo que en el álgebra de los números reales, no se puede encontrar una solución única para estos tres sistemas de ecuaciones.

4.2.2. Rango de una matriz

Siempre que en un sistema de ecuaciones el número de ecuaciones independientes sea inferior al número de incógnitas, el determinante de la matriz de coeficientes correspondiente es nulo (0).

Si el número de ecuaciones independientes es igual al número de incógnitas (n) menos 2 (es decir $n - 2$), el determinante de la matriz de los coeficientes es nulo y además todos los subdeterminantes de dimensión $(n - 1)$ son también nulos.

Si el número de ecuaciones independientes es $(n - 3)$, serán también nulos todos los subdeterminantes de dimensión $(n - 2)$ etc.

Sea por ejemplo la matriz:

$$\mathbf{A} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{vmatrix}$$

esta matriz tiene dimensión 3, pero dos dependencias lineales, es decir, solamente tiene una ecuación independiente.

El determinante de \mathbf{A} es:

$$|\mathbf{A}| = 1 \times 4 \times 9 + 2 \times 6 \times 3 + 3 \times 2 \times 6 - 1 \times 6 \times 6 - 2 \times 2 \times 9 - 3 \times 4 \times 3 = 0$$

Y todos los subdeterminantes de orden 2 son también nulos:

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 1 \times 4 - 2 \times 2 = 0; \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 1 \times 6 - 2 \times 3 = 0$$

$$\begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix} = 2 \times 6 - 4 \times 3 = 0; \begin{vmatrix} 1 & 2 \\ 3 & 6 \end{vmatrix} = 1 \times 6 - 3 \times 2 = 0$$

$$\begin{vmatrix} 1 & 3 \\ 3 & 9 \end{vmatrix} = 1 \times 9 - 3 \times 3 = 0; \begin{vmatrix} 2 & 3 \\ 6 & 9 \end{vmatrix} = 2 \times 9 - 6 \times 3 = 0$$

$$\begin{vmatrix} 2 & 4 \\ 3 & 6 \end{vmatrix} = 2 \times 6 - 3 \times 4 = 0; \begin{vmatrix} 2 & 6 \\ 3 & 9 \end{vmatrix} = 2 \times 9 - 3 \times 6 = 0$$

$$\begin{vmatrix} 4 & 6 \\ 6 & 9 \end{vmatrix} = 4 \times 9 - 6 \times 6 = 0;$$

En cambio los elementos de la matriz, que se pueden considerar como subdeterminantes de dimensión 1, no son todos nulos. El rango de esta matriz es 1, que es la dimensión del mayor subdeterminante no nulo.

EL RANGO DE UNA MATRIZ ES IGUAL A LA DIMENSION DE SU MAYOR SUBDETERMINANTE NO NULO Y SE CORRESPONDE CON EL NUMERO DE ECUACIONES INDEPENDIENTES QUE TIENE EL SISTEMA DE ECUACIONES CORRESPONDIENTE.

Una matriz singular, cuyo determinante es cero, tiene un rango incompleto.

Una matriz cuyo rango es igual a su dimensión, tiene rango completo.

Observación: una matriz no cuadrada, también tiene rango. Por ejemplo la matriz $\begin{pmatrix} 1 & 2 & 3 \\ 6 & 9 & 8 \end{pmatrix}$ es de rango 2, dado que, por ejemplo, la submatriz de dimensión 2 $\begin{pmatrix} 1 & 2 \\ 6 & 9 \end{pmatrix}$ tiene un determinante de -3.

4.3. Operadores elementales y sistemas de ecuaciones

4.3.1. Operaciones elementales en filas y columnas

Las siguientes operaciones con las filas o columnas de una matriz \mathbf{A} , se llaman operaciones elementales:

1. intercambio de dos filas (columnas)

2. multiplicación de todos los elementos de una fila (columna) por un escalar k
3. adición a una fila (columna), de otra multiplicada por un escalar

Cuando se lleva a cabo alguna de estas operaciones, la matriz original \mathbf{A} se transforma en otra matriz \mathbf{B} , cuyas filas (columnas) son combinaciones lineales de las filas (columnas) de \mathbf{A} . Entonces se dice que \mathbf{A} y \mathbf{B} son MATRICES EQUIVALENTES.

El determinante de la matriz \mathbf{B} no es independiente del determinante de \mathbf{A} . A partir de las propiedades de los determinantes se pueden comprobar fácilmente las siguientes afirmaciones:

- Si \mathbf{B} es el resultado de una operación elemental del primer o tercer tipo, entonces el determinante de \mathbf{B} es igual al determinante de \mathbf{A} : $|\mathbf{B}| = |\mathbf{A}|$
- Si \mathbf{B} es el resultado de una operación lineal del segundo tipo, el determinante de \mathbf{B} es igual al determinante de \mathbf{A} multiplicado por el escalar k : $|\mathbf{B}| = k |\mathbf{A}|$

Una operación elemental con filas de \mathbf{A} se lleva a cabo premultiplicando \mathbf{A} por una determinada matriz cuadrada no singular \mathbf{L} . En la i -ésima fila de \mathbf{L} se encuentran los escalares por los que hay que multiplicar cada fila de \mathbf{A} para obtener la i -ésima fila de \mathbf{B} .

Por ejemplo, para intercambiar las filas primera y segunda en la matriz \mathbf{A} , dando lugar a \mathbf{B}_1 :

$$\mathbf{A} = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix}; \mathbf{B}_1 = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 5 & -3 \\ 2 & 2 & 2 \end{bmatrix}$$

La primera fila de \mathbf{B}_1 es igual a:

$$\begin{aligned} & 0 \times \text{la primera fila de } \mathbf{A} \\ + & 1 \times \text{la segunda fila de } \mathbf{A} \\ + & 0 \times \text{la tercera fila de } \mathbf{A} \end{aligned}$$

La segunda fila de \mathbf{B}_1 es igual a:

$$\begin{aligned} & 1 \times \text{la primera fila de } \mathbf{A} \\ + & 0 \times \text{la segunda fila de } \mathbf{A} \\ + & 0 \times \text{la tercera fila de } \mathbf{A} \end{aligned}$$

La tercera fila de \mathbf{B}_1 es igual a:

$$\begin{aligned}
& 0 \times \text{la primera fila de } \mathbf{A} \\
+ & 0 \times \text{la segunda fila de } \mathbf{A} \\
+ & 1 \times \text{la tercera fila de } \mathbf{A}
\end{aligned}$$

Los escalares por los que hay que multiplicar cada fila de \mathbf{A} para obtener las filas de \mathbf{B}_1 son por lo tanto:

- para la primera fila de \mathbf{B}_1 : 0, 1 y 0
- para la segunda fila de \mathbf{B}_1 : 1, 0 y 0
- para la tercera fila de \mathbf{B}_1 : 0, 0 y 1

$$\mathbf{L}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

\mathbf{L}_1 es el operador elemental que intercambia las filas 1 y 2:

$$\mathbf{L}_1 \mathbf{A} = \mathbf{B}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 5 & -3 \\ 2 & 2 & 2 \end{bmatrix}$$

Un segundo ejemplo: ahora queremos transformar la matriz \mathbf{A} en otra matriz \mathbf{B}_2 . La tercera fila de \mathbf{B}_2 ha de ser igual a la suma de la primera fila de \mathbf{A} multiplicada por 5 y de la segunda fila de \mathbf{A} .

$$\mathbf{B}_2 = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 5 \times 4 + 2 & 5 \times 5 + 2 & 5 \times (-3) + 2 \end{bmatrix} = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 22 & 27 & -13 \end{bmatrix}$$

La primera fila de \mathbf{B}_2 es igual a:

$$\begin{aligned}
& 1 \times \text{la primera fila de } \mathbf{A} \\
+ & 0 \times \text{la segunda fila de } \mathbf{A} \\
+ & 0 \times \text{la tercera fila de } \mathbf{A}
\end{aligned}$$

La segunda fila de \mathbf{B}_2 es igual a:

$$\begin{aligned}
& 0 \times \text{la primera fila de } \mathbf{A} \\
+ & 1 \times \text{la segunda fila de } \mathbf{A} \\
+ & 0 \times \text{la tercera fila de } \mathbf{A}
\end{aligned}$$

La tercera fila de \mathbf{B}_2 es igual a:

$$\begin{aligned}
& 5 \times \text{la primera fila de } \mathbf{A} \\
+ & 0 \times \text{la segunda fila de } \mathbf{A} \\
+ & 1 \times \text{la tercera fila de } \mathbf{A}
\end{aligned}$$

Los escalares son, por lo tanto:

- para la primera fila de \mathbf{B}_2 : 1, 0 y 0
- para la segunda fila de \mathbf{B}_2 : 0, 1 y 0
- para la tercera fila de \mathbf{B}_2 : 5, 0 y 1

$$\mathbf{L}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

El operador elemental es \mathbf{L}_2 :

$$\mathbf{L}_2 \mathbf{A} = \mathbf{B}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 22 & 27 & -13 \end{bmatrix}$$

Para llevar a cabo en \mathbf{A} una operacion elemental con columnas, hay que post-multiplicar \mathbf{A} por una matriz cuadrada no singular \mathbf{M} . Esta matriz \mathbf{M} es la traspuesta de la matriz de los escalares por los que hay que multiplicar las columnas de \mathbf{A} para transformarla en \mathbf{B} .

Por ejemplo, transformemos \mathbf{A} en \mathbf{B}_3 , de modo que las columnas primera y segunda de \mathbf{B}_3 sean respectivamente iguales a las columnas segunda y primera de \mathbf{A} (intercambio de columnas) y la tercera columna de \mathbf{B}_3 sea igual a la suma de las columnas segunda y tercera de \mathbf{A} .

Se trata de dos operaciones elementales: suma de dos columnas e intercambio otras dos. La primera operación se puede llevar a cabo mediante postmultiplicación por la matriz \mathbf{M}_1 , transformando después el producto en otra matriz equivalente postmultiplicando por la matriz \mathbf{M}_2 :

$$\mathbf{B}_3 = (\mathbf{A} \times \mathbf{M}_1) \times \mathbf{M}_2$$

aunque también se puede hacer en un solo paso:

$$\mathbf{B}_3 = \mathbf{A} \times (\mathbf{M}_1 \times \mathbf{M}_2) = \mathbf{A} \times \mathbf{M}_3$$

Aquí se mostrará únicamente la matriz \mathbf{M}_3 . Como ejercicio se recomienda hallar las matrices \mathbf{M}_1 y \mathbf{M}_2 y comprobar que su producto es la matriz \mathbf{M}_3 .

La primera columna de \mathbf{B}_3 es igual a:

$$\begin{aligned} & 0 \times \text{la primera columna de } \mathbf{A} \\ + & 1 \times \text{la segunda columna de } \mathbf{A} \\ + & 0 \times \text{la tercera columna de } \mathbf{A} \end{aligned}$$

La segunda columna de \mathbf{B}_3 es igual a:

$$\begin{aligned} & 1 \times \text{la primera columna de } \mathbf{A} \\ + & 0 \times \text{la segunda columna de } \mathbf{A} \\ + & 0 \times \text{la tercera columna de } \mathbf{A} \end{aligned}$$

La tercera columna de \mathbf{B}_3 es igual a:

$$\begin{aligned} & 0 \times \text{la primera columna de } \mathbf{A} \\ + & 1 \times \text{la primera columna de } \mathbf{A} \\ + & 0 \times \text{la primera columna de } \mathbf{A} \end{aligned}$$

Los escalares necesarios son por lo tanto:

- para la primera columna de \mathbf{B}_3 : 0, 1 y 0
- para la segunda columna de \mathbf{B}_3 : 1, 0 y 0
- para la tercera columna de \mathbf{B}_3 : 0, 1 y 1

$$\mathbf{M}_3' = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Y la transformación: $\mathbf{A} \times \mathbf{M}_3 = \mathbf{B}$

$$\begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 4 & 2 \\ -1 & 1 & 1 \\ 2 & 2 & 4 \end{bmatrix}$$

4.3.2. Forma canónica de una matriz cuadrada. Cálculo del rango de una matriz.

Cualquier matriz cuadrada puede transformarse en una matriz diagonal equivalente por medio de operadores elementales.

Veamos el modo de diagonalizar la matriz de nuestro ejemplo:

- para ello transformamos primero \mathbf{A} en una matriz triangular, por medio de operaciones con filas: $\mathbf{L} \times \mathbf{A} = \mathbf{T}$ (triangular).
- a continuación transforma \mathbf{T} en una matriz diagonal, mediante operaciones con columnas: $\mathbf{T} \times \mathbf{M} = \mathbf{D}$ (diagonal)

Operaciones con filas: el objetivo es conseguir que haya ceros en los lugares (2,1), (3,1) y (3,2). Una de las muchas posibilidades es, por ejemplo, llevar a cabo los dos siguientes pasos:

(1): $\mathbf{L}_1 \times \mathbf{A} = \mathbf{B}$

- Primera fila de \mathbf{B}_1 = primera fila de \mathbf{A}
- Segunda fila de \mathbf{B}_1 = segunda fila de \mathbf{A} , menos .25 por la primera fila de \mathbf{A}
- Tercera fila de \mathbf{B}_1 = tercera fila de \mathbf{A} , menos .5 por la primera fila de \mathbf{A}

$$\begin{bmatrix} 0 & 1 & 0 \\ -1/4 & 0 & 1 \\ -1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & -3 \\ 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 5 & -3 \\ 0 & -9/4 & 11/4 \\ 0 & -1/2 & 7/2 \end{bmatrix}$$

(2): $\mathbf{L}_2 \times \mathbf{B}_1 = \mathbf{B}_2$

- Primera fila de \mathbf{B}_2 = primera fila de \mathbf{B}_1
- Segunda fila de \mathbf{B}_2 = segunda fila de \mathbf{B}_1
- Tercera fila de \mathbf{B}_2 = tercera fila de \mathbf{B}_1 , menos 2/9 de la primera fila de \mathbf{B}_1

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2/9 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & -3 \\ 1 & -9/4 & 11/4 \\ 0 & -1/2 & 7/2 \end{bmatrix} = \begin{bmatrix} 4 & 5 & -3 \\ 0 & -9/4 & 11/4 \\ 0 & 0 & 26/9 \end{bmatrix}$$

Operaciones con columnas: el objetivo ahora es conseguir ceros en las posiciones (1,2), (1,3) y (2,3). Una posibilidad es la siguiente:

(3): $\mathbf{B}_2 \times \mathbf{M}_1 = \mathbf{B}_3$

- Primera columna de \mathbf{B}_3 = primera columna de \mathbf{B}_2
- Segunda columna de \mathbf{B}_3 = segunda columna, menos 5/4 de la primera columna de \mathbf{B}_2
- Tercera columna de \mathbf{B}_3 = tercera columna más 3/4 de la primera columna de \mathbf{B}_2

$$\begin{bmatrix} 4 & 5 & -3 \\ 0 & -9/4 & 11/4 \\ 0 & 0 & 26/9 \end{bmatrix} \begin{bmatrix} 1 & 5/4 & 3/4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -9/4 & 11/4 \\ 0 & 0 & 26/9 \end{bmatrix}$$

(4): $\mathbf{B}_3 \times \mathbf{M}_2 = \mathbf{B}$

- Primera columna de \mathbf{B} = primera columna de \mathbf{B}_3
- Segunda columna de \mathbf{B} = segunda columna de \mathbf{B}_3
- Tercera columna de \mathbf{B} = tercera columna más 11/9 de la segunda columna de \mathbf{B}_3

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & -9/4 & 11/4 \\ 0 & 0 & 26/9 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 11/9 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -9/4 & 0 \\ 0 & 0 & 26/9 \end{bmatrix}$$

Por medio de operadores elementales se ha llevado a cabo una transformación lineal de la matriz \mathbf{A} en una matriz diagonal equivalente:

$$\mathbf{L}_2 \times \mathbf{L}_1 \times \mathbf{A} \times \mathbf{M}_1 \times \mathbf{M}_2 = \mathbf{L} \times \mathbf{A} \times \mathbf{M} = \mathbf{D} \text{ (diagonal)}$$

No hay una única matriz diagonal equivalente de \mathbf{A}

El determinante de una matriz \mathbf{D} diagonal y equivalente de \mathbf{A} es igual al determinante de \mathbf{A} , o al determinante de \mathbf{A} multiplicado por un escalar, ya que la transformación de \mathbf{A} en \mathbf{B} se hace mediante operadores elementales (ver 3.3.1). En nuestro ejemplo,

$$|\mathbf{D}| = 4 \times (-9/4) \times (26/9) = -26 = |\mathbf{A}|$$

El determinante de una matriz diagonal es igual al producto de los elementos de la diagonal. Si \mathbf{A} tiene rango completo, no puede haber ningún cero en la diagonal de \mathbf{D} , porque, si no, el determinante de \mathbf{D} sería cero y esto no puede ser, ya que hemos dicho que $|\mathbf{D}| = k |\mathbf{A}|$ donde ni k ni $|\mathbf{A}|$ son cero. En cambio, si \mathbf{A} no tiene rango completo, tiene que haber necesariamente por lo menos un cero sobre la diagonal. El número de ceros que se encuentran en la diagonal de \mathbf{D} es igual al número de filas dependientes que se encuentran en \mathbf{A} . El número de elementos no nulos en la diagonal de \mathbf{D} es igual al rango de \mathbf{A} .

Una matriz equivalente especial es la FORMA CANONICA de \mathbf{A} . Esta matriz esta compuesta únicamente por ceros y unos y tiene la siguiente estructura:

$$\begin{bmatrix} \mathbf{I}_r & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix}$$

donde \mathbf{I}_r es una matriz unidad de dimensión r , igual al rango de \mathbf{A} , y el resto de submatrices son matrices nulas.

Por ejemplo, la forma canónica de una matriz de dimensión 4×4 y rango 2 es:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

4.3.3. Eigenvalues and eigenvectors

Cuando existe una matriz \mathbf{U} , tal que $\mathbf{U}^{-1} \times \mathbf{A} \times \mathbf{U} = \mathbf{D}$ donde \mathbf{D} es una matriz diagonal, se denomina “eigenvalues” a los valores de la diagonal de \mathbf{D} y “eigenvectors” a las filas de \mathbf{U} .

La mayor parte de los sistemas de ecuaciones que hay que resolver en mejora genética son sistemas cuyas matrices de coeficientes son simétricas. En estos casos, el operador por columnas \mathbf{U} es igual a la traspuesta del operador por filas \mathbf{U}^{-1} : $\mathbf{U}' = \mathbf{U}^{-1}$. Se dice entonces que esta matriz es ORTOGONAL. Las matrices ortogonales son muy valiosas para simplificar la solución de sistemas de ecuaciones.

4.3.4. Operaciones elementales en sistemas de ecuaciones

Sea \mathbf{A} la matriz de coeficientes de un sistema de ecuaciones son rango completo: $\mathbf{A}\mathbf{b} = \mathbf{y}$

En nuestro ejemplo, \mathbf{A} es la matriz de coeficientes del siguiente sistema de ecuaciones:

(1)

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$2 \times x_1 + 2 \times x_2 + 2 \times x_3 = 12$$

Si cambiamos el orden de las ecuaciones:

(2)

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$2 \times x_1 + 2 \times x_2 + 2 \times x_3 = 12$$

la solución del nuevo sistema (2) es idéntico a la del sistema (1) e idéntica a la de los sistemas (3) y (4) que se escriben a continuación:

(3)

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$2 \times x_1 - 2 \times x_2 + 4 \times x_3 = 10$$

$$2 \times x_1 + 2 \times x_2 + 2 \times x_3 = 12$$

cuya segunda ecuación es la segunda de (1), pero multiplicados todos sus términos por dos.

(4)

$$4 \times x_1 + 5 \times x_2 - 3 \times x_3 = 5$$

$$x_1 - x_2 + 2 \times x_3 = 5$$

$$4 \times x_1 + 8 \times x_3 = 12$$

cuya tercera ecuación es igual a la tercera ecuación de (1), sumándole dos veces la segunda.

(1), (2), (3) y (4) son diferentes formas de un mismo sistema de ecuaciones. La solución para x_1 , x_2 y x_3 se puede calcular indistintamente a partir de cualquiera de ellos.

En notación matricial, la transformación de (1) en (2), (3) ó (4) se escribe de la siguiente forma:

$$\mathbf{LAb} = \mathbf{Ly}$$

donde \mathbf{L} es un operador elemental, o un producto de operadores elementales ($\mathbf{L} = \mathbf{L}_1 \times \mathbf{L}_2 \times \mathbf{L}_3 \times \dots \times \mathbf{L}_n$).

Resolviendo el sistema de ecuaciones (ver [Inversa del producto de dos matrices cuadradas](#)):

$$\begin{aligned}\mathbf{b} &= (\mathbf{LA})^{-1}\mathbf{Ly} \\ &= \mathbf{A}^{-1}\mathbf{L}^{-1}\mathbf{Ly}\end{aligned}$$

$$= \mathbf{A}^{-1}\mathbf{Iy} = \mathbf{A}^{-1}\mathbf{y}$$

Es decir, la solución para el sistema transformado es idéntica a la solución para el sistema original, con la única condición de que \mathbf{L} tenga inversa.

Será interesante llevar a cabo una transformación de este tipo, cuando la matriz equivalente (\mathbf{LA}) sea más fácil de invertir que la matriz \mathbf{A} .

Las matrices más fáciles de invertir son las matrices diagonales. Si se transforma la matrix \mathbf{A} por medio de operadores elementales \mathbf{L} y \mathbf{M} en una matriz diagonal: $\mathbf{LAM} = \mathbf{D}$ (diagonal), teniendo \mathbf{L} y \mathbf{M} inversa, el sistema de ecuaciones queda de la siguiente forma:

$$\mathbf{Ab} = \mathbf{y}$$

$$\mathbf{A Ib} = \mathbf{y}$$

$$\begin{aligned}\mathbf{A}\mathbf{M}\mathbf{M}^{-1}\mathbf{b} &= \mathbf{y} \\ \mathbf{L}\mathbf{A}\mathbf{M}\mathbf{M}^{-1}\mathbf{b} &= \mathbf{L}\mathbf{y} \\ \mathbf{D}\mathbf{M}^{-1}\mathbf{b} &= \mathbf{L}\mathbf{y}\end{aligned}$$

y su solución es:

$$\begin{aligned}\mathbf{b} &= (\mathbf{D}\mathbf{M}^{-1})^{-1}\mathbf{L}\mathbf{y} \\ &= (\mathbf{M}^{-1})^{-1}\mathbf{D}^{-1}\mathbf{L}\mathbf{y} \\ &= \mathbf{M}\mathbf{D}^{-1}\mathbf{L}\mathbf{y}\end{aligned}$$

donde \mathbf{D} es una matriz diagonal y por lo tanto muy fácil de invertir.

Esta solución

$$\mathbf{b} = \mathbf{M}\mathbf{D}^{-1}\mathbf{L}\mathbf{y} = \mathbf{M}(\mathbf{L}\mathbf{A}\mathbf{M})^{-1}\mathbf{L}\mathbf{y} = \mathbf{M}(\mathbf{M}^{-1}\mathbf{A}^{-1}\mathbf{M}^{-1})\mathbf{L}\mathbf{y} = \mathbf{A}^{-1}\mathbf{y}$$

es idéntica a la solución del sistema de ecuaciones original, pero se ha calculado de una forma mucho más fácil.

4.4. Solución de un sistema de ecuaciones con rango incompleto

4.4.1. Inversa generalizada de una matriz

Muy a menudo nos encontramos con sistemas de ecuaciones con rango incompleto:

$$\mathbf{A}\mathbf{b} = \mathbf{y}$$

donde

$$|\mathbf{A}| = 0$$

En este caso, \mathbf{A} no tiene inversa. Ya no es posible resolver el sistema.

Un sistema como éste tiene un número infinito de soluciones, que se pueden encontrar por medio de “inversas generalizadas”. Una vez hallada una solución, no se puede olvidar que NO ES LA ÚNICA SOLUCIÓN y hay que considerar qué sentido tiene esa solución encontrada.

Una inversa generalizada, o g-inversa de la matriz \mathbf{A} es otra matriz \mathbf{G} , tal que:

$$\mathbf{A}\mathbf{G}\mathbf{A} = \mathbf{A}$$

A menudo se representa una g-inversa de \mathbf{A} como \mathbf{A}^- .

Si \mathbf{A} tiene el rango completo, su inversa es también una g-inversa de \mathbf{A} , ya que cumple que:

$$\mathbf{A}\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}\mathbf{A} = \mathbf{A}$$

Una matriz puede tener no solamente una, sino muchas (un número infinito) de g-inversas.

Observación: matrices no cuadradas también tienen g-inversas.

4.4.2. Cálculo de g-inversas de una matriz

Sea la matriz cuadrada \mathbf{A} , de dimensión n y rango r , donde $r < n$. Vamos a hallar una g-inversa de \mathbf{A} , por dos métodos, uno general y otro simplificado, siendo este último el que realmente se utiliza en la práctica ⁵.

4.4.2.1. Método general: Como se ha visto anteriormente, es posible transformar la matriz \mathbf{A} , por medio de matrices no singulares \mathbf{P} y \mathbf{Q} , en otra que tenga $(n - r)$ filas y columnas ocupadas únicamente por ceros:

$$\mathbf{B} = \mathbf{P}\mathbf{A}\mathbf{Q} = \begin{bmatrix} \mathbf{C}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix}$$

donde $\mathbf{C}_{r \times r}$ es una submatriz de dimensión $r \times r$ y rango completo y el resto de las submatrices son matrices nulas.

Como la matriz \mathbf{C} tiene rango completo, tiene inversa \mathbf{C}^{-1} . Entonces la matriz:

$$\mathbf{G} = \mathbf{Q} \begin{bmatrix} \mathbf{C}_{r \times r}^{-1} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \mathbf{P}$$

es una g-inversa de \mathbf{A} : $\mathbf{G} = \mathbf{A}^{-1}$

En efecto, sea $\mathbf{F} = \mathbf{A}\mathbf{G}\mathbf{A}$, pre y postmultiplicando \mathbf{F} por \mathbf{P} y \mathbf{Q} respectivamente, se obtiene la siguiente expresión (se explicitan las dimensiones de las submatrices, para comprobar que los pasos son correctos):

⁵Mencionaremos dos métodos más muy frecuentes: la [g-inversa de Moore-Penrose](#), y un método muy usado en mejora genética animal: la g-inversa obtenida mediante nulificación de pivotes de valor 0 durante la descomposición de Cholesky.

$$\begin{aligned}
\mathbf{PFQ} &= \mathbf{PAGAQ} = \mathbf{PAQ} \begin{bmatrix} \mathbf{C}_{r \times r}^{-1} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \mathbf{PAQ} \\
&= \begin{bmatrix} \mathbf{C}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{r \times r}^{-1} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \mathbf{PAQ} \\
&= \begin{bmatrix} \mathbf{I}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{r \times r}^{-1} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \mathbf{PAQ} \\
&= \begin{bmatrix} \mathbf{C}_{r \times r}^{-1} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} = \mathbf{PAQ}
\end{aligned}$$

En resumen: $\mathbf{PFQ} = \mathbf{PAQ}$

\mathbf{P} y \mathbf{Q} son no singulares y por lo tanto tienen inversas \mathbf{P}^{-1} y \mathbf{Q}^{-1} .

Pre y postmultiplicando ambos miembros de la expresión anterior por \mathbf{P}^{-1} y \mathbf{Q}^{-1} respectivamente, se obtiene:

$$\mathbf{P}^{-1}\mathbf{PFQ}\mathbf{Q}^{-1} = \mathbf{P}^{-1}\mathbf{PAQ}\mathbf{Q}^{-1}$$

$$\mathbf{IFI} = \mathbf{IAI}$$

$$\mathbf{F} = \mathbf{A}$$

y como $\mathbf{F} = \mathbf{AGA}$

$$\mathbf{AGA} = \mathbf{A}$$

y \mathbf{G} es una g-inversa de \mathbf{A} .

Ejemplo: en el siguiente sistema de ecuaciones existe una dependencia entre las ecuaciones:

$$x_1 + 2 \times x_2 + 3 \times x_3 = 30$$

$$2 \times x_1 + 2 \times x_2 + x_3 = 26$$

$$4 \times x_1 + 6 \times x_2 + 7 \times x_3 = 86$$

La tercera ecuación es igual a la segunda más dos veces la primera. En notación matricial:

$$\mathbf{Ab} = \mathbf{y} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 4 & 6 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 30 \\ 26 \\ 86 \end{bmatrix}$$

El determinante de \mathbf{A} es cero:

$$|\mathbf{A}| = 1 \times 2 \times 7 + 2 \times 1 \times 4 + 3 \times 2 \times 6 - 3 \times 2 \times 4 - 2 \times 2 \times 7 - 1 \times 1 \times 6 = 0$$

Dos de las muchas inversas de \mathbf{A} serán \mathbf{G}_1 y \mathbf{G}_2 .

a) las matrices \mathbf{P}_1 y \mathbf{Q}_1 son no singulares: $|\mathbf{P}_1| = |\mathbf{Q}_1| = 1$

$$\mathbf{P}_1 = \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

;

$$\mathbf{Q}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -5/4 & 1 & 0 \\ 1/2 & -2 & 1 \end{bmatrix}$$

el producto $\mathbf{P}_1 \mathbf{A} \mathbf{Q}_1$ tiene una fila y una columna compuestas exclusivamente por ceros:

$$\mathbf{P}_1 \mathbf{A} \mathbf{Q}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -8 & 7 \end{bmatrix}$$

Eliminando la fila y la columna nulas, se obtiene la submatriz \mathbf{C}_1 , con rango completo e inversa:

$$\mathbf{C}_1 = \begin{bmatrix} 0 & 1 \\ -8 & 7 \end{bmatrix}; \mathbf{C}_1^{-1} = \begin{bmatrix} 7/8 & -1/8 \\ 1 & 0 \end{bmatrix}$$

Sustituyendo en $\mathbf{P}_1 \mathbf{A} \mathbf{Q}_1$ la submatriz \mathbf{C}_1 por su inversa \mathbf{C}_1^{-1} , se obtiene una nueva matriz \mathbf{C}_1^- :

$$\mathbf{C}_1^- = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 7/8 & -1/8 \\ 0 & 1 & 0 \end{bmatrix}$$

Entonces una g-inversa de \mathbf{A} es la matriz \mathbf{G}_1 :

$$\mathbf{G}_1 = \mathbf{Q}_1 \mathbf{C}_1^- \mathbf{P}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -5/4 & 1 & 0 \\ 1/2 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 7/8 & -1/8 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{G}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 7/8 & -1/8 \\ 0 & -3/4 & 1/4 \end{bmatrix}$$

Es fácil comprobar que \mathbf{G}_1 es una inversa generalizada de \mathbf{A} , llevando a cabo el producto de $\mathbf{A} \mathbf{G}_1 \mathbf{A}$, que tiene que ser igual a \mathbf{A} .

b) las matrices \mathbf{P}_2 y \mathbf{Q}_2 son no singulares: $|\mathbf{P}_2| = |\mathbf{Q}_2| = 1$

$$\mathbf{P}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & -1 \\ -4 & 0 & 1 \end{bmatrix}; \mathbf{Q}_2 = \begin{bmatrix} 1 & -4/5 & -3 \\ 0 & 1 & 0 \\ 0 & -2/5 & 1 \end{bmatrix}$$

siguiendo el mismo proceso anterior:

$$\mathbf{P}_2 \mathbf{A} \mathbf{Q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

Eliminando la segunda fila y la segunda columna:

$$\mathbf{C}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -5 \end{bmatrix}; \mathbf{C}_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1/5 \end{bmatrix}; \mathbf{C}_2^- = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1/5 \end{bmatrix}$$

Una g-inversa de \mathbf{A} es entonces \mathbf{G}_2 :

$$\mathbf{G}_2 = \mathbf{Q}_2 \mathbf{C}_2^- \mathbf{P}_2 = \begin{bmatrix} -7/5 & 0 & 3/5 \\ 0 & 0 & 0 \\ 4/5 & 0 & -1/5 \end{bmatrix}$$

Se puede comprobar fácilmente que $\mathbf{A} \mathbf{G}_2 \mathbf{A} = \mathbf{A}$

Del mismo modo se pueden encontrar otras g-inversas de \mathbf{A} , por ejemplo por medio de las siguientes transformaciones:

$$\mathbf{P}_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & -1 \end{bmatrix}; \mathbf{Q}_3 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -5/2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{P}_4 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & -1 & 1 \end{bmatrix}; \mathbf{Q}_4 = \begin{bmatrix} 1 & -2 & 2 \\ 0 & 1 & -5/2 \\ 0 & 0 & 1 \end{bmatrix}$$

etc.

4.4.2.2. Método simplificado Hay algunas g-inversas de \mathbf{A} que son muy fáciles de hallar. Para ello debemos conocer primeramente el rango de \mathbf{A} . Si su dimensión es n y su rango es r , se eliminan $(n - r)$ filas y columnas de \mathbf{A} , se invierte la matriz resultante y se remplazan en la inversa las filas y columnas eliminadas por ceros. El resultado es una g-inversa de \mathbf{A} .

En el ejemplo anterior existía una dependencia entre ecuaciones y el rango de la matriz \mathbf{A} era 2, mientras que su dimensión era 3. Sustituimos por tanto $(3 - 2)$ filas y columnas (o sea una fila y una columna cualesquiera) por ceros y obtenemos, por ejemplo la matriz \mathbf{B} :

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_3 & 0 \\ 0 & 0 \end{bmatrix} \text{ donde } \mathbf{C}_3 = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$$

$$\text{Sea entonces } \mathbf{G}_3 = \begin{bmatrix} \mathbf{C}_3^{-1} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

La matriz \mathbf{G}_3 también es una g-inversa de \mathbf{A} , como se puede comprobar llevando a cabo el producto de $\mathbf{AG}_3\mathbf{A}$.

4.4.3. Solución de sistemas de ecuaciones con rango incompleto por medio de g-inversas. Funciones estimables.

Para poder resolver sistemas de ecuaciones con rango incompleto se exige que las ecuaciones sean CONSISTENTES, es decir, que se den las mismas dependencias a ambos lados del signo de igual.

En nuestro ejemplo, tanto en la matriz \mathbf{A} como en el vector \mathbf{y} , la tercera fila es igual a la segunda fila más dos veces la primera ($86 = 2 \times 30 + 26$). Las ecuaciones del sistema son consistentes.

Si \mathbf{G} es una g-inversa de \mathbf{A} , se puede probar la consistencia de las ecuaciones llevando a cabo el producto de \mathbf{AGy} . Si el producto de $\mathbf{AGy} = \mathbf{y}$, las ecuaciones son consistentes. En caso contrario no lo son.

Una vez probada la consistencia, se puede calcular una solución para el sistema:

$$\mathbf{b} = \mathbf{Gy}$$

En el ejemplo:

$$\mathbf{b}_1 = \mathbf{G}_1\mathbf{y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 7/8 & -1/8 \\ 0 & -3/4 & 1/4 \end{bmatrix} \begin{bmatrix} 30 \\ 26 \\ 86 \end{bmatrix} = \begin{bmatrix} 0 \\ 12 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\mathbf{b}_2 = \mathbf{G}_2 \mathbf{y} = \begin{bmatrix} -7/5 & 0 & 3/5 \\ 0 & 0 & 0 \\ 4/5 & 0 & -1/5 \end{bmatrix} \begin{bmatrix} 30 \\ 26 \\ 86 \end{bmatrix} = \begin{bmatrix} 48/5 \\ 0 \\ 34/5 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\mathbf{b}_3 = \mathbf{G}_3 \mathbf{y} = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 30 \\ 26 \\ 86 \end{bmatrix} = \begin{bmatrix} -4 \\ 17 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Cada g-inversa ha dado lugar a una solución diferente. La pregunta ahora sería: ¿Cual de las tres es “la buena”?; ¿ x_1 es igual a 0, 48/5 o -4?. No podemos aceptar ninguno de los tres valores, porque la solución encontrada depende de que g-inversa hayamos usado y hemos podido usar \mathbf{G}_1 , \mathbf{G}_2 , \mathbf{G}_3 , o cualquier otra. Sin embargo hay algunas relaciones entre valores o combinaciones de valores que sí son constantes e independientes de qué g-inversa se haya utilizado. Así podemos comprobar en nuestro ejemplo, que la función $x_1 - 2x_3$, tiene el mismo valor sea cual sea la solución encontrada para \mathbf{b} :

- $\mathbf{b}_1 : x_1 - 2x_3 = 0 - 2 \times 2 = -4$
- $\mathbf{b}_2 : x_1 - 2x_3 = 48/5 - 2 \times 34/5 = -4$
- $\mathbf{b}_3 : x_1 - 2x_3 = -4 - 2 \times 0 = -4$

Del mismo modo, $x_2 + 5/2x_3$ da siempre 17:

- $\mathbf{b}_1 : x_2 + 5/2x_3 = 0 + 5/2 \times 2 = 17$
- $\mathbf{b}_2 : x_2 + 5/2x_3 = 0 + 5/2 \times 34/5 = 17$
- $\mathbf{b}_3 : x_2 + 5/2x_3 = 17 + 5/2 \times 0 = 17$

Estas combinaciones (FUNCIONES) de x_1 , x_2 y x_3 toman el mismo valor con las tres soluciones y también con cualquier otra de las muchas soluciones posibles. Cuales son estas funciones depende únicamente de la estructura del sistema de ecuaciones y no del modo de resolverlo y se denominan FUNCIONES ESTIMABLES.

La tarea de quien se encuentra ante la resolución de un sistema de ecuaciones con rango incompleto es encontrar FUNCIONES ESTIMABLES QUE TENGAN UN SENTIDO.

Con la valoración genética de los animales se pretende encontrar aquellos individuos que son los “mejores” dentro de una población. Los animales se clasifican de acuerdo a valores relativos, de modo que el mejor obtiene el valor más alto. Desde este punto de vista, el conocimiento del valor absoluto de los animales no es ineludiblemente necesario. Siempre que las diferencias entre los individuos sean funciones estimables, se podrán utilizar sistemas de ecuaciones con rango incompleto para su evaluación⁶.

⁶En la mayoría de programas utilizados en evaluación, la g-inversa utilizada no es explícita para el utilizador. Sin embargo, las funciones de interés (diferencia entre valores de animales, o, por ejemplo, entre niveles de un efecto fijo) son casi siempre estimables.

Las funciones se representan por medio de un vector \mathbf{k}' . Las funciones estimables que hemos visto en el ejemplo serán:

$$x_1 - 2x_3 = (1 \quad 0 \quad -2) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}; \mathbf{k}'_1 = (1 \quad 0 \quad -2)$$

$$x_2 + \frac{5}{2}x_3 = (1 \quad 0 \quad 5/2) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}; \mathbf{k}'_2 = (105/2)$$

Para comprobar si una función es estimable o no, se lleva a cabo el producto $\mathbf{k}'\mathbf{GA} = \mathbf{k}'$, la función es estimable y en caso contrario no lo es.

Como ejercicio se puede comprobar que las funciones \mathbf{k}'_1 y \mathbf{k}'_2 son estimables, y que en cambio no lo son ni la función (100), ni la función (010). Esto último quiere decir que las incógnitas no son estimables.

4.5. Absorción de ecuaciones

Normalmente el mayor problema que se encuentra es la inversión de la matriz. La dificultad y tiempo necesario para el cálculo de determinantes e inversas depende del tamaño y estructura de la matriz a invertir.

Con la dimensión de la matriz, aumentan exponencialmente.

Algunas matrices, como por ejemplo las matrices diagonales, son fáciles de invertir aún en el caso de que su dimensión sea muy grande.

Un procedimiento para facilitar la inversión de algunas matrices de gran dimensión, es la absorción de ecuaciones.

Sea el siguiente sistema de ecuaciones a resolver: $\mathbf{Ab} = \mathbf{y}$, y sea su matriz de coeficientes \mathbf{A} tal que tenga una gran dimensión, por ejemplo un millón de filas y columnas, pero una estructura especial, que se puede explicar mostrando \mathbf{A} en forma de matriz particionada:

$$\mathbf{A} = \begin{bmatrix} \mathbf{B}_{m \times m} & \mathbf{C}_{m \times n} \\ \mathbf{D}_{n \times m} & \mathbf{E}_{n \times n} \end{bmatrix}$$

donde:

- la submatriz \mathbf{E} es una matriz de dimensión grande, pero muy fácil de invertir, como por ejemplo una matriz diagonal.
- La submatriz \mathbf{B} tiene una dimensión lo suficientemente pequeña, como para que su inversión no exija más memoria de ordenador de la disponible.

entonces el sistema de ecuaciones se puede escribir de la siguiente forma:

$$\begin{bmatrix} \mathbf{B}_{m \times m} & \mathbf{C}_{m \times n} \\ \mathbf{D}_{n \times m} & \mathbf{E}_{n \times n} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{m \times 1} \\ \mathbf{d}_{n \times 1} \end{bmatrix} = \begin{bmatrix} \mathbf{w}_{m \times 1} \\ \mathbf{z}_{n \times 1} \end{bmatrix}$$

donde

$$\begin{bmatrix} \mathbf{c}_{m \times 1} \\ \mathbf{d}_{n \times 1} \end{bmatrix} = \mathbf{b}; \begin{bmatrix} \mathbf{w}_{m \times 1} \\ \mathbf{z}_{n \times 1} \end{bmatrix} = \mathbf{y}$$

Multiplicando las matrices queda:

(1)

$$\mathbf{B}_{m \times m} \mathbf{c}_{m \times 1} + \mathbf{C}_{m \times n} \mathbf{d}_{n \times 1} = \mathbf{w}_{m \times 1}$$

(2)

$$\mathbf{D}_{n \times m} \mathbf{c}_{m \times 1} + \mathbf{E}_{n \times n} \mathbf{d}_{n \times 1} = \mathbf{z}_{n \times 1}$$

Se muestran las dimensiones de las submatrices, para comprobar que coinciden.

Si suponemos que \mathbf{E} es fácil de invertir, se puede despejar \mathbf{d} en la ecuación (2):

$$\mathbf{E} \mathbf{d} = \mathbf{z} - \mathbf{D} \mathbf{c}$$

$$\mathbf{d} = \mathbf{E}^{-1}(\mathbf{z} - \mathbf{D} \mathbf{c})$$

y sustituirlo en la ecuación (1):

$$\mathbf{B} \mathbf{c} + \mathbf{C} \mathbf{E}^{-1}(\mathbf{z} - \mathbf{D} \mathbf{c}) = \mathbf{w}$$

$$(\mathbf{B} - \mathbf{C} \mathbf{E}^{-1} \mathbf{D}) \mathbf{c} = \mathbf{w} - \mathbf{C} \mathbf{E}^{-1} \mathbf{z}$$

y la solución para \mathbf{C} es:

$$\mathbf{c} = (\mathbf{B} - \mathbf{C} \mathbf{E}^{-1} \mathbf{D})^{-1}(\mathbf{w} - \mathbf{c} \mathbf{E}^{-1} \mathbf{z})$$

donde la matriz $(\mathbf{B} - \mathbf{C} \mathbf{E}^{-1} \mathbf{D})$ tiene dimensión que hace posible llevar a cabo su inversión.

A partir de \mathbf{c} se puede calcular también $\mathbf{d} = \mathbf{E}^{-1}(\mathbf{z} - \mathbf{D} \mathbf{c})$.

Este procedimiento suele exigir bastante tiempo de cálculo, pero hace posible la inversión de la matriz.

5. Modelos lineales en Producción Animal

5.1. Definición de un modelo lineal

Los análisis estadísticos se llevan a cabo para explicar qué causas han dado lugar al valor de una determinada medida (normalmente en nuestro caso, un rendimiento animal): qué factores tienen una influencia sobre la magnitud de la medida y en cuánto se puede cifrar esa influencia.

Esas causas y factores de influencia son:

- el mismo animal. La constitución genética de un animal está condicionada por las de sus antecesores, pero definitivamente determinada por el azar.
- el ambiente en el que se encuentra el animal.

Si conociéramos estos factores de influencia y sus efectos, podríamos predecir el futuro rendimiento de un animal, o dirigir en una dirección determinada el nivel de rendimiento de una población.

Para llevar a cabo un análisis de este tipo se necesita:

1. una población, o una muestra de la población, que nos proporcione los datos.
2. una idea preconcebida de cuáles son los factores que pueden tener una influencia significativa sobre los valores a analizar. Esa idea previa da lugar al “modelo”.

Por medio del modelo se intenta describir, cómo se ha llegado a realizar el rendimiento de un animal. Este rendimiento se divide en efectos. Cuando estos efectos simplemente se suman, se ha construido un “modelo lineal”:

$$y = A + B + C + \dots$$

Una explicación de este tipo, sólo puede ser una aproximación. El número de factores de influencia es prácticamente infinito y no se pueden considerar todos en un modelo.

Se quedan sin considerar:

1. factores de influencia cuyo efecto es muy pequeño.
2. factores de influencia para los que es desconocida la relación entre el animal medido y el factor.

Para adecuar el modelo a las posibilidades de cálculo, es necesario hacer una serie de asunciones, que normalmente no se corresponden con la realidad. Estas asunciones pertenecen al modelo y hay que explicitarlas juntamente con la fórmula matemática del mismo. Los efectos del modelo son cantidades que pueden explicar una parte del rendimiento, pero queda otra parte que “no

sabemos explicar” (la suma de los efectos de los factores no considerados). Esta parte no explicada se llama “residuo” (e).

La fórmula del modelo completo es:

$$y = A + B + C + \dots + e$$

Si se estima el rendimiento del animal “ i ” - cuyo rendimiento realizado nos es conocido - por medio de la suma de los efectos de los factores de influencia del modelo que han actuado sobre ese animal, el residuo e_i para esa medida es el error que se ha cometido en la estimación. El residuo e_i para la medida de “ i ” es la diferencia entre el rendimiento realizado y el rendimiento explicado:

$$e_i = y_i - \hat{y}_i = y_i - A - B - C - \dots$$

debido a que no sabemos ABSOLUTAMENTE NADA sobre las causas que han dado lugar a e_i , debemos hacer una primera ASUNCIÓN general sobre e_i :

El residuo e_i está determinado por el azar y por lo tanto es un suceso independiente cuya esperanza matemática es cero y que pertenece a una población de sucesos independientes con una determinada distribución de frecuencias. Para algunos análisis es necesario aceptar además que dicha distribución es normal:

$$e_i \sim N(0, \sigma_e^2)$$

Dicho de otra manera: e_i tiene valor determinado, pero no sabemos por qué causas e_i ha tomado ese valor. En principio no tenemos ninguna idea de cuánto debe valer e_i y su valor podría ser cualquiera (e_i es aleatorio). Sin embargo sería muy improbable que el valor de e_i sobrepasara ciertos límites. Incluso dentro de ese rango, hay ciertos intervalos de magnitud en los que es más probable que se encuentre el valor de e_i que en otros.

Una vez definido el modelo, se pueden calcular valores para los efectos considerados en el mismo, si se dispone de datos registrados en una muestra de la población. Estos valores calculados, son ESTIMACIONES; si hubiéramos tomado otra muestra distinta, habríamos obtenido otros valores distintos. Se dice que una estimación tiene poca precisión, cuando con distintas muestras se obtienen valores muy distintos. En cambio, un procedimiento preciso proporciona resultados parecidos con muestras distintas y sólo muestras muy poco representativas de la población dan lugar a que los valores estimados con ellas se alejen de los demás.

Por lo tanto para llevar a cabo la estimación se necesita:

- una muestra representativa de la población

- un modelo, que se corresponda con la realidad lo más posible (o sea un modelo en el que todos los factores de influencia realmente importantes están considerados).

Vamos a utilizar un ejemplo muy simplificado para llevar a cabo el proceso de construcción de un modelo. Para ello utilizaremos un conjunto de datos extremadamente pequeño, de modo que se pueda seguir sin problemas. Esto tiene el inconveniente de que el análisis parece no tener mucho sentido, pero la intención por ahora es únicamente describir el proceso y no hacer una utilización práctica del mismo.

Ejemplo: un criador de caracoles quisiera tener una idea del peso de los animales de su explotación, pero no quiere pesarlos todos uno a uno.

1. Material:

Su población consiste en 4 grupos de caracoles de 1, 2, 3 y 4 meses de edad, respectivamente. A modo de muestra, ha pesado un caracol de cada grupo y quiere calcular parámetros de estimación a partir de estos datos.

Los datos registrados son:

Animal número	Edad (meses)	Peso (g)
1	1	3.5
2	2	4.5
3	3	6.5
4	4	9.6

2. Método:

El peso de un caracol es la suma de su peso al nacimiento y del aumento de peso desde entonces y éste último depende de la edad del animal. Por ello tomaremos como factores de influencia el peso al nacimiento y la edad.

- Efecto del peso al nacimiento: desconocemos los pesos que tuvieron los cuatro caracoles al nacer. Dado que no hay ningún criterio para clasificarlos, como pudieran ser diferencias genéticas entre ellos o distintos ambientes hasta el nacimiento, tenemos que ASUMIR que el efecto “peso al nacimiento” es igual para todos.
- Efecto de la edad: el efecto de la edad es el aumento de peso. Cuanto más viejos son los animales, más pesados son. ASUMIMOS que la relación entre peso y edad es lineal:

$$\text{Aumento de peso} = b \times \text{Edad}$$

donde b es una constante igual para todos los animales.

Bajo esas condiciones, el modelo resulta:

$$y_i = a + b \times x_i + e_i$$

donde:

y_i : peso de i-ésimo caracol

x_i : edad del i-ésimo caracol en el momento de la pesada

a : efecto “peso al nacimiento”

b : coeficiente para el efecto “aumento de peso”, que depende de la edad

e_i : residuo para el i-ésimo caracol

Ahora podemos, a partir de los datos de la muestra, estimar valores para los parámetros desconocidos a y b . Los valores estimados se representan \hat{a} y \hat{b} .

Una vez hallados \hat{a} y \hat{b} , podremos estimar el peso actual del j-ésimo caracol, que tiene una edad x_j , mediante la expresión:

$$\hat{y}_j = \hat{a} + \hat{b} \times x_j$$

Este es el valor que utilizaremos como aproximación al peso real del j-ésimo caracol, en la esperanza de que el residuo (que es igual al error que cometemos), sea poco importante.

5.2. Método de los mínimos cuadrados (LSQ)

5.2.1. Ecuaciones normales con rango completo

Antes del análisis :

- desconocemos el valor de los parámetros a y b
- conocemos la edad y el peso de los caracoles de la muestra

Si escribimos la expresión del modelo particularizada para cada uno de los caracoles, obtenemos un sistema de ecuaciones:

$$\text{Animal número 1: } 35 = a \times 1 + b \times 1 + e_1$$

$$2: 45 = a \times 1 + b \times 2 + e_2$$

$$3: 65 = a \times 1 + b \times 3 + e_3$$

$$4: 96 = a \times 1 + b \times 4 + e_4$$

En notación matricial: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$

$$\begin{pmatrix} 35 \\ 45 \\ 65 \\ 96 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}$$

donde:

- \mathbf{y} es el vector de los pesos registrados
- $\boldsymbol{\beta}$ es el vector de los parámetros a estimar
- \mathbf{e} es el vector de los residuos

La matriz \mathbf{X} se llama matriz de incidencia y describe la estructura de la muestra: qué parámetros corresponden a cada peso y en qué relación.

El sistema de ecuaciones tiene solo 4 ecuaciones (tantas como registros se obtienen de la muestra), pero 6 incógnitas: los dos parámetros a y b y los cuatro residuos e_1, e_2, e_3 y e_4 . Por lo tanto tiene un rango incompleto y un número infinito de soluciones.

Sin embargo, no todas las incógnitas tienen el mismo sentido. Los parámetros a y b son los valores que queremos encontrar para utilizarlos en la estimación de los pesos de los animales. El valor de cada uno de los residuos en cambio, no tiene una aplicación práctica. Si sabemos que estos residuos son iguales a los errores que cometemos al hacer la estimación y por lo tanto lo único que nos interesa de ellos es que tengan la menor magnitud posible.

De las muchas soluciones posibles del sistema de ecuaciones, nos interesa únicamente aquella que haga que el error en conjunto sea mínimo. A esta solución le llamamos la OPTIMA (BEST) solución.

Por medio del “Método de los Mínimos Cuadrados”, podemos encontrar esta óptima solución. Este método tiene en cuenta a la par el modelo (el sistema de ecuaciones) y la condición de que el error general sea mínimo. Como el residuo puede tomar valores positivos y negativos, la suma de todos ellos no es un buen parámetro para cuantificar el error en su conjunto. En el método de los mínimos cuadrados, se minimiza la suma de los cuadrados de todos los residuos.

En resumen, el método de los mínimos cuadrados se basa en la toma en consideración de:

1. el modelo y
2. la condición de que la suma de los cuadrados de los residuos de los registros de muestra sea mínima.

Donde:

1. Modelo $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$
2. Función a minimizar: $L = \sum_i e_i^2 = \mathbf{e}'\mathbf{e}$

La función L alcanza un mínimo cuando las primeras derivadas parciales de L respecto a los parámetros a y b se igualan a cero:

$$\begin{pmatrix} \frac{\partial L}{\partial a} \\ \frac{\partial L}{\partial b} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

La función es :

$$\begin{aligned} L = \mathbf{e}'\mathbf{e} &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = (\mathbf{y}' - \boldsymbol{\beta}'\mathbf{X}')(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned}$$

y sus derivadas parciales:

$$\frac{\partial L}{\partial a} = 0 - \mathbf{y}'\mathbf{X} \frac{\partial \boldsymbol{\beta}}{\partial a} - \frac{\partial \boldsymbol{\beta}'}{\partial a} \mathbf{X}'\mathbf{y} + \frac{\partial \boldsymbol{\beta}'}{\partial a} \mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X} \frac{\partial \boldsymbol{\beta}}{\partial a}$$

$$\frac{\partial L}{\partial b} = 0 - \mathbf{y}'\mathbf{X} \frac{\partial \boldsymbol{\beta}}{\partial b} - \frac{\partial \boldsymbol{\beta}'}{\partial b} \mathbf{X}'\mathbf{y} + \frac{\partial \boldsymbol{\beta}'}{\partial b} \mathbf{X}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X} \frac{\partial \boldsymbol{\beta}}{\partial b}$$

donde:

$$\frac{\partial \boldsymbol{\beta}'}{\partial a} = \frac{\partial (a \quad b)}{\partial a} = (1 \quad 0)$$

$$\frac{\partial \boldsymbol{\beta}'}{\partial b} = \frac{\partial (a \quad b)}{\partial b} = (0 \quad 1)$$

$$\frac{\partial \boldsymbol{\beta}'}{\partial a} \mathbf{X}'\mathbf{y} = (1 \quad 0) \mathbf{X}'\mathbf{y}$$

etc.

Nótese que:

- los sumandos de las dos ecuaciones son matrices de dimensión 1
- algunos sumandos son iguales a la matriz traspuesta de otros sumandos, como por ejemplo :

$$\mathbf{y}'\mathbf{X} \frac{\partial \boldsymbol{\beta}}{\partial a} = \left(\frac{\partial \boldsymbol{\beta}'}{\partial a} \mathbf{X}'\mathbf{y} \right)'$$

Dado que la matriz traspuesta de una matriz de dimensión 1, es la misma matriz, se pueden simplificar las expresiones anteriores.

Igualando las derivadas parciales a cero, se obtiene la siguiente expresión:

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} \frac{\partial L}{\partial a} \\ \frac{\partial L}{\partial b} \end{pmatrix} = \begin{pmatrix} -2(1 \ 0) \mathbf{X}'\mathbf{y} + 2(1 \ 0) \mathbf{X}'\mathbf{X}\hat{\beta} \\ -2(0 \ 1) \mathbf{X}'\mathbf{y} + 2(0 \ 1) \mathbf{X}'\mathbf{X}\hat{\beta} \end{pmatrix} = \\ &= -2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{X}'\mathbf{y} + 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{X}'\mathbf{X}\hat{\beta} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\beta} \end{aligned}$$

Observación: al igualar las derivadas parciales a cero, se ha cumplido la condición de que la suma de los cuadrados sea mínima. Por ello se ha sustituido el vector β que contiene a las variables a y b (las infinitas posibles soluciones del sistema), por el vector $\hat{\beta}$, que es una sola de esas soluciones, que llamamos la solución óptima. El vector $\hat{\beta}$ es el vector de los estimadores de los parámetros: $\hat{\beta}' = (\hat{a} \ \hat{b})$.

Dividiendo por dos a la izquierda y a la derecha del igual queda:

$$0 = -\mathbf{X}'\mathbf{y} + \mathbf{X}'\mathbf{X}\hat{\beta}$$

o también

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$$

Este es un nuevo sistema de ecuaciones. Estas ecuaciones se conocen como “las ecuaciones normales”.

Su solución:

$$\boxed{\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}}$$

es la mejor de todas las soluciones que podemos encontrar para ese modelo con la muestra disponible: la solución para la que la suma de cuadrados de los residuos es mínima.

Los valores \hat{a} y \hat{b} son los estimadores LSQ (least-square) de los parámetros a y b .

En nuestro ejemplo:

$$\begin{aligned}
\hat{\beta} &= \left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 35 \\ 45 \\ 65 \\ 96 \end{pmatrix} = \\
&\begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 35 \\ 45 \\ 65 \\ 96 \end{pmatrix} = \\
&\frac{1}{20} \begin{pmatrix} 30 & -10 \\ -10 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 35 \\ 45 \\ 65 \\ 96 \end{pmatrix} = \\
&\frac{1}{20} \begin{pmatrix} 190 \\ 406 \end{pmatrix} = \begin{pmatrix} 095 \\ 203 \end{pmatrix}
\end{aligned}$$

Nuestros estimadores son:

$$\hat{a} = 095$$

$$\hat{b} = 203$$

El estimador del peso del j-ésimo caracol que tiene la edad x_j es:

$$\hat{y}_j = 095 + 203 \times x_j$$

Los residuos son las diferencias entre los valores medidos por un lado y la suma de los efectos correspondientes por otro. De los cuatro animales de la muestra conocemos los verdaderos pesos, pero de los efectos sólo tenemos estimaciones. Por ello, la diferencia entre el valor realizado y el estimado, es una estimación del residuo, aunque no el residuo mismo.

Los estimadores de los residuos de nuestro ejemplo son:

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\beta}$$

$$\hat{\mathbf{e}} = \begin{pmatrix} 35 \\ 45 \\ 65 \\ 96 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 095 \\ 203 \end{pmatrix} = \begin{pmatrix} 052 \\ -051 \\ -054 \\ 053 \end{pmatrix}$$

y la suma de cuadrados de los residuos:

$$\hat{\mathbf{e}}'\hat{\mathbf{e}} = \begin{pmatrix} 052 & -051 & -054 & 053 \end{pmatrix} \begin{pmatrix} 052 \\ -051 \\ -054 \\ 053 \end{pmatrix} = 1103$$

es el valor mínimo de la función L . Si se acepta que todos los residuos son variables aleatorias que pertenecen a una distribución de media igual a cero y varianza σ_e^2 , el valor 1.103 dividido por el número de observaciones menos uno es un estimador de σ_e^2 .

Ejercicio: sustituir en las ecuaciones del modelo las variables a y b por números diferentes y comprobar que $\hat{\mathbf{e}}'\hat{\mathbf{e}}$ es en todos los casos mayor que 1.103.

5.2.2. Ecuaciones normales con rango incompleto. Restricciones en el vector β

La matriz de coeficientes de las ecuaciones normales puede tener rango incompleto.

Como se verá más adelante (capítulo 5), $(\mathbf{X}'\mathbf{X})$ tiene rango incompleto cuando en el modelo se contempla más de un factor de influencia principal (con niveles). Normalmente las covariables no dan lugar a dependencias en las ecuaciones normales.

En esos casos, hay dos posibilidades para hallar estimadores para los parámetros:

1. Construir las ecuaciones normales y hallar una solución mediante una inversa generalizada
2. introducir restricciones en el vector β

Las restricciones son nuevas ecuaciones, que complementan la falta información suficiente. Sólomente se utilizan para posibilitar el cálculo: las restricciones son arbitrarias y la falta de información persiste.

En estas condiciones, el método de los mínimos cuadrados consiste en encontrar el vector β que minimice $(\mathbf{e}'\mathbf{e})$ bajo las condiciones adicionales o restricciones.

La minimización condicionada de una función se puede llevar a cabo mediante multiplicadores de Lagrange.

Cuando el número de parámetros a estimar es n y el rango de la matriz de coeficientes de las ecuaciones normales es r , hay que introducir $(n-r)$ restricciones y $(n-r)$ multiplicadores de Lagrange. Estos multiplicadores forman el vector \mathbf{g} .

Por ejemplo, sea un modelo con tres factores de influencia, cuyos efectos son:

- a para el primer factor
- c_1, c_2 y c_3 para el segundo factor
- d_1 y d_2 para el tercer factor

y en cuya matriz de coeficientes de las ecuaciones normales hay dos dependencias lineales: $n - r = 2$.

Se puede introducir para el segundo y tercer factores la restricción de que la suma de los efectos de todos sus niveles sea cero:

$$\begin{aligned}c_l + c_2 + c_3 &= 0 \\d_l + d_2 &= 0\end{aligned}$$

En notación matricial: $\mathbf{C}'\beta = \mathbf{f}$

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ c_l \\ c_2 \\ c_3 \\ d_l \\ d_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

La minimización de la función $(\mathbf{e}'\mathbf{e})$ bajo la condición de que $\mathbf{C}'\beta = \mathbf{f}$ da lugar a unas nuevas ecuaciones normales:

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{C}' \\ \mathbf{C}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \beta \\ \mathbf{g} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{f} \end{pmatrix}$$

Las antiguas ecuaciones normales se han extendido con las restricciones.

Con distintas matrices de restricciones \mathbf{C}' , se encuentran distintas soluciones para β , del mismo modo que con distintas inversas generalizadas se encuentran distintas soluciones a un sistema de ecuaciones. El procedimiento de los multiplicadores de Lagrange tiene con respecto a la utilización de inversas generalizadas el inconveniente de que no se puede probar qué funciones de los parámetros son estimables.

En el capítulo 5 se ilustrará este procedimiento por medio de un ejemplo.

5.3. Comparación entre dos modelos. Sesgo.

Un criador de caracoles vecino del anterior decide utilizar el procedimiento desarrollado por éste en su propia granja. Como tiene dos tipos distintos de caracoles, unos amarillos y otros marrones, toma una muestra de doble tamaño.

Los datos son:

Animal número	Tipo	Edad (meses)	Peso (g)
1	amarillo	1	3.1
2	marrón	1	3.9

Animal número	Tipo	Edad (meses)	Peso (g)
3	amarillo	2	4.9
4	marrón	2	6.1
5	amarillo	3	7.1
6	marrón	3	8.2
7	amarillo	4	8.9
8	marrón	4	9.8

El modelo es:

$$\begin{pmatrix} 31 \\ 39 \\ 49 \\ 61 \\ 71 \\ 82 \\ 89 \\ 98 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 4 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \mathbf{e}$$

La ecuaciones normales son: $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 4 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 \end{pmatrix} \begin{pmatrix} 31 \\ 39 \\ 49 \\ 61 \\ 71 \\ 82 \\ 89 \\ 98 \end{pmatrix}$$

$$\begin{pmatrix} 8 & 20 \\ 20 & 60 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 520 \\ 1477 \end{pmatrix}$$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = 1/80 \begin{pmatrix} 60 & -20 \\ -20 & 8 \end{pmatrix} \begin{pmatrix} 520 \\ 1477 \end{pmatrix} = \begin{pmatrix} 2075 \\ 1770 \end{pmatrix}$$

Los pesos de los otros caracoles se estiman mediante la siguiente expresión:

$$\hat{y}_j = 2075 + 177 \times x_j$$

donde x_j es la edad en meses.

2.075 y 1.77 son los mejores estimadores para los parámetros a y b de ese modelo: los valores con los que $(\hat{\mathbf{e}}'\hat{\mathbf{e}})$ es mínimo.

Los estimadores de los residuos son:

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \begin{pmatrix} 31 \\ 39 \\ 49 \\ 61 \\ 71 \\ 82 \\ 89 \\ 98 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 4 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2075 \\ 1770 \end{pmatrix} = \begin{pmatrix} -0745 \\ 0055 \\ -0715 \\ 0485 \\ -0285 \\ 0815 \\ -0255 \\ 0645 \end{pmatrix}$$

y la suma de cuadrados $\hat{\mathbf{e}}'\hat{\mathbf{e}} = 2531$

En este modelo no se ha tenido en cuenta que existe una diferencia genética entre caracoles. La suposición de que todos los animales, independientemente de su tipo genético (amarillos o marrones) tienen el mismo peso al nacimiento parece cuestionable. Por eso sugerimos al segundo criador de caracoles que utilice otro modelo, en el cual se consideren dos pesos al nacimiento diferentes. Los factores de influencia y sus efectos serían en este nuevo modelo:

- Factor de influencia peso al nacimiento: la influencia del peso al nacimiento tiene dos niveles o clases:
 - efecto peso al nacimiento de tipo amarillo: a_1
 - efecto peso al nacimiento de tipo marrón: a_2
- Factor de influencia edad (x_i): efecto ganancia de peso bx_i

Seguimos ASUMIENDO que el aumento de peso con la edad es lineal e igual para los dos tipos genéticos.

Las ecuaciones del nuevo modelo son:

$$\begin{pmatrix} 31 \\ 39 \\ 49 \\ 61 \\ 71 \\ 82 \\ 89 \\ 98 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 1 & 3 \\ 1 & 0 & 4 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ b \end{pmatrix} + \mathbf{e}$$

Las ecuaciones normales :

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$$

$$\begin{pmatrix} 4 & 0 & 10 \\ 0 & 4 & 10 \\ 10 & 10 & 60 \end{pmatrix} \hat{\beta} = \begin{pmatrix} 240 \\ 280 \\ 1477 \end{pmatrix}$$

Los estimadores:

$$\hat{\beta} = \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{b} \end{pmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{pmatrix} 1570 \\ 2575 \\ 1770 \end{pmatrix}$$

Los residuos:

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\beta} = \begin{pmatrix} 31 \\ 39 \\ 49 \\ 61 \\ 71 \\ 82 \\ 89 \\ 98 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 1 & 3 \\ 1 & 0 & 4 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1570 \\ 2575 \\ 1770 \end{pmatrix} = \begin{pmatrix} -0240 \\ -0685 \\ -0210 \\ -0015 \\ 0220 \\ 0315 \\ 0250 \\ 0145 \end{pmatrix}$$

Y la suma de cuadrados de los residuos: $\hat{\mathbf{e}}'\hat{\mathbf{e}} = 08023$

A partir de una misma muestra hemos encontrado, para dos modelos distintos, los mejores estimadores de los parámetros de cada modelo. En un caso, la suma de los cuadrados de los residuos fue 2.531 y en el otro 0.8023; el segundo modelo ha resultado ser más adecuado que el primero, porque las previsiones que hacemos con él, en general, se ajustan más a la realidad conocida.

Comparemos ahora los estimadores de los residuos de los dos modelos:

Animal número	Tipo	Residuo Modelo 1	Residuo Modelo 2
1	amarillo	-0.745	-0.240
2	marrón	0.055	-0.682
3	amarillo	-0.715	-0.210
4	marrón	0.485	-0.015
5	amarillo	-0.285	0.220
6	marrón	0.815	0.315
7	amarillo	-0.255	0.250
8	marrón	0.645	0.145

Se observa que con el primer modelo, todos los caracoles amarillos se han sobreestimado y todos los marrones se han subestimado. Con el primer modelo se comete un ERROR SISTEMÁTICO.

Esto es lógico si aceptamos que el segundo modelo es mejor. Se ha utilizado como estimador de todos pesos la única expresión

$$\hat{y}_j = 2075 + 177 \times x_j$$

en vez de aplicar:

- para los amarillos $\hat{y}_j = 1570 + 177 \times x_j$
- y para los marrones $\hat{y}_j = 2575 + 177 \times x_j$

donde:

$$1570 < 2075 < 2575$$

Asumiendo que todos los caracoles, independientemente del tipo tienen el mismo peso al nacimiento, se comete un error sistemático en la estimación y los estimadores son SENGADOS.

SI SE DEJA DE CONSIDERAR UN FACTOR DE INFLUENCIA IMPORTANTE, SE OBTIENEN ESTIMADORES SENGADOS

5.4. Vector de esperanzas matemáticas y matriz de varianzas-covarianzas de un vector de variables aleatorias.

5.4.1. Definición

El vector de esperanzas matemáticas de un vector de variables aleatorias es el vector de las esperanzas matemáticas de las variables individuales que lo componen.

La matriz de varianzas-covarianzas de un vector de variables aleatorias es una matriz cuadrada, simétrica, de dimensión n , igual a la dimensión del vector. En la diagonal de esa matriz se encuentran las varianzas y fuera de la diagonal, las covarianzas entre las distribuciones de las variables del vector.

Sean r , s , t y u cuatro variables aleatorias con esperanzas matemáticas $E(r)$, $E(s)$, $E(t)$ y $E(u)$, respectivamente y con varianzas σ_r^2 , σ_s^2 , σ_t^2 y σ_u^2 . Sean las covarianzas entre ellas σ_{rs} , σ_{st} , σ_{ru} , etc.

El vector $\mathbf{v} = \begin{pmatrix} r \\ s \\ t \\ u \end{pmatrix}$ es un vector de variables aleatorias.

La esperanza matemática de \mathbf{v} es el vector: $E(\mathbf{v}) = \begin{pmatrix} E(r) \\ E(s) \\ E(t) \\ E(u) \end{pmatrix}$

La matriz de varianzas-covarianzas es:

$$Var(\mathbf{v}) = \begin{pmatrix} \sigma_r^2 & \sigma_{rs} & \sigma_{rt} & \sigma_{ru} \\ \sigma_{rs} & \sigma_s^2 & \sigma_{st} & \sigma_{su} \\ \sigma_{rt} & \sigma_{st} & \sigma_t^2 & \sigma_{tu} \\ \sigma_{ru} & \sigma_{su} & \sigma_{tu} & \sigma_u^2 \end{pmatrix}$$

5.4.2. Matriz de varianzas-covarianzas de un vector con esperanza matemática nula.

Supongamos que las esperanzas matemáticas de todas las variables aleatorias son nulas:

$$E(\mathbf{v}) = \begin{pmatrix} E(r) \\ E(s) \\ E(t) \\ E(u) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

La variable r es una variable aleatoria con una determinada distribución de frecuencias y media cero. Hay un número N (infinito) de valores que puede tomar la variable r con mayor o menor probabilidad. La media de todos los casos posibles ponderados por su frecuencia o probabilidad es cero (la esperanza matemática es cero).

La varianza de la distribución de r es la suma de los cuadrados de las desviaciones respecto a la media de todos los casos, dividido por el número de casos:

$$\sigma_r^2 = \frac{1}{N} \sum_i (r_i - E(r))^2 = \frac{1}{N} \sum_i (r_i)^2 = E(r^2)$$

La varianza de una variable aleatoria de esperanza matemática igual a cero, es igual a la esperanza matemática del cuadrado de la variable.

Del mismo modo, la covarianza entre r y s es:

$$\sigma_{rs} = \frac{1}{N} \sum_i (r_i - E(r))(s_i - E(s)) = \frac{1}{N} \sum_i r_i s_i = E(rs)$$

La covarianza entre las variables r y s es igual a la esperanza matemática de la variable $(r \times s)$.

En este caso, la matriz de varianzas y covarianzas del vector \mathbf{v} queda como sigue:

$$\begin{aligned}\mathbf{R} = Var(\mathbf{v}) &= \begin{pmatrix} E(r^2) & E(rs) & E(rt) & E(ru) \\ E(rs) & E(s^2) & E(st) & E(su) \\ E(rt) & E(st) & E(t^2) & E(tu) \\ E(ru) & E(su) & E(tu) & E(u^2) \end{pmatrix} = \\ &= E \begin{pmatrix} (r^2) & (rs) & (rt) & (ru) \\ (rs) & (s^2) & (st) & (su) \\ (rt) & (st) & (t^2) & (tu) \\ (ru) & (su) & (tu) & (u^2) \end{pmatrix} = E \begin{pmatrix} r \\ s \\ t \\ u \end{pmatrix} (r \quad s \quad t \quad u) = E(\mathbf{v}\mathbf{v}') \\ \mathbf{R} &= E(\mathbf{v}\mathbf{v}')$$

\mathbf{R} es igual a la matriz de la esperanza matemática de $\mathbf{v}\mathbf{v}'$.

Sea \mathbf{L} un operador elemental, que transforma el vector \mathbf{v} en otro vector de variables aleatorias \mathbf{w} : $\mathbf{w} = \mathbf{L}\mathbf{v}$

\mathbf{L} es una matriz de constantes; por ello la esperanza matemática del nuevo vector \mathbf{w} es igual al producto de \mathbf{L} por el vector de esperanzas matemáticas de \mathbf{v} :

$$E(\mathbf{w}) = E(\mathbf{L}\mathbf{v}) = \mathbf{L}E(\mathbf{v})$$

La matriz de varianzas y covarianzas del nuevo vector es:

$$Var(\mathbf{w}) = Var(\mathbf{L}\mathbf{v}) = E((\mathbf{L}\mathbf{v})(\mathbf{L}\mathbf{v})') = E(\mathbf{L}\mathbf{v}\mathbf{v}'\mathbf{L}')$$

y como \mathbf{L} y \mathbf{L}' son constantes:

$$Var(\mathbf{w}) = \mathbf{L}E(\mathbf{v}\mathbf{v}')\mathbf{L}' = \mathbf{L}\mathbf{R}\mathbf{L}' = \mathbf{L}Var(\mathbf{v})\mathbf{L}'$$

5.5. Método de los mínimos cuadrados generalizado (GLS)

En un modelo lineal, el vector \mathbf{e} es un vector de variables aleatorias:

$$\mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}$$

El residuo e_1 pertenece a una distribución de media cero y varianza $\sigma_{e_1}^2$.

El residuo e_2 pertenece a una distribución de media cero y varianza $\sigma_{e_2}^2$.

El residuo e_n pertenece a una distribución de media cero y varianza $\sigma_{e_n}^2$.

Dado que no sabemos nada acerca de e_1 , e_2 , etc., no tenemos ningún criterio para juzgar si la varianza de la distribución de la que se ha obtenido e_1 es mayor o menor que la varianza de la distribución de donde ha salido e_2 . Por eso tenemos que asumir, que la varianza de la distribución es la misma para todos los residuos:

$$\sigma_{e_1}^2 = \sigma_{e_2}^2 = \dots = \sigma_{e_n}^2 = \sigma_e^2$$

Del mismo modo, no tenemos ninguna razón para pensar que se producen covarianzas entre los residuos. Existe una covarianza entre dos variables, cuando las dos tienen causas comunes. Pero hemos partido de la base de que no sabemos absolutamente nada acerca de las causas de e_1 , e_2 , etc. Por lo tanto, en principio debemos asumir, que todas las covarianzas $\mathbf{R} = \text{Var}(\mathbf{e})$ son cero.

Así la esperanza matemática y la matriz de varianzas-covarianzas de los residuos de un modelo LSQ son:

$$E(\mathbf{e}) = \begin{pmatrix} E(e_1) \\ E(e_2) \\ E(e_3) \\ \dots \\ \dots \\ E(e_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} \sigma_e^2 & 0 & 0 & \dots & \dots & 0 \\ 0 & \sigma_e^2 & 0 & \dots & \dots & 0 \\ 0 & 0 & \sigma_e^2 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & \sigma_e^2 \end{pmatrix} = \mathbf{I}\sigma_e^2$$

$E(\mathbf{e})$ es un vector nulo y $\text{Var}(\mathbf{e})$ es una matriz escalar. Esta es la primera ASUNCIÓN de modelo (ver apdo. 4.1), expresada en notación matricial.

Pero también podría darse el caso de que sí hubiera algún criterio que hiciera esperar, que las varianzas de los residuos fueran de diferente magnitud, o que hay covarianza entre residuos. En este caso, \mathbf{R} ya no sería una matriz escalar.

Ejemplos de ellos podrían ser los siguientes:

Ejemplo primero: se pretende valorar datos de crecimiento de terneros, parte de los cuales se han criado en una gran estación de testaje, mientras que el resto han sido controlados individualmente sin salir de sus granjas de origen.

En una estación, las condiciones ambientales están estandarizadas: todos los terneros son tratados y alimentados de la misma manera, de acuerdo con unas normas prefijadas y han permanecido todos en el mismo lugar. En cambio, las condiciones en que se han criado los terneros que permanecen en sus lugares de origen variarán de unas explotaciones a otras, sin que sea normalmente posible recoger información suficiente sobre ellas. En estas condiciones, los factores ambientales desconocidos presentarán más variabilidad entre los animales probados en estación que entre los probados en campo. El efecto de estos factores está contenido en el residuo, de modo que tenemos razones para sospechar, que habrá diferentes varianzas del residuo en los casos de prueba en estación, que en los casos de testaje en campo. En este caso la matriz \mathbf{R} sería:

$$\mathbf{R} = \begin{pmatrix} \sigma_{e_s}^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_{e_s}^2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \sigma_{e_s}^2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \sigma_{e_c}^2 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \sigma_{e_c}^2 \end{pmatrix}$$

donde $\sigma_{e_s}^2$ y $\sigma_{e_c}^2$ serían las varianzas de los residuos en estación y en campo, respectivamente.

Ejemplo segundo: se pretende valorar con un modelo resultados de lactaciones de vacas. De cada lactación se han recogido varios caracteres a saber, cantidad de leche, cantidad de grasa y cantidad de proteína producida. Así la vaca número 2345678, contribuye al modelo con tres rendimientos (leche, grasa y proteína), para los que se pueden escribir tres ecuaciones:

$$\begin{aligned} y_{l2345678} &= \sum efectos + e_{l2345678} \\ y_{g2345678} &= \sum efectos + e_{g2345678} \\ y_{p2345678} &= \sum efectos + e_{p2345678} \end{aligned}$$

Los tres residuos, $e_{l2345678}$, $e_{g2345678}$ y $e_{p2345678}$ del vector \mathbf{e} corresponden a caracteres distintos. Las varianzas de los tres no pueden ser iguales.

Además los tres residuos pertenecen a la misma vaca número 2345678. Los tres son variables aleatorias, que tienen causas desconocidas. Estas causas son en gran parte efectos ambientales, que no podemos registrar. Pero como los tres pertenecen a la misma vaca, las condiciones ambientales fueron para los tres las mismas. Las causas de los tres restos son, en parte las mismas, es decir, podemos pensar que habrá covarianza entre los tres residuos.

En este modelo (modelo de caracteres múltiples), la matriz \mathbf{R} no es ni siquiera diagonal.

Podríamos mencionar muchos otros casos, en los que los residuos de las diferentes ecuaciones del modelo no tienen igual importancia, o en los que hay relaciones entre los residuos. En todos estos casos, la función a minimizar para el método

de los mínimos cuadrados es otra distinta a $(\mathbf{e}'\mathbf{e})$ (ver apdo. 4.2), porque en esta función:

- todos los residuos se consideran de la misma forma, cuando su importancia relativa no es igual y
- las relaciones entre residuos no se tienen en cuenta.

Así se desprecia una parte de información importante. La búsqueda de una nueva función a minimizar que tuviera en cuenta estas nuevas informaciones y el posterior desarrollo de las nuevas ecuaciones normales sería enormemente complicado, si no se dispusiera del conocimiento del álgebra de las matrices, en la que se simplifican los problemas más complicados.

Sea nuestro modelo **(1)**:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$$

donde $E(\mathbf{e}) = \mathbf{0}$ y $Var(\mathbf{e}) = \mathbf{R}$

Dado que \mathbf{R} no es una matriz escalar, no podemos aplicar el método de los mínimos cuadrados a nuestro modelo. Necesitaríamos otro modelo **(2)**:

$$\mathbf{z} = \mathbf{Q}\beta + \mathbf{f}$$

donde $E(\mathbf{f}) = \mathbf{0}$ y $Var(\mathbf{f}) = \mathbf{S}(\text{escalar})$

en el que los parámetros de β fueran los mismos de nuestro modelo original. Si dispusiéramos de datos para este segundo modelo, sería posible aplicar el método de los mínimos cuadrados y hacer una estimación de β .

Como se vio en el apdo 3.3.4, se puede transformar un sistema de ecuaciones por medio de operaciones elementales, sin que varíe la solución del mismo. Nuestro problema es pues, encontrar una matriz operador elemental, que transforme el sistema de ecuaciones (1) en el sistema de ecuaciones (2), de modo que $E(\mathbf{f})$ sea cero y \mathbf{S} sea una matriz escalar.

Una propiedad de las matrices de varianzas-covarianzas (como por ejemplo \mathbf{R}) y que no podemos demostrar dentro de las dimensiones de estos apuntes, es que para ellas existe siempre una única matriz no singular \mathbf{P} tal que:

$$\mathbf{P}'\mathbf{P} = \mathbf{R}$$

Se puede probar que la inversa de la matriz \mathbf{P} , es la matriz que buscábamos para llevar a cabo la transformación del sistema.

En efecto, si llevamos a cabo la transformación, premultiplicando por \mathbf{P}^{-1} a ambos lados del signo igual:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$$

$$\mathbf{P}^{-1}\mathbf{y} = \mathbf{P}^{-1}\mathbf{X}\beta + \mathbf{P}^{-1}\mathbf{e}$$

Si llamamos a las matrices:

$$\mathbf{P}^{-1}\mathbf{y} = \mathbf{z}$$

$$\mathbf{P}^{-1}\mathbf{X} = \mathbf{Q}$$

$$\mathbf{P}^{-1}\mathbf{e} = \mathbf{f}$$

tenemos un nuevo sistema de ecuaciones:

$$\mathbf{z} = \mathbf{Q}\beta + \mathbf{f}$$

La esperanza matemática de \mathbf{f} es: $E(\mathbf{f}) = E(\mathbf{P}^{-1}\mathbf{e}) = \mathbf{P}^{-1}E(\mathbf{e}) = \mathbf{0}$

La matriz de varianzas-covarianzas de \mathbf{f} es:

$$Var(\mathbf{f}) = Var(\mathbf{P}^{-1}\mathbf{e}) = \mathbf{P}^{-1}Var(\mathbf{e})(\mathbf{P}^{-1})' = \mathbf{P}^{-1}\mathbf{R}(\mathbf{P}^{-1})'$$

La matriz \mathbf{P} es simétrica y su inversa es igualmente simétrica, lo que quiere decir que: $(\mathbf{P}^{-1})' = \mathbf{P}^{-1}$

Entonces: $Var(\mathbf{f}) = \mathbf{P}^{-1}\mathbf{R}\mathbf{P}^{-1} = \mathbf{P}^{-1}\mathbf{P}'\mathbf{P}\mathbf{P}^{-1} = \mathbf{I}$

La matriz de varianzas-covarianzas del vector de los nuevos residuos es una matriz unidad o identidad y por lo tanto es una matriz escalar, con lo que se cumple lo que exigíamos a la transformación.

Aplicando el método de los mínimos cuadrados en el modelo (2) :

$$\mathbf{Q}'\mathbf{Q}\hat{\beta} = \mathbf{Q}'\mathbf{z}$$

donde:

$$\mathbf{Q} = \mathbf{P}^{-1}\mathbf{X}; \mathbf{Q}' = \mathbf{X}'\mathbf{P}^{-1}; \mathbf{z} = \mathbf{P}^{-1}\mathbf{y}$$

$$\mathbf{X}'\mathbf{P}^{-1}\mathbf{P}^{-1}\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{R}^{-1}\mathbf{y}$$

$$\boxed{\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{R}^{-1}\mathbf{y}}$$

Esta última expresión se corresponde con las ecuaciones normales del método de los mínimos cuadrados generalizado (GLS), a aplicar cuando la matriz de varianzas-covarianzas de los residuos no es escalar.

5.6. Efectos fijos y al azar. Un modelo mixto

Además del residuo, otros efectos del modelo pueden ser variables aleatorias.

Pensemos en un modelo para analizar el rendimiento lechero en ganado vacuno. Los factores de influencia sobre la producción de leche de una vaca son su ambiente y su constitución genética.

Influencia ambiental:

Un modelo real será mucho más complicado, pero para el ejemplo vamos a considerar, que todos los efectos ambientales se pueden resumir en un solo efecto, que llamaremos “efecto rebaño”. Este sería un efecto común a todas las vacas presentes en ese rebaño: como el ambiente es el mismo para todas las vacas del rebaño, todas reciben el mismo efecto.

Factores genéticos:

Consideremos un animal. Según la genética de poblaciones ⁷ y la genética cuantitativa ⁸, en una población de animales que se reproducen (una raza), existe una cierta variabilidad genética aditiva. Un animal que sale de dicha población no puede tener un valor genético “cualquiera”, ya que su valor genético estará condicionado por la variabilidad genética de la población. Un animal cuyos padres se conocen tendrá, a su vez, su variabilidad condicionada por el valor genético de sus padres y una cierta desviación llamada muestreo mendeliano, que se explica ahora.

Por ejemplo, todas las vacas hijas del toro número 7654 reciben de su padre una parte de su genotipo que se ha transmitido a través de un espermatozoide. A poco que sepamos de la formación de gametos y pensando que la producción de leche es un carácter regulado por un número muy grande de genes de efectos individuales muy pequeños, no podemos en ningún caso aceptar, que todos los gametos producidos por el toro número 7654 son iguales, ni por lo tanto, que tienen el mismo valor. Por tanto, el promedio de todas las hijas posibles de un toro es igual a $1/2$ del valor genético del toro, pero cada vaca recibe un gameto diferente al azar.

Por ambas razones, se dice que el valor genético de un animal es un “efecto aleatorio” ya que se visualiza el valor genético del animal como debido en parte al azar, ya que podría haber tenido otros genes o recibido otros gametos de sus padres.

Por lo tanto, en nuestro modelo habrá dos tipos diferentes de efectos:

- El efecto “rebaño” es un EFECTO FIJO, porque *no* se considera que haya una población de rebaños de la que nuestro rebaño sería uno entre muchos.

⁷Gillespie, John H. Population genetics: a concise guide. JHU Press, 2004.

⁸Falconer, D. S., & Mackay, T. F. C. (1996). Introduction to quantitative genetics. Longman New York.

- El efecto “animal” es un EFECTO AL AZAR O ALEATORIO, porque consideramos que su valor es una variable aleatoria dado que los animales salen de poblaciones con cierta variación genética.

La diferencia entre efectos fijos y efectos al azar no es siempre clara. Por ejemplo, se puede discutir nuestra afirmación anterior de que el efecto “rebaño” porque en realidad todos los rebaños de una raza tienden a parecerse. Los efectos de un factor de influencia se pueden considerar unas veces como fijos y otras como al azar, dependiendo de las circunstancias. La definición de un efecto como fijo o como al azar, pertenece a las ASUNCIONES del modelo.

Un modelo en el que se encuentran al mismo tiempo efectos fijos y efectos al azar es un MODELO MIXTO (Mixed Model). Por supuesto, en un modelo mixto, los efectos fijos y los efectos al azar no se pueden tratar de la misma manera.

Las esperanzas matemáticas y matrices de varianzas-covarianzas de los efectos al azar serán, para el efecto genético w :

$$E(\mathbf{w}) = \mathbf{0}; Var(\mathbf{w}) = \mathbf{G}$$

y para el residuo e :

$$E(\mathbf{e}) = \mathbf{0}; Var(\mathbf{e}) = \mathbf{R}$$

El modelo es (3):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{w} + \mathbf{e}$$

donde se puede definir un vector de “nuevos” residuos $\mathbf{f} = \mathbf{Z}\mathbf{w} + \mathbf{e}$ tal que un modelo equivalente es $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{f}$.

La esperanza matemática y la matriz de varianzas-covarianzas del vector \mathbf{f} es:

$$E(\mathbf{f}) = E(\mathbf{Z}\mathbf{w} + \mathbf{e}) = E(\mathbf{Z}\mathbf{w}) + E(\mathbf{e}) = \mathbf{Z}\mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$Var(\mathbf{f}) = Var(\mathbf{Z}\mathbf{w} + \mathbf{e}) = Var(\mathbf{Z}\mathbf{w}) + Var(\mathbf{e}) = \mathbf{ZGZ}' + \mathbf{R} = \mathbf{V}$$

Aplicando el método generalizado de los mínimos cuadrados (GLS), las ecuaciones normales son:

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

5.7. BLUP (Best Linear Unbiased Prediction)

Una vez obtenido $\hat{\beta}$, los estimadores (o predictores) de los valores genéticos w se pueden obtener de esta manera:

$$\hat{\mathbf{w}} = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta})$$

Estos estimadores $\hat{\mathbf{w}}$ son la *esperanza condicional* de los valores genéticos dados los datos (\mathbf{y}), se denominan Best Linear Unbiased Prediction (BLUP), y tienen una serie de propiedades óptimas que no detallaremos aquí.

Estas ecuaciones, en la mayor parte de los casos de interés para la mejora genética animal resultaban imposibles de resolver. Uno de los principales problemas para ello es la inversión de la matriz \mathbf{V} de las varianzas-covarianzas entre los rendimientos individuales cuya dimensión es $n \times n$, donde n es el número de observaciones.

La matriz \mathbf{R} tiene la misma dimensión de \mathbf{V} , pero normalmente tiene una estructura interna que permite invertirla fácilmente (muchas veces \mathbf{R} es una matriz diagonal o diagonal de bloques). La matriz \mathbf{G} puede tener una dimensión menor que \mathbf{V} y \mathbf{R} y su inversa es además fácil de calcular. Por esto, la aplicación de modelos lineales en la mejora genética dio un gran paso adelante, cuando Henderson y Searle consiguieron demostrar, que las mismas soluciones para \mathbf{b} y \mathbf{w} se pueden obtener también mediante la siguiente expresión, que se denomina habitualmente las ecuaciones del modelo mixto:

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}$$

Las soluciones de este último sistema de ecuaciones son:

- $\hat{\beta}$: estimadores GLS (o BLUE) de los efectos fijos β
- $\hat{\mathbf{w}}$: predicciones BLUP de los efectos al azar

Los efectos genéticos calculados son:

- predictores óptimos lineales (Best Linear Predictions), porque se han obtenido mediante aplicación del método de los mínimos cuadrados generalizados, a partir de un modelo lineal.
- insesgados (Unbiased), porque se ha tenido en cuenta al mismo tiempo los factores de influencia con efectos fijos (ver 4.3) .

Por eso estos valores se llaman “best linear unbiased predictors”, o abreviado “BLUP” de los efectos genéticos considerados.

6. Ejemplo de utilización práctica de un modelo lineal

Como ejemplo para facilitar la comprensión de la teoría de los modelos lineales, se va a calcular BLUP de los valores de cría para el rendimiento en cantidad de leche, de una población de diez vacas que parieron por primera vez en el mismo año y estación y que han acabado recientemente su primera lactación.

6.1. Datos

Los datos recogidos de la población son los siguientes:

Vaca	Raza	Padre	Edad al primer parto	Rebaño	kg de leche
6	1	1	20	1	6251.1
7	1	2	40	1	6173.5
8	1	3	30	2	5775.0
9	1	2	26	2	5775.6
10	1	3	38	3	5973.8
11	1	1	24	3	6050.9
12	2	4	26	4	4150.6
13	2	5	32	4	4249.7
14	2	5	36	5	3849.1
15	2	4	38	5	3750.3

Media de la edad al primer parto: 30

6.2. Factores de influencia

El modelo considera los siguientes factores de influencia:

6.2.1. Influencia genética sobre el rendimiento de una vaca

Se puede descomponer en:

- Media de la población (μ) la media de la población corresponde al efecto “vaca”, es decir la esperanza matemática del rendimiento de una vaca, sin consideración de otros factores de influencia.

El valor de μ responde a la pregunta de cuántos kg de leche dará una vaca de la que no se conoce absolutamente nada. Esta pregunta no tiene ningún sentido especial para la mejora genética. Si lo tendría si quisiéramos comparar especies: qué animal da más leche, una hembra de ratón, una vaca o una elefanta.

- Raza (R): la raza de la vaca es un factor de influencia, cuyos efectos fijos se clasifican en dos niveles o clases:
 - Efecto R1 para una vaca de raza 1

- Efecto R2 para una vaca de raza 2

R1 es la desviación sobre μ (esperanza matemática de la producción de una vaca), de la esperanza matemática de la producción de una vaca de la raza 1. R1 es la contestación a la pregunta de cuánta leche más (o menos) esperamos que dé una vaca de la que solo sabemos que es de la raza 1, que otra vaca de la que no sabemos absolutamente nada. Por medio de los efectos R1 y R2 intentamos comparar las razas.

— Nivel genético de la k-ésima vaca, cuyos padre y madre son los dos de una misma raza.

Este efecto le corresponde únicamente a la k-ésima vaca, que ya es un individuo particular. Este efecto es la contestación a la pregunta de cuánto mejor o peor es la k-ésima vaca que la “vaca indeterminada” de su raza. Es un efecto al azar (ver 5.6). La vaca proviene de un espermatozoide de su padre y de un óvulo de su madre, y ambos gametos contenían solamente una parte de los genes de los padres.

6.2.2. Factores de influencia ambientales

En este caso tenemos:

— Edad al primer parto (EPP)

Es sabido que una vaca que pare por primera vez a una edad relativamente temprana, produce en su primera lactación menos leche de la que hubiera producido en caso de que su edad al primer parto hubiera sido más alta, manteniéndose todos los demás factores de influencia iguales. La edad es una variable cuantitativa continua. Se puede considerar como una covariable y explicar su efecto por medio de una ecuación de regresión. Es decir, el efecto de la edad al primer parto es:

$$Efecto - EPP = b(EPP_k - E\bar{P}P)$$

donde ⁹

- EPP_k : es la edad a la que parió por primera vez la k-ésima vaca
- $E\bar{P}P$ es la edad media de las vacas al primer parto

— Rebaño

Ya se habló en el apartado 5.6 del efecto explotación, de modo que ya no se insistirá más aquí sobre él. Para no confundirlo con el efecto “raz#”, se representarán los efectos rebaño, como H_i . A la vista de la estructura de los datos se puede observar que en cada rebaño hay únicamente una raza. Un efecto distribuido de este modo dentro de los niveles o clases de otro efecto, recibe la denominación de efecto “nested”, jerarquizado (en inglés literalmente “anidado”)

⁹técnicamente sustraer $E\bar{P}P$ no es necesario

. Hay una estructura jerárquica, de modo que la clase “raza 1” del factor de influencia de rango jerárquico superior, engloba a todos los animales de las clases 1, 3 y 3 del factor subordinado, que es el rebaño. Todo los animales de los rebaños 1, 2 y 3 pertenecen a la raza 1; del mismo modo, todos los animales de los rebaños 4 y 5 pertenecen a la raza 2. Esto tiene una influencia sobre el rango de la matriz de coeficientes de las ecuaciones del modelo mixto (MME) y sobre la estimabilidad de las distintas funciones de parámetros.

Estos cinco factores descritos no son todos los que normalmente se consideran en la realidad. Se trata solamente de un ejemplo, que ilustre sobre la estructura de la matriz de coeficientes de las MME.

6.2.3. Modelo

A la vista de los factores de influencia descritos más arriba, el modelo se escribe de la siguiente forma:

$$y_{ijk} = \mu + b(EPP_k - E\bar{P}P) + R_i + H_{i(j)} + e_{ijk}$$

donde:

- y_{ijk} : rendimiento observado en la k-ésima vaca, que se encuentra en el j-ésimo rebaño, donde las vacas son de la i-ésima raza
- μ : media de la población
- b : coeficiente de regresión
- EPP_k : edad al primer parto de la k-ésima vaca
- $E\bar{P}P$: edad al primer parto media de las vacas
- R_i : efecto fijo de la i-ésima raza
- $H_{i(j)}$: efecto fijo del j-ésimo rebaño, subordinado a la i-ésima raza
- u_k resto del valor aditivo de la k-ésima vaca, o efecto al azar de la propia vaca
- e_{ijk} : residuo de la observación de la k-ésima vaca

En notación matricial el modelo queda:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

donde :

- \mathbf{y} : vector de las observaciones
- \mathbf{X} : matriz de incidencia de los efectos fijos
- $\boldsymbol{\beta}$: vector de los efectos fijos
- \mathbf{Z} : matriz de incidencia de los efectos aleatorios
- \mathbf{u} : vector de los efectos aleatorios
- \mathbf{e} : vector de los residuos

El conjunto de las ecuaciones del modelo en notación matricial se presentó previamente y es:

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}$$

6.2.4. Asunciones del modelo

- La varianza del residuo es igual para todas las vacas y los residuos son independientes (no correlacionados)

$$\mathbf{R} = \mathbf{I}\sigma_e^2$$

$$\mathbf{R}^{-1} = \mathbf{I}\frac{1}{\sigma_e^2}$$

Se asume la varianza residual de $\sigma_e^2 = 300000$

- No hay correlación entre los efectos genéticos aleatorios y residuos:
 $Cov(u, e) = 0$
- El único parentesco entre los animales es conocido a los partir de los datos. Se asumen padres desconocidos para los padres y como las madres de las vacas se ignoran, asumimos un pedigree con esta forma:

animal	padre	madre
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	1	0
7	2	0
8	3	0
9	2	0
10	3	0
11	1	0
12	4	0
13	5	0
14	5	0
15	4	0

- La varianza aditiva de la población, σ_u^2 , permanece constante de una generación a otra. La matriz de covarianza entre valores aditivos es, entonces,

$$\mathbf{G} = \mathbf{A}\sigma_u^2$$

donde \mathbf{A} es la matriz de relaciones aditivas (“numerator relationship matrix”). Por tanto

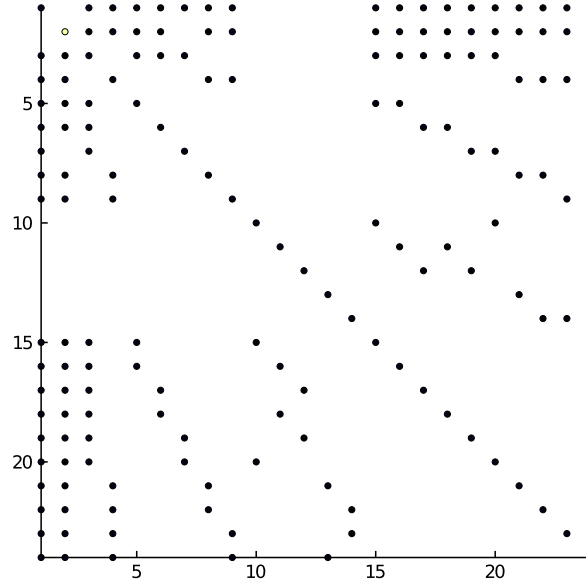
$$\mathbf{G}^{-1} = \mathbf{A}^{-1} \frac{1}{\sigma_u^2}$$

Asumimos $\sigma_u^2 = 300000$, lo que corresponde a una heredabilidad $h^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} = \frac{300000}{300000 + 900000} = 0.25$.

Dado que $\mathbf{R}^{-1} = \mathbf{I}_{\sigma_e^2}^{-1}$ y $\mathbf{G}^{-1} = \mathbf{A}^{-1} \frac{1}{\sigma_u^2}$, las ecuaciones del modelo mixto se pueden expresar *en este caso* como

$$\begin{pmatrix} \mathbf{X}'\mathbf{X}/\sigma_e^2 & \mathbf{X}'\mathbf{Z}/\sigma_e^2 \\ \mathbf{Z}'\mathbf{X}/\sigma_e^2 & \mathbf{Z}'\mathbf{Z}/\sigma_e^2 + \mathbf{A}^{-1}/\sigma_u^2 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{w}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y}/\sigma_e^2 \\ \mathbf{Z}'\mathbf{y}/\sigma_e^2 \end{pmatrix}$$

La matriz de las MME se muestra en el gráfico que sigue. Para que sea más fácil de ver, se han sustituido los ceros por puntos.



La matriz de coeficientes de las MME es una matriz cuadrada simétrica. Su dimensión es igual al número de parámetros a estimar, es decir:

- μ : 1
- coeficiente de regresión para EPP: 1
- efectos raza: 2
- efectos rebaño: 5
- efectos animales (padres y vacas): 15

Para un total de 24.

Por ejemplo, las 9 primeras líneas y columnas reflejan los efectos fijos y sus cruces con el efecto vaca (terminos $\mathbf{X}'\mathbf{X}/\sigma_e^2$, $\mathbf{X}'\mathbf{Z}/\sigma_e^2$, $\mathbf{Z}'\mathbf{X}/\sigma_e^2$). Las líneas y columnas 10 a 24 corresponden a las ecuaciones de los animales $\mathbf{Z}'\mathbf{Z}/\sigma_e^2 + \mathbf{A}^{-1}\sigma_u^2$. Se puede observar que las líneas y columnas 10 a 14 casi no tienen elementos, y es porque corresponden a los padres, que no tienen dato de producción y entran en las ecuaciones por relaciones de parentesco.

La primera fila de la inversa de la matriz, multiplicada por el vector $\begin{pmatrix} \mathbf{X}'\mathbf{y}/\sigma_e^2 \\ \mathbf{Z}'\mathbf{y}/\sigma_e^2 \end{pmatrix}$ da la solución para μ . La segunda fila es la que se utiliza para obtener el coeficiente de regresión. Las filas tercera y cuarta toman parte en el cálculo de los estimadores de los efectos de las razas 1 y 2 respectivamente etc.

Esta matriz de coeficientes no tiene rango completo. Se puede comprobar que existen las siguientes dependencias:

1. La suma de las filas de R1 y R2 es igual a la fila de μ
2. La suma de las filas de los rebaños 1, 2 y 3 es igual a la fila de la raza 1, en la que están jerarquizados esos rebaños.
3. La suma de las filas de los rebaños 4 y 5 es igual a la fila de la raza 2, en la que están jerarquizados esos rebaños.

Además y como consecuencia, la suma de todas las filas correspondientes a los rebaños es igual a la suma de las filas correspondientes a las razas e igual a la fila de μ .

Cuando se consideran en el modelo varios factores de influencia con efectos principales clasificatorios (no covariables) fijos, la suma de las filas de todos los efectos de un factor de influencia es igual a la fila de la media. Si en el modelo se considera una media y un número m de tales factores de influencia, habrá un número m de dependencias en la matriz de coeficientes. Si no se incluye la media, el número de dependencias será como mínimo de $m-1$.

Si es un modelo jerárquico, donde los efectos del j -ésimo factor de influencia están subordinados en los efectos del i -ésimo factor de influencia, entonces para el j -ésimo factor, en vez de una, aparecerán en la matriz de coeficientes tantas dependencias como clases tiene el i -ésimo factor. Así en el ejemplo, para el factor rebaño, subordinado en el factor raza, en vez de una aparecen dos dependencias, una para cada raza. Las covariables normalmente no dan lugar a ninguna dependencia en la matriz, pero existen covariables anidadas ("nested"), por ejemplo si el efecto de EPP dependiera de la raza. En ese caso sí que se generan dependencias.

En modelos complejos, el número de dependencias puede ser mayor y no siempre es posible determinarlo *a priori*.

En cambio, las dependencias que podrían aparecer para las filas de los efectos al azar se destruyen al añadir \mathbf{G}^{-1} .

Como las ecuaciones normales del modelo mixto no tienen rango completo, se pueden resolver mediante g—inversas. Para ello se borran tantas filas y sus correspondientes columnas como dependencias hay en la matriz (ver 4.4.2, procedimiento simplificado). La matriz resultante es la matriz **C** de 4.4.2 y su dimensión es igual al rango de la matriz de coeficientes. El número de ecuaciones que restan para cada factor ha de ser igual al número de grados de libertad de ese factor. En nuestro ejemplo:

- μ : 1
- coeficiente de regresión para EPP : 1
- efectos raza: 2-1=1
- efectos rebaño dentro de raza 1: 3-1=2
- efectos rebaño dentro de raza 2: 2-1=1
- efectos animal: 15

Otra posibilidad de encontrar soluciones para las MME, es introducir restricciones en el sistema de ecuaciones (ver 4.2.2).

Para hacer evidente la influencia del método de solución sobre la solución misma, se ha hecho con las 12 siguientes g—inversas o restricciones:

1—9: se borraron las filas y columnas correspondientes los siguientes tres efectos cada vez:

1. μ , H1 y H4
2. μ , H3 y H5
3. μ , H2 y H4
4. R1, H1 y H4
5. R1, H3 y H5
6. R1, H2 y H4
7. R2, H1 y H4
8. R2, H3 y H5
9. R2, H2 y H4

10-12: se introdujeron las siguientes restricciones:

10:

- $1/2 R1 + 1/2 R2 = 0$
- $1/3 H1 + 1/3 H2 + 1/3 H3 = 0$
- $1.2 H4 + 1.2 H5 = 0$

11:

- $1/2 R1 + 1/2 R2 = 0$
- $1/3 H1 + 1/3 H2 - 1/3 H3 = 0$
- $1.2 H4 + 1.2 H5 = 0$

12:

- $1/2 R1 + 1/2 R2 = 0$
- $0.3 H1 + 0.3 H2 + 0.4 H3 = 0$
- $1.2 H4 + 1.2 H5 = 0$

Por ultimo tenemos los casos

13:

- solucion mediante el método iterativo de gradientes conjugados preconditionados (PCG)

14:

- solucion mediante la funcion “Matrix Division” o “left-division” de Julia.

Las soluciones y distintas funciones de las soluciones para cada uno de los 11 casos se muestran en los cuadros 6 a 9. De las tablas siguientes se pueden sacar las siguientes consecuencias:

Tablas 6 y 7: las soluciones para los efectos fijos NO SON ESTIMABLES: dependen de qué g—inversa o grupo de restricciones se haya utilizado.

Las soluciones para los coeficientes de regresión (Tabla 6) y para los efectos al azar (Tablas 8 y 9) SI SON ESTIMABLES.

Cuadro 6: Soluciones para los efectos fijos media, edad al parto y raza con las 14 restricciones

μ	b	Raza1	Raza2
0.0	-3.42	6208.64	4193.32
0.0	-3.42	6012.01	3820.2
0.0	-3.42	5765.64	4193.32
6208.64	-3.42	0.0	-2015.32
6012.01	-3.42	0.0	-2191.82
5765.64	-3.42	0.0	-1572.32
4193.32	-3.42	2015.32	0.0
4193.32	-3.42	2015.32	0.0
4193.32	-3.42	1572.32	0.0
5001.09	-3.42	994.34	-994.34
4984.51	-3.42	977.75	-977.75
5001.92	-3.42	995.17	-995.17
1733.32	-3.42	2131.05	1136.72
17369.7	-3.42	48034.3	-642.0

Cuadro 7: Soluciones para los efectos fijos rebaño con las 14 restricciones

H1	H2	H3	H4	H5
0.0	-443.0	-196.63	0.0	-373.12
196.63	-246.37	0.0	373.12	0.0
443.0	0.0	246.37	0.0	-373.12
0.0	-443.0	-196.63	0.0	-373.12
196.63	-246.37	0.0	373.12	0.0
443.0	0.0	246.37	0.0	-373.12
0.0	-443.0	-196.63	0.0	-373.12
0.0	-443.0	-196.63	0.0	-373.12
443.0	0.0	246.37	0.0	-373.12
213.21	-229.79	16.58	186.56	-186.56
246.37	-196.63	49.75	186.56	-186.56
211.55	-231.45	14.92	186.56	-186.56
2344.26	1901.26	2147.64	1323.28	950.16
-59195.4	-59638.4	-59392.0	-12534.4	-12907.5

Cuadro 8: soluciones para los padres (p1...p5) con las 14 restricciones

p1	p2	p3	p4	p5
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44
2.34	-1.35	-0.98	-12.44	12.44

Cuadro 9: soluciones para las vacas (v6...v15) con las 14 restricciones

v6	v7	v8	v9	v10	v11	v12	v13	v14	v15
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17
1.91	-1.42	0.8	-1.96	-3.25	3.93	-16.94	16.94	14.17	-14.17

Tabla 10: las diferencias entre los efectos rebaño dentro de una raza SI SON ESTIMABLES: se pueden comparar rebaños dentro de razas. Sin embargo, las diferencias entre los efectos de dos rebaños que no tienen la misma raza NO SON ESTIMABLES. Para un factor jerarquizado dentro de Otro, solo se pueden comparar los efectos que están jerarquizados dentro de una misma clase del factor superior.

Cuadro 10: Contrastes de efecto rebaño

H1-H2	H1-H3	H1-H4	H4-H5
443.0	196.63	0.0	373.12
443.0	196.63	-176.5	373.12
443.0	196.63	443.0	373.12
443.0	196.63	0.0	373.12
443.0	196.63	-176.5	373.12
443.0	196.63	443.0	373.12
443.0	196.63	0.0	373.12
443.0	196.63	0.0	373.12
443.0	196.63	443.0	373.12
443.0	196.63	26.65	373.12
443.0	196.63	59.81	373.12
443.0	196.63	24.99	373.12
443.0	196.63	1020.98	373.12
443.0	196.63	-46661.0	373.12

Tabla 11: La diferencia entre las razas, en este caso NO ES ESTIMABLE. Normalmente, las diferencias entre efectos de un mismo factor si lo son. En este caso no es así, debido a que dentro de la raza hay otro factor jerarquizado. por eso no se puede separar exactamente los efectos raza y rebaño. Sólo las funciones “media+raza+rebaño” son estimables, pero no lo son las funciones “media+raza” ni la diferencia “(media+R1)— (media+R2)”. Debido a que los distintos efectos “rebaño” no están distribuidos sobre todas las razas, la pregunta de “¿qué raza es mejor?” , no se puede contestar.

Cuadro 11: contrastes implicando media, raza y rebaño

R1-R2	$\mu+R1$	$\mu+R2$	$(\mu+R1$ $-\mu+R2)$	$(\mu+R1+H1)$ $-(\mu+R1+H2)$	$(\mu+R1+H1)$ $-(\mu+R2+H4)$
2015.32	6208.64	4193.32	2015.32	443.0	2015.32
2191.82	6012.01	3820.2	2191.82	443.0	2015.32
1572.32	5765.64	4193.32	1572.32	443.0	2015.32
2015.32	6208.64	4193.32	2015.32	443.0	2015.32
2191.82	6012.01	3820.2	2191.82	443.0	2015.32
1572.32	5765.64	4193.32	1572.32	443.0	2015.32
2015.32	6208.64	4193.32	2015.32	443.0	2015.32
2015.32	6208.64	4193.32	2015.32	443.0	2015.32
1572.32	5765.64	4193.32	1572.32	443.0	2015.32
1988.67	5995.43	4006.76	1988.67	443.0	2015.32
1955.51	5962.26	4006.76	1955.51	443.0	2015.32
1990.33	5997.09	4006.76	1990.33	443.0	2015.32
994.34	3864.37	2870.04	994.34	443.0	2015.32
48676.3	65404.0	16727.7	48676.3	443.0	2015.32

La pregunta “¿cuánta leche da un vaca?” (media de producción de la especie) , tampoco se puede contestar con nuestro modelo (Tabla 3). La “media mínimo—cuadrática” sólo es estimable cuando no se considera ningún Otro factor fijo en el modelo.

El valor aditivo de una vaca es igual a la suma de sus efectos genéticos: raza+efecto propio. Como el efecto raza no es estimable, el valor aditivo de una vaca, incluyendo el efecto raza, TAMPOCO ES ESTIMABLE. Sin embargo, las diferencias entre los valores aditivos de dos vacas de la misma raza SI SON ESTIMABLES (Tabla 12): se pueden comparar las vacas unas con otras dentro de una raza (Y ESTE ES EL OBJETIVO DE LA EVALUACION GENETICA DE LOS ANIMALES!).

Sin embargo, dada la estructura de nuestros datos, la comparación entre valores de dos vacas de distinta raza NO ES ESTIMABLE (Tabla 12). Si algunos rebaños hubieran tenido vacas de las dos razas (estructura no jerárquica) , los efectos “raza” y “rebaño” se hubizan podido separar y hubieran sido estimables las

diferencias entre razas y entre valores aditivos de vacas de dos razas diferentes.

R1+vaca6	R1+vaca7	R2+vaca12	R1+vaca6 - (R2+vaca12)	R1+vaca6 -(R1+vaca7)
6206.73	6210.06	4189.39	2009.48	-0.49
6010.1	6013.43	3816.27	2185.97	-0.49
5763.72	5767.06	4189.39	1566.48	-0.49
-1.91	1.42	-2019.25	2009.48	-0.49
-1.91	1.42	-2195.75	2185.97	-0.49
-1.91	1.42	-1576.25	1566.48	-0.49
2013.41	2016.74	-3.93	2009.48	-0.49
2013.41	2016.74	-3.93	2009.48	-0.49
1570.41	1573.74	-3.93	1566.48	-0.49
992.42	995.76	-998.27	1982.83	-0.49
975.84	979.17	-981.68	1949.66	-0.49
993.25	996.59	-999.1	1984.49	-0.49
2129.14	2132.48	1132.79	988.49	-0.49
48032.4	48035.7	-645.93	48670.5	-0.49

Conclusión: a pesar de las dependencias lineales en las MME se pudo llevar a cabo una comparación entre animales (vacas y padres) dentro de cada raza . Sin embargo la comparación entre raza y entre animales de distinta raza no fue posible. LAS SOLUCIONES DE LAS MME TIENEN QUE SER INTERPRETADAS SIEMPRE BAJO LA CONSIDERACION DE SU ESTIMABILIDAD Y POR LO TANTO DE SU VALIDEZ.

6.3. Apéndice: script Julia para el ejemplo presentado

ejemplo_5 in julia

#Los datos recogidos de la población son los siguientes:

#

#Vaca Raza Padre Edad al primer parto Rebaño kg de leche

#-----

<i>#6</i>	<i>1</i>	<i>1</i>	<i>20</i>	<i>1</i>	<i>6251.1</i>
<i>#7</i>	<i>1</i>	<i>2</i>	<i>40</i>	<i>1</i>	<i>6173.5</i>
<i>#8</i>	<i>1</i>	<i>3</i>	<i>30</i>	<i>2</i>	<i>5775.0</i>
<i>#9</i>	<i>1</i>	<i>2</i>	<i>26</i>	<i>2</i>	<i>5775.6</i>
<i>#10</i>	<i>1</i>	<i>3</i>	<i>38</i>	<i>3</i>	<i>5973.8</i>
<i>#11</i>	<i>1</i>	<i>1</i>	<i>24</i>	<i>3</i>	<i>6050.9</i>
<i>#12</i>	<i>2</i>	<i>4</i>	<i>26</i>	<i>4</i>	<i>4150.6</i>
<i>#13</i>	<i>2</i>	<i>5</i>	<i>32</i>	<i>4</i>	<i>4249.7</i>
<i>#14</i>	<i>2</i>	<i>5</i>	<i>36</i>	<i>5</i>	<i>3849.1</i>
<i>#15</i>	<i>2</i>	<i>4</i>	<i>38</i>	<i>5</i>	<i>3750.3</i>

```

using LinearAlgebra
using SparseArrays

function solve_densem_pcg(A::Array{<:AbstractFloat,2},rhs::Array{<:AbstractFloat,1})
    # solve A sol = rhs by preconditioned conjugate gradient for densem A
    # Preconditioner can be block-diagonal if blksize>1
    m=diag(A)
    m=ifelse.(m .<= 0,0,1 ./m)
    n=length(rhs)
    sol=zeros(n)
    p=zeros(n)
    z=zeros(n)
    w=zeros(n)
    r=zeros(n)
    oldtau=1
    r[:]=rhs # to avoid pass by reference
    for k in 1:1000
        z=m .* r
        tau=z'r
        if (k == 1)
            beta=0
            p=z
        else
            beta=tau/oldtau
            p=z+beta.*p
        end
        w=A*p
        alpha=tau/p'w
        sol=sol + alpha.*p
        if ((k% 100) != 0)
            r=r-alpha.*w
        else
            r=rhs-A*sol
        end
        conv=r'r/rhs'rhs
        println(k," iterations,    convergence criterion=",conv)
        if (conv < 1e-20)
            break
        end
        oldtau = tau
    end
    sol
end

function getA(ped::Array{Int,2})

```

```

# tabular method
nanim=size(ped,1)
A=zeros(nanim,nanim)
for i in 1:nanim
    m=ped[i,3]
    p=ped[i,2]
    A[i,i]=1
    if ((m!=0) & (p!=0))
        A[i,i]=A[i,i]+0.5*A[m,p]
    end
    for j in 1:(i-1)
        #println(i," ",j)
        if m!=0
            A[i,j]=A[i,j]+0.5*A[m,j]
        end
        if p!=0
            A[i,j]=A[i,j]+0.5*A[p,j]
        end
        A[j,i]=A[i,j]
    end
end
A
end

X_mu=[ 1 1 1 1 1 1 1 1 1 1]'
X_raza=[1 1 1 1 1 1 0 0 0 0
        0 0 0 0 0 0 1 1 1 1]'
X_edad=[ 20
        40
        30
        26
        38
        24
        26
        32
        36
        38]
# center the covariate
X_edad = X_edad .- mean(X_edad)
X_rebano=[1 1 0 0 0 0 0 0 0 0
          0 0 1 1 0 0 0 0 0 0
          0 0 0 0 1 1 0 0 0 0
          0 0 0 0 0 0 1 1 0 0
          0 0 0 0 0 0 0 0 1 1]'
y = [6251.1
     6173.5

```

```

5775.0
5775.6
5973.8
6050.9
4150.6
4249.7
3849.1
3750.3]

# we extract a simple pedigree from the data; dams are assumed unknown
ped=
[
1 0 0
2 0 0
3 0 0
4 0 0
5 0 0
6      1 0
7      2 0
8      3 0
9      2 0
10     3 0
11     1 0
12     4 0
13     5 0
14     5 0
15     4 0
]
# Asumimos sigma2u=300000 sigma2e=900000 h2=0.25
sigma2u=300000; sigma2e=900000

# get A-inverse brutally
Ai=inv(getA(ped))

# build MME
# fixed effects
X=[X_mu X_edad X_raza X_rebano]
# random
Z=zeros(10,15)
# 6th cow has 1st record, 2nd cow has 2nd record and so on
for i in 6:15
    rec=i-5
    Z[rec,i]=1
end

# MME

```

```

LHS=[X'X./sigma2e X'Z./sigma2e
      Z'X./sigma2e Z'Z./sigma2e + Ai./sigma2u]
RHS=[X'y./sigma2e
      Z'y./sigma2e]

# show pattern (this is not needed for the computations)
# using Plots
# savefig(Plots.spy(sparse(LHS),markersize=3,legend=nothing),"sparsity.pdf")

# julia library solution
sol=LHS \ RHS
# iterative solution
sol2=solve_densem_pcg(LHS,RHS)

# both solutions are *not* identical in the fixed effects because
# the LHS is not full rank, but solutions for random effects and
# for estimable functions are identical

# Now get estimable functions with restrictions
all_solutions=zeros(24,14)
# our two old solutions
all_solutions[:,13]=sol2
all_solutions[:,14]=sol
# borrado 1 a 9
LHS=[X'X./sigma2e X'Z./sigma2e
      Z'X./sigma2e Z'Z./sigma2e + Ai./sigma2u]
a_borrar=zeros(9,3)
a_borrar[1,:]=[1 5 8]
a_borrar[2,:]=[1 7 9]
a_borrar[3,:]=[1 6 8]
a_borrar[4,:]=[3 5 8]
a_borrar[5,:]=[3 7 9]
a_borrar[6,:]=[3 6 8]
a_borrar[7,:]=[4 5 8]
a_borrar[8,:]=[4 5 8]
a_borrar[9,:]=[4 6 8]
# deleting rows and columns is actually not easy in Julia.
# I found this
#https://stackoverflow.com/questions/58033504/julia-delete-rows-and-columns-from-an-array-or
for i in 1:9
    LHSmodif=LHS[setdiff(1:24, a_borrar[i,:]),setdiff(1:24, a_borrar[i,:])]
    RHSmodif=RHS[setdiff(1:24, a_borrar[i,:])]
    # put solutions back we need to use this setdiff again
    all_solutions[setdiff(1:24, a_borrar[i,:]),i]=LHSmodif \ RHSmodif
end
# other restrictions

```

```
# I (AL) use Henderson's 3.14 to introduce the restriction
```

```
# Restriction 10
```

```
M=[0 0 1/2 1/2 0 0 0 0
    0 0 0 0 1/3 1/3 1/3 0
    0 0 0 0 0 0 0 1/2 1/2]'
```

```
LHS=[X'X./sigma2e+M*M' X'Z./sigma2e
      Z'X./sigma2e Z'Z./sigma2e + Ai./sigma2u]
```

```
all_solutions[:,10]=LHS \ RHS
```

```
# Restriction 11
```

```
M=[0 0 1/2 1/2 0 0 0 0
    0 0 0 0 1/3 1/3 -1/3 0
    0 0 0 0 0 0 0 1/2 1/2]'
```

```
LHS=[X'X./sigma2e+M*M' X'Z./sigma2e
      Z'X./sigma2e Z'Z./sigma2e + Ai./sigma2u]
```

```
all_solutions[:,11]=LHS \ RHS
```

```
# Restriction 12
```

```
M=[0 0 1/2 1/2 0 0 0 0
    0 0 0 0 0.3 0.3 0.4 0
    0 0 0 0 0 0 0 1/2 1/2]'
```

```
LHS=[X'X./sigma2e+M*M' X'Z./sigma2e
      Z'X./sigma2e Z'Z./sigma2e + Ai./sigma2u]
```

```
all_solutions[:,12]=LHS \ RHS
```

```
# soluciones efectos fijos
```

```
display(round.(all_solutions[1:9,:]',digits=2))
```

```
# aleatorios
```

```
display(round.(all_solutions[10:24,:]',digits=2))
```

```
# contrastes
```

```
# H1-H2 H1-H3 H1-H4 H4-H5
```

```
solH=all_solutions[5:9,:]
```

```
contrastH=zeros(4,14)
```

```
contrastH[1,:]=solH[1,:]-solH[2,:]
```

```
contrastH[2,:]=solH[1,:]-solH[3,:]
```

```
contrastH[3,:]=solH[1,:]-solH[4,:]
```

```
contrastH[4,:]=solH[4,:]-solH[5,:]
```

```
display(round.(contrastH',digits=2))
```

```
# R1-R2  $\mu_{R1}$   $\mu_{R2}$  ( $\mu_{R1}-\mu_{R2}$ ) ( $\mu_{R1+H1}-\mu_{R1+H2}$ ) ( $\mu_{R1+H1}-\mu_{R1+H2}$ )
```

```
contrastH=zeros(6,14)
```

```
contrastH[1,:]=all_solutions[3,:]-all_solutions[4,:]
```

```
contrastH[2,:]=all_solutions[1,:]+all_solutions[3,:]
```

```
contrastH[3,:]=all_solutions[1,:]+all_solutions[4,:]
```

```
contrastH[4,:]=all_solutions[1,:]+all_solutions[3,:]
```

```
-(all_solutions[1,:]+all_solutions[4,:])
```

```
contrastH[5,:]=all_solutions[1,:]+all_solutions[3,:]+all_solutions[5,:]
```



```

        -(all_solutions[1,:]+all_solutions[3,:]+all_solutions[6,:])
contrastH[6,:]=all_solutions[1,:]+all_solutions[3,:]+all_solutions[5,:]-
        -(all_solutions[1,:]+all_solutions[4,:]+all_solutions[8,:])
display(round.(contrastH',digits=2))

# R1+vaca6 R1+vaca7 R2+vaca12 R1+vaca6-(R2+vaca12) R1+vaca6-(R1+vaca7)
contrastH=zeros(5,14)
contrastH[1,:]=all_solutions[3,:]-all_solutions[15,:]
contrastH[2,:]=all_solutions[3,:]-all_solutions[16,:]
contrastH[3,:]=all_solutions[4,:]-all_solutions[20,:]
contrastH[4,:]=all_solutions[3,:]-all_solutions[15,:]-
        all_solutions[4,:]-all_solutions[20,:]
contrastH[5,:]=all_solutions[3,:]-all_solutions[15,:]-
        all_solutions[3,:]-all_solutions[16,:]
display(round.(contrastH',digits=2))

```