Capítulo 2

Álgebra de matrices

2.1. Operaciones con matrices

La forma de denotar los elementos de una matriz ya la introdujimos en (1.11). Una matriz de $m \times n$ (m filas y n columnas) es de la forma

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

En general, denotamos por a_{ij} al elemento que se encuentra en la fila i y columna j de la matriz. Denotaremos con $\mathscr{M}_{m\times n}$ al conjunto de todas las matrices de $m \times n$. Los elementos diagonales son a_{11}, $a_{22}, \ldots$, y forman la *diagonal principal* de la matriz.

Ejemplo 2.1. Una *matriz diagonal* es una matriz cuadrada (de $n \times n$) con $a_{ij} = 0$ si $i \neq j$, es decir, los únicos elementos no nulos son los de la diagonal principal:

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Un ejemplo de matriz diagonal es la matriz identidad I_n

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.1}$$

es decir, la matriz $n \times n$ compuesta de unos en la diagonal principal y ceros en el resto de elementos.

La matriz 0 es aquella cuyos elementos son todos nulos $\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$.

Consideremos de nuevo la **interpretación de la estructura de una matriz como un conjunto de vectores columna ordenados** (ver (1.11)) Usando las propiedades de las operaciones (1.8–1.9) entre vectores columna se pueden definir análogamente las operaciones elementales entre matrices:

- La multiplicación por un escalar c se hace multiplicando todos los vectores columna, es decir todos los elementos uno a uno:

$$c \cdot \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \Big[c\,\mathbf{a}_1 \,\Big|\, c\,\mathbf{a}_2 \,\Big|\, \cdots \,\Big|\, c\,\mathbf{a}_n \Big] = \begin{bmatrix} c\,a_{11} & c\,a_{12} & \cdots & c\,a_{1n} \\ c\,a_{21} & c\,a_{22} & \cdots & c\,a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c\,a_{m1} & c\,a_{m2} & \cdots & c\,a_{mn} \end{bmatrix}$$

- La suma se realiza columna a columna, es decir elemento a elemento

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

$$= \Big[\mathbf{a}_1 \,\Big|\, \mathbf{a}_2 \,\Big|\, \cdots \,\Big|\, \mathbf{a}_n \Big] + \Big[\mathbf{b}_1 \,\Big|\, \mathbf{b}_2 \,\Big|\, \cdots \,\Big|\, \mathbf{b}_n \Big] = \Big[\mathbf{a}_1 + \mathbf{b}_1 \,\Big|\, \mathbf{a}_2 + \mathbf{b}_2 \,\Big|\, \cdots \,\Big|\, \mathbf{a}_n + \mathbf{b}_n \Big] =$$

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

Ciertas propiedades de los vectores columna (Proposición 1.18) se generalizan directamente a matrices, como sigue.

Teorema 2.2 (Propiedades algebraicas de las matrices.).

Para todas las matrices A, B, C de $m \times n$ y todos los escalares $r, s \in \mathbf{R}$:

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

Multiplicación de matrices. Definimos ahora el producto de matrices, utilizando de nuevo la interpretación columnar de una matriz. Conocemos ya la operación de multiplicar una matriz A por un vector $\mathbf{x}$ (definición 1.23) Vamos a estudiar qué sucede si multiplicamos consecutivamente un vector por dos matrices:

$$A \cdot (B \cdot \mathbf{x})$$

Es decir, primero multiplicamos a $\mathbf{x}$ por B para obtener un vector $B \cdot \mathbf{x}$, y a este vector resultante lo multiplicamos por A.

El primer producto $B\mathbf{x}$ (es habitual, como entre números reales, omitir el símbolo de multiplicación) se puede expresar por columnas como

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \Rightarrow \quad B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$$

si $\mathbf{b}_1, \ldots, \mathbf{b}_n$ son las columnas de B. Multiplicando por la matriz A

$$A(B\mathbf{x}) = A(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = x_1 A\mathbf{b}_1 + \cdots + x_n A\mathbf{b}_n$$

por las propiedades del producto matriz-vector. Convirtiendo esta expresión vectorial de nuevo a notación matricial

$$A(B\mathbf{x}) = \left[\, A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n \,\right] \mathbf{x} \equiv C\mathbf{x}$$

¡ Resulta que el vector obtenido es el mismo que se obtiene multiplicando $\mathbf{x}$ por una matriz $C = \left[\, A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n \,\right]$! Definimos esta matriz como el producto $C = AB$ de las matrices A y B.

Definición 2.3. *Sea A una matriz $m \times p$ y B una matriz $p \times n$, cuyas columnas son $\mathbf{b}_1$, $\ldots$, $\mathbf{b}_n$. El producto AB se define como la matriz de $m \times n$ de columnas $A\mathbf{b}_1, \ldots, A\mathbf{b}_n$:*

$$AB = A\left[\, \mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_n \,\right] = \left[\, A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n \,\right] \tag{2.2}$$

Es importante darse cuenta de que no se pueden multiplicar dos matrices cualesquiera. El número de columnas de la matriz de la izquierda A debe ser igual al número de columnas de la matriz de la derecha B.

Ejemplo 2.4. Calculad AB si

$$A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$$

Respuesta: por columnas

$$\begin{aligned} AB &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} \\ &= \left[\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} \,\middle|\, \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} \,\middle|\, \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix} \right] = \left[\begin{array}{r|r|r} 11 & 0 & 21 \\ 1 & 13 & -9 \end{array}\right] \end{aligned}$$

El cálculo se puede organizar directamente mediante la llamada regla fila-columna:

$$\begin{aligned} AB &= \begin{bmatrix} 2\times 4 + 3\times 1 & 2\times 3 + 3\times(-2) & 2\times 6 + 3\times 3 \\ 1\times 4 + (-5)\times 1 & 1\times 4 + (-5)\times(-2) & 1\times 6 + (-5)\times 3 \end{bmatrix} \\ &= \begin{bmatrix} 11 & 0 & 21 \\ 1 & 13 & -9 \end{bmatrix} \end{aligned}$$

Al menos hay tres maneras de organizar e interpretar el producto de dos matrices A de $m \times p$ y B de $p \times n$:

1. La **multiplicación por columnas** es la que hemos usado para definir el producto: *la columna i-ésima de AB es una combinación lineal de las columnas de A, con pesos dados por la columna i-ésima de B.* Es decir, calculamos las columnas $A\mathbf{b}_i$ de $\left[A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n \right]$ por columnas:

$$AB = \left[b_{11}\mathbf{a}_1 + b_{21}\mathbf{a}_2 + \cdots + b_{p1}\mathbf{a}_p \mid \cdots \mid b_{1n}\mathbf{a}_1 + b_{2n}\mathbf{a}_2 + \cdots + b_{pn}\mathbf{a}_p \right]$$

2. Es bastante más eficiente calcular el producto usando la **regla fila-columna** para calcular cada elemento de la matriz producto: La regla fila-columna se puede usar para productos de matrices igual que se usó para el producto de una matriz por un vector columna. Para ello, basta usarla para calcular los productos $A\mathbf{b}_i$ de $\left[A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n \right]$.

3. Finalmente, existe la **multiplicación por filas**. Se puede observar que la fila i-ésima de la matriz producto es una combinación lineal de las filas de B, con pesos dados por los elementos de la fila i-ésima de A:

$$AB = \left[\begin{array}{c} a_{11}(\text{fila 1 de } B) + a_{12}(\text{fila 2 de } B) + \cdots + a_{1p}(\text{fila } p \text{ de } B) \\ \hline a_{21}(\text{fila 1 de } B) + a_{22}(\text{fila 2 de } B) + \cdots + a_{2p}(\text{fila } p \text{ de } B) \\ \hline \vdots \\ \hline a_{m1}(\text{fila 1 de } B) + a_{m2}(\text{fila 2 de } B) + \cdots + a_{mp}(\text{fila } p \text{ de } B) \end{array}\right] \tag{2.3}$$

Para más información, consultad el teorema 2.34 de la sección 2.5 sobre matrices partidas. En todo caso, el elemento (i,j) (fila i, columna j) de la matriz producto AB es:

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Teorema 2.5. *Dadas una matriz A de $m \times n$ y B y C tales que las sumas y productos siguientes se pueden realizar, tenemos las siguientes propiedades.*

1. $A(BC) = (AB)C$ (*propiedad asociativa*)
2. $A(B+C) = AB + AC$ (*propiedad distributiva*)
3. $(A+B)C = AC + AB$ (*propiedad distributiva*)
4. $c(AB) = (cA)B = A(cB)$ ($c \in \mathbf{R}$ *escalar*)
5. $I_m A = A = A I_n$

Demostración. Todas las propiedades son fáciles de demostrar con la definición por columnas. Consideremos el producto $A(BC)$. Sobre una columna $\mathbf{c}_i$ de C, la definición (2.2) es en realidad haber exigido asociatividad al producto matricial $AB\mathbf{x}$:

$$A(B\mathbf{c}_i) = (AB)\mathbf{c}_i$$

por tanto tenemos que $A(BC) = A\left[\, B\mathbf{c}_1 \mid B\mathbf{c}_2 \mid \cdots \mid B\mathbf{c}_p \,\right] = \left[\, AB\mathbf{c}_1 \mid AB\mathbf{c}_2 \mid \cdots \mid AB\mathbf{c}_p \,\right]$ $= AB\left[\, \mathbf{c}_1 \mid \mathbf{c}_2 \mid \cdots \mid \mathbf{c}_p \,\right] = (AB)C$. Las demás se demuestran análogamente. □

Ejercicio 2.6. Comprobad que las matrices $\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ y $\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$ *no* conmutan (es decir, $AB \neq BA$). Comparad las multiplicaciones por filas y por columnas en ambos productos.

Proposición 2.7. *En general:*

1. *$AB \neq BA$,*
2. *$AB = AC$ no implica que $B = C$, incluso si $A \neq 0$,*
3. *$AB = 0$, no implica que $A = 0$ ó $B = 0$ (es decir, hay $A \neq 0$, $B \neq 0$ tales que $AB = 0$)*

La transpuesta de una matriz.

Definición 2.8. *Si A es una matriz $m \times n$, la matriz **transpuesta** A^T es una matriz $n \times m$ cuyas columnas son las filas de A.*

Ejercicio 2.9. Calculad las transpuestas de

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}.$$

Teorema 2.10. *Si A y B son matrices tales que las operaciones siguientes se pueden realizar, tenemos que*

1. *$(A^T)^T = A$,*
2. *$(A + B)^T = A^T + B^T$,*
3. *$(cA)^T = cA^T$ para todo escalar $c \in \mathbf{R}$,*
4. *$(AB)^T = B^T A^T$.*

Demostración. Sólo la propiedad 4 no es obvia, aunque se demuestra expresando una fila genérica de $(AB)^T$. □

Ejercicio 2.11. Calculad la fórmula de $(ABC)^T$ en términos de A^T, B^T y C^T.

2.2. Inversa de una matriz

Definición 2.12. *Una matriz A de $n \times n$ es **invertible** si existe una matriz C de $n \times n$ tal que*

$$CA = I \quad y \quad AC = I.$$

Cuando existe C, se denota por $A^{-1} = C$ y se denomina matriz inversa *de A.*

Obsérvese que la definición requiere que la matriz A sea *cuadrada* (número de filas igual a número de columnas) De una matriz no invertible A se dice que es **singular**.

Proposición 2.13. *Si una matriz A tiene una inversa A^{-1}, entonces esta matriz es única.*

Demostración. Imaginemos que C y D son dos inversas de A: entonces $AC = I$, $CA = I$, $AD = I$ y $DA = I$. Multiplicando la primera por D a la izquierda $DAC = D$ y como $DA = I$, entonces $C = D$. □

Por ello se puede hablar de *la* inversa de una matriz dada.

Ejercicio 2.14. Demostrad que si una matriz tiene una inversa por la izquierda y otra por la derecha, entonces ambas inversas son iguales. Es decir, si dada A existen C y D tales que $CA = I$ y $AD = I$.

Ejemplo 2.15. Comprobad que la inversa de $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ es $A^{-1} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$.

¿ Cómo se calcula la inversa ? Con matrices 2×2 es muy sencillo.

Proposición 2.16 (Inversa de matrices 2×2). *Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Si $ad - bc \neq 0$ entonces A es invertible y*

$$A^{-1} = \frac{1}{ab - cd} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Si $ad - bc = 0$ entonces A es singular.

La cantidad $ad - bc$ se denomina **determinante** de A, denotándose

$$\det A = ad - bc \tag{2.4}$$

Por tanto, según la proposición, una matriz de 2×2 es singular si $\det A = 0$, y es invertible si $\det A \neq 0$, en cuyo caso la inversa se construye dividiendo por el determinante la matriz resultante de intercambiar los elementos de la diagonal principal, y cambiar de signo los elementos de la otra diagonal.

Teorema 2.17. *Si A es invertible, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución, y es única, para todo $\mathbf{b}$. Esta solución es $\mathbf{x} = A^{-1}\mathbf{b}$.*

Demostración. Multipliquemos la ecuación por la izquierda por A^{-1}:

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \Rightarrow \mathbf{x} = A^{-1}\mathbf{b}.$$

Este cálculo implica que si $\mathbf{x}$ es solución del sistema, obligatoriamente ha de ser igual a $A^{-1}\mathbf{b}$ (unicidad de la solución) Multiplicando esta única candidata a solución por A tenemos que $A(A^{-1}\mathbf{b}) = \mathbf{b}$, luego es, efectivamente, una solución. □

Ejercicio 2.18. Demostrad que el teorema 2.17 implica que una matriz A invertible satisface tanto el teorema S como el teorema I del capítulo 1.

Se puede pensar que la extracción de la inversa es algo así como "dividir" por una matriz. Es decir, dada la ecuación $A\mathbf{x} = \mathbf{b}$ se puede, si A es invertible, despejar $\mathbf{x}$ de la forma $\mathbf{x} = A^{-1}\mathbf{b}$. Sin embargo, este punto de vista es muy peligroso: no todas las matrices pueden "dividir", sólo las (cuadradas) invertibles y además dividir ... ¿ por la izquierda o por la derecha ? La falta de conmutatividad complica esta visión, y es mejor pensar escrictamente en multiplicaciones por inversas, a la izquierda o la derecha según convenga, y abstenerse de hablar de división alguna.

Ejemplo 2.19. Resolved el sistema

$$\begin{aligned} -7x_1 - 5x_2 &= b_1 \\ 3x_1 + 2x_2 &= b_2 \end{aligned}$$

en términos de b_1 y b_2.

Teorema 2.20 (Propiedades de la inversa)**.**

1. *Si A es invertible, entonces A^{-1} es invertible con*
$$(A^{-1})^{-1} = A.$$
2. *Si A y B son matrices invertibles del mismo tamaño, su producto AB también lo es y*
$$(AB)^{-1} = B^{-1}A^{-1}.$$
3. *Si A es invertible, entonces A^T es invertible con*
$$(A^T)^{-1} = (A^{-1})^T.$$

Demostración. Las ecuaciónes que definen A^{-1} como inverso de A, $A^{-1}A = AA^{-1} = I$, también definen a A como inverso de A^{-1}. Además $B^{-1}A^{-1}AB = B^{-1}B = I = ABB^{-1}A^{-1}$ y que $(A^{-1})^TA^T = (AA^{-1})^T = I^T = I = A^T(A^{-1})^T = (A^{-1}A)^T$. □

Proposición 2.21. *El producto de varias matrices invertibles $A_1A_2\cdots A_p$ es invertible, siendo su inversa el producto de las inversas en el orden opuesto: $A_p^{-1}A_{p-1}^{-1}\cdots A_1^{-1}$.*

Vamos a estudiar las propiedades de las matrices invertibles. Este estudio, que comenzamos aquí pero se extenderá durante varios temas, dará lugar a un enorme teorema que reflejará muchísimas propiedades equivalentes a la invertibilidad. Se trata del **teorema de la matriz invertible** que apodaremos "*Teorema B*" y cuya primera aparición es el ejercicio 2.18 y el siguiente teorema.

Teorema 2.22. *Una matriz A de $n \times n$ es invertible si y sólo si es equivalente por filas a la matriz identidad I_n. Cualquier sucesión de operaciones elementales de fila que reduce a A a la identidad, aplicada sobre la identidad, produce A^{-1}.*

Demostración. El ejercicio 2.18 asegura que hay una posición pivote en cada columna y en cada fila de A. Pero A es cuadrada, así que la matriz reducida equivalente ha de ser la identidad I_n. Hemos demostrado entonces que "invertible $\Rightarrow$ equivalente a I".

Hay que demostrar ahora la implicación contraria, el "sólo si", es decir "equivalente a $I \Rightarrow$ invertible". Si una matriz es equivalente a I, es que existe una secuencia de operaciones elementales de fila que la transforman en la identidad. Cada operación elemental realiza combinaciones lineales de filas, y una secuencia de operaciones elementales produce una matriz cuyas filas son ciertas combinaciones lineales de las filas de la matriz inicial. Observando la interpretación por filas (2.3) del producto de matrices, concluimos que la matriz final, que en nuestro caso es I, se puede escribir como el producto de una matriz E (de $n \times n$) por la izquierda por la matriz A:

$$EA = I$$

Evidentemente, E es una inversa por la izquierda de A. Como A es equivalente a la identidad, los n sistemas de ecuaciones simultáneos $A\mathbf{x}_1 = \mathbf{e}_1$, $A\mathbf{x}_2 = \mathbf{e}_2$, $\dots$, $A\mathbf{x}_n = \mathbf{e}_n$, tienen soluciones $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ (por el teorema S) únicas (ejercicio 1.46) Esas soluciones son las columnas de una matriz X que es inversa por la derecha de A, porque $AX = I$. El teorema **??** demuestra que $E = X = A^{-1}$ □

Ejercicio 2.23. Sea A de $n \times n$. Demostrad el recíproco del teorema 2.17, es decir, muéstrese que si el sistema $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución (es consistente) para todo $\mathbf{b} \in \mathbf{R}^n$, entonces A debe ser invertible. Sugerencia: considerad si A es equivalente por filas a I_n, usando el teorema S.

Algoritmo para calcular A^{-1}. El teorema 2.22 permite formular un algoritmo sencillo para hallar la inversa, basado en operaciones elementales de filas.

Dada una matriz A de $n \times n$,

1. formar la matriz aumentada $\left[\, A \mid I \,\right]$,
2. reducir por filas esa matriz aumentada hasta la forma escalonada reducida,
3. si A es invertible, la matriz aumentada reducida será de la forma $\left[\, I \mid A^{-1} \,\right]$.

Las operaciones realizadas sobre A para transformarla en I, se están realizando simultáneamente sobre I. Por el teorema 2.22 el resultado de estas operaciones sobre I es la inversa A^{-1}.

Ejemplo 2.24. Calcular la inversa, si existe, de $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$.

$$\begin{aligned} \left[\, A \mid I \,\right] &= \left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & -3 & 8 & 0 & 0 & 1 \end{array}\right] \\ &\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array}\right] \end{aligned}$$

Podemos observar este procedimiento como si estuviéramos resolviendo n sistemas de ecuaciones simultáneos $A\mathbf{x} = \mathbf{e}_1$, $A\mathbf{x} = \mathbf{e}_2$, ..., $A\mathbf{x} = \mathbf{e}_n$, que resolvemos de golpe aumentando la matriz A con todos los lados derechos $\mathbf{e}_i$ a la vez:

$$\left[\, A \mid \mathbf{e}_1 \,\right], \ldots, \left[\, A \mid \mathbf{e}_n \,\right] \quad \Leftrightarrow \quad \left[\, A \mid I \,\right]. \tag{2.5}$$

Ejemplo 2.25. Este procedimiento se puede generalizar para la situación en la que tenemos p sistemas de ecuaciones que resolver $A\mathbf{x} = \mathbf{b}_1, \ldots,$ $A\mathbf{x} = \mathbf{b}_p$, que podemos resolver poniendo todos los lados derechos $\mathbf{b}_i$ como columnas en una matriz $B = \left[\, \mathbf{b}_1 \mid \cdots \mid \mathbf{b}_p \,\right]$ y reduciendo

$$\left[\, A \mid B \,\right] \sim \left[\, I \mid C \,\right]. \tag{2.6}$$

Las soluciones son las columnas de la matriz

$$C = A^{-1}B = \left[\, A^{-1}\mathbf{b}_1 \mid \cdots \mid A^{-1}\mathbf{b}_p \,\right].$$

Es interesante interpretar todavía de otra manera el algoritmo para calcular la inversa. Resulta el producto de una matriz A^{-1} por una matriz dividida en dos submatrices como $\left[\, A \mid I \,\right]$ se puede organizar *por bloques*

$$A^{-1}\left[\, A \mid I \,\right] = \left[\, A^{-1}A \mid A^{-1}I \,\right] = \left[\, I \mid A^{-1} \,\right] \tag{2.7}$$

como se deduce fácilmente de la definición de multiplicación de matrices (2.2). Este es un ejemplo importante de la multiplicación de matrices por bloques, que se discute en el párrafo optativo 2.5. Otro ejemplo es el de la solución simultánea de ecuaciones (2.6)

$$A^{-1}\left[\, A \mid B \,\right] = \left[\, I \mid A^{-1}B \,\right] = \left[\, I \mid A^{-1}\mathbf{b}_1 \mid \cdots \mid A^{-1}\mathbf{b}_n \,\right] \tag{2.8}$$

2.3. Caracterizaciones de matrices invertibles

Teorema 2.26 (Teorema de la matriz invertible)**.** *Sea A una matriz $n\times n$ (cuadrada). Las siguientes afirmaciones son equivalentes.*

a. *A es invertible.*

b. *A es equivalente por filas a la matriz identidad I_n.*

c. *A tiene n posiciones pivote (es decir, n pivotes no nulos).*

d. *La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ admite sólo la solución trivial.*

e. *Las columnas de A forman un conjunto linealmente independiente.*

f. *

g. *La ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para todo $\mathbf{b} \in \mathbf{R}^n$.*

h. *Las columnas de A generan $\mathbf{R}^n$.*

i.

j. *Existe una matriz C $n\times n$ tal que $CA = I$.*

k. *Existe una matriz D $n\times n$ tal que $AD = I$.*

l. *A^T es invertible.*

*dejamos algunos apartados vacíos para concidir con la notación de Lay [1].

Demostración. Para demostrar un teorema que consiste en la equivalencia de tantas proposiciones, lo más eficiente sería crear una cadena de implicaciones lógicas, como por ejemplo $a. \Rightarrow b. \Rightarrow \cdots \Rightarrow l. \Rightarrow a$:. Para adaptarnos a las particularidades de cada afirmación, realizaremos varias subcadenas lógicas que acabarán conectando todas las afirmaciones realizadas.

Comencemos por equivalencias ya conocidas. Por el teorema 2.22, $a. \Leftrightarrow b.$ y su demostración implica que $a. \Rightarrow c.$. Pero $c. \Rightarrow b.$ porque n posiciones pivote en una matriz $n \times n$ implican que la matriz escalonada reducida es la indentidad I_n. El teorema 2.17 implica que $a. \Leftrightarrow d.$ y $a. \Leftrightarrow g.$ por el ejercicio 2.23 y $g. \Leftrightarrow h.$ por definición. Además $g. \Rightarrow k.$ y $k. \Rightarrow g.$ por el ejercicio 2.27. Por definición, y la definición de producto matriz-vector, $d. \Leftrightarrow e.$. El teorema 2.20 3. implica que $a. \Leftrightarrow l.$ y ésto implica que $j.$ y $k.$ son equivalentes. □

Ejercicio 2.27. Demuéstrese que si existe una matriz D $n \times n$ tal que $AD = I$ entonces $A\mathbf{x} = \mathbf{b}$ tiene siempre al menos una solución. Sugerencia: aplíquese esta ecuación matricial a $\mathbf{b}$.

2.4. Matrices elementales*

Matrices elementales. Vamos a representar las operaciones elementales por fila mediante multiplicaciones por la izquierda de ciertas matrices especiales, denominadas matrices elementales. Con su ayuda, demostraremas que toda matriz invertible es equivalente por filas a la matriz identidad, y encontraremos un algoritmo para calcular la inversa.

La siguiente proposición utiliza un hecho que es la clave de la aplicabilidad del Álgebra Lineal en Ciencia, y especialmente en Ingeniería. Una matriz es un **operador**, es decir, un objeto que actúa sobre los vectores, realizando sobre ellos alguna operación concreta de interés. Puede actuar, como veremos, como una operación de clara interpretación geométrica, un giro, una proyección, etc. O como una operación elemental por filas, como es el caso aquí. Su manera de actuar es multiplicándose por la izquierda por el vector. Si una matriz actúa sobre vectores, puede actúar también sobre matrices, actuando simultáneamente sobre cada vector columna. Y, otra vez, una matriz actúa multiplicando por la izquierda al vector o matriz sobre el que actúa.

Proposición 2.28. *Realizar una operacion elemental sobre una matriz A de $m \times n$ es equivalente a multiplicarla por la izquierda por una matriz E de $m \times m$, siendo la matriz obtenida EA. La matriz E es precisamente la que se obtiene al realizar esa misma operación elemental sobre la matriz identidad I_m.*

Demostración. La interpretación del producto por filas (1.12) (v. también (2.9)) de $E \cdot I$, junto con la igualdad $E \cdot A = E \cdot I \cdot A$, demuestra la proposición. □

Definición 2.29. *Una* ***matriz elemental*** *es la obtenida al realizar una operación elemental por filas a la matriz identidad.*

Como hay tres tipos de operaciones elementales por filas, habrá tres tipos de matrices elementales. En el siguiente ejemplo los ilustramos.

Ejemplo 2.30. Permutación: $P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

multiplicación por constante de una fila: $L_2(3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

suma del mútliplo de una fila a otra: $L_{31}(4) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$.

Es interesante la siguiente observación: toda matriz elemental es invertible. Ello es consecuencia de la invertibilidad de las operaciones elementales, como averiguamos cuando las definimos.

Ejercicio 2.31. Compruébese que las inversas de las matrices elementales del ejemplo 2.30 son:

$$P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow P_{12}^{-1} = P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$L_2(3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow L_2(3)^{-1} = L_2(1/3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$L_{31}(4) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \Rightarrow L_{31}(4)^{-1} = L_{31}(-4) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$$

A partir de ahora podemos visualizar las operaciones elementales sobre una matriz como multiplicaciones por la izquierda por matrices elementales apropiadas. Este punto de vista computacional nos facilitará llegar a varios resultados prácticos y teóricos de gran importancia. Escribiremos, por tanto

$$E_p E_{p-1} \cdots E_2 E_1 A = U \quad \Leftrightarrow \quad EA = U$$

siendo U la matriz reducida y donde las E_i codifican las operaciones elementales individuales, y E es el producto de todas las matrices elementales correspondientes.

2.5. Matrices partidas*

Es interesante conocer que hay distintas formas de organizar el cálculo del producto de dos matrices. En realidad ya hemos visto tres maneras, la multiplicación por columnas, por filas y el uso de la regla fila-columna. Todo ello está muy relacionado con el hecho de que la multiplicación se puede realizar **por bloques** creados dentro de las matrices, pudiéndose ganar en eficiencia, claridad o incluso facilitar la explicación de argumentos teóricos. La técnica básica es la definición de una partición de una matri en subbloques adecuados. Comencemos con un ejemplo.

Ejemplo 2.32. La matriz siguiente se puede partir en bloques

$$A = \left[\begin{array}{rrr|rr|r} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

con

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$
$$A_{12} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} -4 \end{bmatrix}.$$

La observación importante es que las matrices partidas se pueden multiplicar, si todos los tamaños y la disposición de los bloques es compatible, usando la regla fila-columna (o cualquier otra) considerando como si los bloques fueran elementos.

Ejemplo 2.33.

$$A = \left[\begin{array}{rrr|rr} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

$$B = \left[\begin{array}{rr} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

Resulta que

$$
\begin{aligned}
AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} + \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} \\ \begin{bmatrix} 0 & -4 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} + \begin{bmatrix} 7 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} \\ \begin{bmatrix} 14 & -18 \end{bmatrix} + \begin{bmatrix} -12 & 19 \end{bmatrix} \end{bmatrix} = \left[\begin{array}{cc} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right].
\end{aligned}
$$

Hay que señalar que las operaciones de bloques $A_{11}B_1+A_{12}B_2$, etc. son operaciones de matrices, y las multiplicaciones deben ser realizadas en el orden escrito.

Utilizando matrices partidas, podemos dar una formulación sucinta de las tres reglas de multiplicación de matrices. Para ello observamos que dos particiones particulares son la partición en columnas y en filas:

$$
\begin{aligned}
A &= \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right] = \left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \hline & & \vdots & \\ \hline a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right] \\
&= \begin{bmatrix} \operatorname{col}_1(A) & \operatorname{col}_2(A) & \cdots & \operatorname{col}_n(A) \end{bmatrix} = \begin{bmatrix} \operatorname{fil}_1(A) \\ \operatorname{fil}_2(A) \\ \vdots \\ \operatorname{fil}_m(A) \end{bmatrix}.
\end{aligned}
$$

Teorema 2.34. *Si A es $m \times p$ y B es $p \times n$, entonces*

1. ***Regla fila-columna***

$$
\begin{aligned}
A\cdot B &= \begin{bmatrix} \operatorname{fil}_1(A) \\ \operatorname{fil}_2(A) \\ \vdots \\ \operatorname{fil}_m(A) \end{bmatrix} \cdot \begin{bmatrix} \operatorname{col}_1(B) & \operatorname{col}_2(B) & \cdots & \operatorname{col}_n(B) \end{bmatrix} \\
&= \begin{bmatrix} \operatorname{fil}_1(A)\cdot\operatorname{col}_1(B) & \operatorname{fil}_1(A)\cdot\operatorname{col}_2(B) & \cdots & \operatorname{fil}_1(A)\cdot\operatorname{col}_n(B) \\ \operatorname{fil}_2(A)\cdot\operatorname{col}_1(B) & \operatorname{fil}_2(A)\cdot\operatorname{col}_2(B) & \cdots & \operatorname{fil}_2(A)\cdot\operatorname{col}_n(B) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{fil}_m(A)\cdot\operatorname{col}_1(B) & \operatorname{fil}_m(A)\cdot\operatorname{col}_2(B) & \cdots & \operatorname{fil}_m(A)\cdot\operatorname{col}_n(B) \end{bmatrix}
\end{aligned}
$$

2. ***Multiplicación por columnas***

$$
\begin{aligned}
A\cdot B &= \begin{bmatrix} \operatorname{col}_1(A) & \operatorname{col}_2(A) & \cdots & \operatorname{col}_p(A) \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{bmatrix} \\
&= \left[\begin{array}{c|c|c|c} \operatorname{col}_1(A)b_{11} & \operatorname{col}_1(A)b_{12} & & \operatorname{col}_1(A)b_{1n} \\ +\operatorname{col}_2(A)b_{21} & +\operatorname{col}_2(A)b_{22} & \cdots & +\operatorname{col}_2(A)b_{2n} \\ \vdots & \vdots & & \vdots \\ +\operatorname{col}_p(A)b_{p1} & +\operatorname{col}_p(A)b_{p2} & & +\operatorname{col}_p(A)b_{pn} \end{array}\right]
\end{aligned}
$$

3. ***Multiplicación por filas***

$$
\begin{aligned}
A\cdot B &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \cdot \begin{bmatrix} \operatorname{fil}_1(B) \\ \operatorname{fil}_2(B) \\ \vdots \\ \operatorname{fil}_p(B) \end{bmatrix} \\
&= \begin{bmatrix} a_{11}\operatorname{fil}_1(B) + a_{12}\operatorname{fil}_2(B) + \cdots a_{1p}\operatorname{fil}_p(B) \\ a_{21}\operatorname{fil}_1(B) + a_{22}\operatorname{fil}_2(B) + \cdots a_{2p}\operatorname{fil}_p(B) \\ \vdots \\ a_{m1}\operatorname{fil}_1(B) + a_{m2}\operatorname{fil}_2(B) + \cdots a_{mp}\operatorname{fil}_p(B) \end{bmatrix}
\end{aligned} \tag{2.9}
$$

4. ***Regla columna-fila***

$$
\begin{aligned}
A\cdot B &= \begin{bmatrix}\mathrm{col}_1(A) & \mathrm{col}_2(A) & \cdots & \mathrm{col}_p(A)\end{bmatrix}\cdot\begin{bmatrix}\mathrm{fil}_1(B)\\ \mathrm{fil}_2(B)\\ \vdots\\ \mathrm{fil}_p(B)\end{bmatrix}\\
&= \mathrm{col}_1(A)\cdot\mathrm{fil}_1(B)+\mathrm{col}_2(A)\cdot\mathrm{fil}_2(B)+\cdots+\mathrm{col}_p(A)\cdot\mathrm{fil}_p(B)
\end{aligned}
$$

Finalmente observemos que la definición de multiplicación de dos matrices (2.2) es interpretable como una multiplicación de matrices por bloques, como ya hicimos en (2.7) y (2.8).

2.6. Factorización LU *

En esta sección vamos a encontrar una *factorización* de una matriz A, es decir, vamos a escribir A como el producto de dos factores

$$A = L\cdot U$$

siendo esos factores matrices L y U con características interesantes. La primera es que U es una matriz escalonada equivalente por filas a A. La factorización LU no es más que *la codificación en forma matricial del proceso de eliminación gaussiana por filas* para conseguir una matriz escalonada equivalente.

Ya vimos anteriormente que cada operación elemental de filas es equivalente a multiplicar A por una matriz elemental E_i, por la izquierda:

$$E_pE_{p-1}\cdots E_2E_1A = U. \tag{2.10}$$

Supongamos que tenemos una matriz A tal que en su reducción por filas a forma escalonada no es necesario realizar permutaciones. Entonces las posiciones pivote que vayamos encontrando estarán en la fila superior de la submatriz que tratamos. Es posible reducir la matriz a una matriz escalonada U equivalente utilizando solamente operaciones elementales de fila de tipo $L_{i,j}(\alpha)$, sin necesidad de reescalar con $L_i(\alpha)$ las filas. Es más, en la reducción progresiva, la primera fase en la que hacemos cero todos los elementos que están debajo de los pivotes, se utilizan sólo operaciones de fila $L_{i,j}(\alpha)$ con $i > j$. El pivote está en la fila j, y siempre está por encima del elemento que anulamos, que está en la fila $i > j$. Resumiendo, si las permutaciones no son necesarias, la reducción a matriz escalonada se puede hacer con operaciones E_i que son todas de tipo $L_{i,j}(\alpha)$ de adición de múltiplo de una fila superior j a otra inferior i, con $i > j$.

Por ello, todas las matrices E_i de (2.10) son de la forma

$$L_{ij}(\alpha) = \begin{array}{c c} & \begin{array}{ccccccc} & & j & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \end{array} & \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & \alpha & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \end{array} \tag{2.11}$$

(α está en la fila i columna j). Estas matrices tienen dos características interesantes: i) son *triangulares inferiores*, es decir, por encima de la diagonal principal todos los elementos son nulos; ii) tienen la *diagonal unitaria*, es decir, la diagonal principal está compuesta de unos.

Ejercicio 2.35.

1. Demostrar que todo producto $E_p E_{p-1} \cdots E_1$ de matrices del tipo (2.11) es triangular inferior y de diagonal unitaria.

2. Demostrar que la inversa de una matriz (2.11) es del mismo tipo, de hecho $L_{ij}(\alpha)^{-1} = L_{ij}(-\alpha)$.

Poe el ejercicio anterior. La matriz A se reconstruye a partir de U invirtiendo todas las operaciones elementales $A = E_1^{-1} E_2^{-1} \cdots E_p^{-1} U$. Podemos entonces tomar $L = E_1^{-1} E_2^{-1} \cdots E_p^{-1}$ y hemos demostrado el siguiente teorema.

Teorema 2.36 (Factorización LU de una matriz)**.** *Toda matriz A de $m \times n$ que se puede reducir por filas a una forma escalonada U sin realizar permutaciones, se puede factorizar de la forma*

$$A = LU.$$

El factor L es una matriz $m \times m$ triangular inferior con unos en la diagonal e invertible, igual al inverso $L = E_1^{-1} E_2^{-1} \cdots E_p^{-1}$ de las matrices E_i correspondientes a operaciones elementales de fila de tipo $L_{ij}(\alpha)$ que reducen A a U.

Algoritmo de factorización LU. Es muy interesante el hecho de que se puede construir la matriz L de una forma muy sencilla. Como hemos argumentado, L es el producto de las inversas $E_1^{-1}E_2^{-1}\cdots E_p^{-1}$. Resulta que este producto se puede escribir directamente, sin necesidad de multiplicar las matrices. Lo mejor es hacer un ejemplo.

Ejemplo 2.37. Encontrar la factorización LU de

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

El primer pivote es el 2 de la posición 11. Las primeras operaciones elementales son $L_{21}(2)$, $L_{31}(-1)$ y $L_{41}(3)$ que anulan los tres elementos debajo del primer pivote y dejan la matriz en

$$A \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix}$$

Continuando con la segunda columna, siendo el pivote el 3 de la posición 22, realizamos las operaciones $L_{32}(3)$, $L_{42}(-4)$:

$$\sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix}.$$

El pivote se ha desplazado a la cuarta columna, a la posición 34, pero no es necesaria una permutación. Realizamos $L_{43}(-2)$ y obtenemos la matriz escalonada

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

La operación que ha reducido A a U es entonces, en forma matricial $L_{43}(-2)L_{42}(-4)L_{32}(3)L_{41}(3)L_{31}(-1)L_{21}(2)$. Pero a nosotros nos interesa su inversa

$$\begin{aligned} L &= [L_{43}(-2)L_{42}(-4)L_{32}(3)L_{41}(3)L_{31}(-1)L_{21}(2)]^{-1} \\ &= L_{21}(2)^{-1}L_{31}(-1)^{-1}L_{41}(3)^{-1}L_{32}(3)^{-1}L_{42}(-4)^{-1}L_{43}(-2)^{-1} \\ &= L_{21}(-2)L_{31}(1)L_{41}(-3)L_{32}(-3)L_{42}(4)L_{43}(2). \end{aligned}$$

Resulta que este producto es sorprendentemente sencillo:

$$
\begin{aligned}
L &= L_{21}(-2)L_{31}(1)L_{41}(-3)L_{32}(-3)L_{42}(4)L_{43}(2)\\
&= \begin{bmatrix} 1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ -3 & 0 & 0 & 1 \end{bmatrix}\\
&\qquad \cdot \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & -3 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 4 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & \boxed{2} & 1 \end{bmatrix}\\
&= \begin{bmatrix} 1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ -3 & 0 & 0 & 1 \end{bmatrix}\\
&\qquad \cdot \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & -3 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & \boxed{4} & \boxed{2} & 1 \end{bmatrix}\\
&= \begin{bmatrix} 1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ -3 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & \boxed{-3} & 1 & 0\\ 0 & \boxed{4} & \boxed{2} & 1 \end{bmatrix}\\
&= \begin{bmatrix} 1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & -3 & 1 & 0\\ -3 & 4 & 2 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & \boxed{-3} & 1 & 0\\ \boxed{-3} & \boxed{4} & \boxed{2} & 1 \end{bmatrix}\\
&= \begin{bmatrix} 1 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ \boxed{1} & \boxed{-3} & 1 & 0\\ \boxed{-3} & \boxed{4} & \boxed{2} & 1 \end{bmatrix}
= \begin{bmatrix} 1 & 0 & 0 & 0\\ \boxed{-2} & 1 & 0 & 0\\ \boxed{1} & \boxed{-3} & 1 & 0\\ \boxed{-3} & \boxed{4} & \boxed{2} & 1 \end{bmatrix}
\end{aligned}
$$

Debido al orden especial de las operaciones elementales que forman L ¡ los multiplicadores se han colocado en las posiciones correspondientes de la matriz final sin modificación ! Esto no sucede en la matriz $E = E_p \cdots E_2 E_1$ que convierte a A en escalonada, y que es inversa de L. Concluimos que

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -l_{21} & 1 & 0 & 0 \\ -l_{31} & -l_{32} & 1 & 0 \\ -l_{41} & -l_{42} & -l_{43} & 1 \end{bmatrix}.$$

siendo l_{ij} el multiplicador que hemos usado en la operación $L_{ij}(l_{ij})$ al reducir la matriz A.

La regla de construcción de la matriz L es general, válida para toda matriz A que se reduce sin usar permutaciones. Nótese que, sin embargo, el producto inverso L^{-1} que aplicamos a A para reducirla ($L^{-1}A = U$) no es tan sencillo de escribir, porque los coeficientes l_{ij} se mezclan. Afortunadamente, no es L^{-1} la matriz interesante, sino su inversa L.

Ejercicio 2.38. Obténgase la factorización LU de $A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix}$.

La utilidad práctica de la factorización $A = LU$ es inmensa si se necesitan resolver muchos sistemas $A\mathbf{x} = \mathbf{b}_i$ con la misma matriz A pero diferentes lados derechos b_i. El sistema $LU\mathbf{x} = \mathbf{b}$ se convierte en dos sistemas

$$A\mathbf{x} = LU\mathbf{x} = \mathbf{b} \qquad \Leftrightarrow \qquad L\mathbf{y} = \mathbf{b}, \quad U\mathbf{x} = \mathbf{y}$$

que son triangulares, mucho más fáciles de resolver (aunque sean dos) que el sistema original no triangular.

2.7. Cuestiones adicionales*

Número de operaciones para multiplicar dos matrices

Número de operaciones para factorizar LU. El número de operaciones (multiplicaciones) necesario para calcular el primer paso de la eliminación (ceros en la primera columna debajo del pivote) es n para calcular la primera fila que se va a sustraer, por $n-1$ filas que hay debajo de la primera fila. Así, para completar la forma escalonada harán falta

$$\begin{aligned}
&n(n-1)+(n-1)(n-2)+\cdots+2\cdot 1 = \\
&\qquad = n^2-n+(n-1)^2-(n-1)+\cdots+1^2-1 = \\
&\qquad = (1+2^2+\cdots n^2)-(1+2+\cdots+n) = \\
&\qquad = \frac{n(n+1)(2n+1)}{6}-\frac{n(n+1)}{2}=\frac{n^3-n}{3}\approx\frac{n^3}{3}
\end{aligned}$$

Número de operaciones para resolver un sistema triangular o escalonado

La factorización LU permutada. ¿ Qué sucede si son necesarias las permutaciones para obtener la matriz escalonada equivalente a A ? Imaginemos que antes de cada operación elemental E_r de tipo $L_{i,j}(\alpha)$ es necesario realizar una permutación $P_r = P_{ij}$ (que, sólo para este cálculo, podemos considerar puede ser la identidad I si no es necesaria ninguna permutación). Entonces (2.10) se escribirá

$$E_pP_pE_{p-1}P_{p-1}\cdots E_2P_2E_1P_1A = U$$

donde P_r son las permutaciones necesarias (ó I) y las E_r son matrices $L_{i,j}(\alpha)$. Hay que observar que si $P_r = P_{i,j}$ con $i > j$, entonces j es la fila donde situamos el pivote r-ésimo. La operación de reducción que le sigue $E_r = L_{k,j}$ ha de combinar la fila j del pivote con otra por debajo del pivote $k > j$. Es decir, los pares E_rP_r son de la forma $L_{k,j}P_{i,j}$ con $k > j$ e $i > j$, y $k \neq i$, porque i es la fila en la que hemos situado el cero que estaba en la fila j del pivote. Su inversa es

$$A = P_1E_1^{-1}P_2E_2^{-1}\cdots P_{p-1}E_{p-1}^{-1}P_pE_p^{-1}U$$

porque recordemos que $P_r^{-1} = P_r$ al ser una matriz de permutación. Tenemos que desenmarañar el producto que se encuentra delante de la matriz U, y lo primero que se tiene que decir es que, en general, no es una matriz L triangular inferior de diagonal unitaria como antes. Es, como vamos a ver, una versión de ese tipo de matrices, con las filas permutadas.

Ejercicio 2.39. Demuéstrese que, si $P_{i,j}$ es una matriz de permutación y $L_{k,l}(\alpha)$ una matriz elemental de adición de un múltiplo α de la fila l a la fila k, entonces $P_{i,j}L_{k,l}(\alpha)P_{i,j}$ es una matriz de tipo $L_{k',l'}(\alpha)$, aunque posiblemente con k' y l' distintos de k y l: si $k = i$ entonces $k' = j$, si $k = j$ entonces $k' = i$, si $l = i$ entonces $l' = j$ y si $l = j$ entonces $l' = i$. Si ni k ni l coinciden con i ó j, entonces $k' = k$, $l' = l$.

Entonces $E_{r-1}^{-1}P_r = L_{k,j}P_{i,j+1}$, con $i > j+1$ según el ejercicio anterior es

$$P_{i,j+1}L_{k,j}P_{i,j+1} = L_{k',j} \text{ con } k' > j.$$

es decir

$$L_{k,j}P_{i,j+1} = P_{i,j+1}L_{k',j}.$$

Entonces $E_{r-2}^{-1}P_{r-1}E_{r-1}^{-1}P_r = L_{l,j-1}P_{i',j}L_{k,j}P_{i,j+1}$ con $i' > j$ es

$$L_{l,j-1}P_{i',j}L_{k,j}P_{i,j+1} = L_{l,j-1}P_{i',j}P_{i,j+1}L_{k',j} = P_{i',j}L_{l',j-1}P_{i,j+1}L_{k',j}$$

con $l' > j-1$ y

$$P_{i',j}L_{l',j-1}P_{i,j+1}L_{k',j} = P_{i',j}P_{i,j+1}L_{l'',j-1}L_{k',j}$$

con $l'' > j-1$. Parece claro que al final todas las permutaciones se pueden pasar a la izquierda, con el precio de aumentar a veces los primeros índices k de los L_{kj} y quedará.

$$A = PL_{k_1 1}L_{k_2 2}\cdots L_{k_p p}U = PLU$$

escribir
con 1,2,3,... des
el principio

que al ser $k_1 > 1,\ldots,k_p > p$ es el producto de P, una matriz de permutación, por una matriz L triangular inferior con diagonal unitaria. Basta invertir P (obteniéndose otra matriz producto de permutaciones) para obtener el siguiente teorema.

P puede que ter
estructura.

Teorema 2.40 (Factorización LU permutada)**.** *Toda matriz A de $m \times n$ se puede descomponer de la forma*

$$PA = LU$$

donde P es una matriz de permutaciones (producto de $P_{i,j}$'s) L es una matriz triangular inferior (producto de $L_{i,j}$'s) cuadrada de diagonal unidad e invertible y U es una matriz escalonada.

La factorización LU permutada. Existe una factorización un poco más fina, la factorización LDU, que necesitaremos más adelante.

Teorema 2.41 (Descomposición LDU de una matriz)**.** *Toda matriz A de $m \times n$ se puede descomponer de la forma*

$$PA = LDU$$

donde

- *P es una matriz de permutaciones (producto de $P_{i,j}$'s)*
- *L es una matriz $n \times n$ triangular inferior con diagonal unitaria (producto de $L_{i,j}$'s) cuadrada e invertible*

- *D es una matriz $n \times n$ cuadrada diagonal*
- *U es una matriz escalonada con los pivotes igual a uno.*

En el caso de la matriz del ejemplo 2.37

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

la factorización LDU es

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

Proposición 2.42. *La factorización LDU es única salvo permutaciones P.*

La demostración: $L_1 D_1 U_1 = L_2 D_2 U_2$ implica que $U_1 U_2^{-1} = D_1^{-1} L_1^{-1} L_2 D_2$m y esto sólo es posible si $U_1 U_2^{-1} = I \Leftrightarrow U_1 = U_2$. Análogamente se demuestra $D_1 = D_2$ y $L_1 = L_2$.

Ejercicio. La factorización $A = LDU$ de una matriz simétrica $A = A^t$ es ser de la forma $A = LDL^t$.

2.8. Respuestas a los ejercicios

2.14 Multiplicando $CA = I$ por D por la derecha (si ello es posible) se tiene que

$$CAD = I \cdot D, \quad \Leftrightarrow \quad C \cdot I = D, \quad \Leftrightarrow \quad C = D$$

Hay restricciones para que el cálculo anterior sea posible:

2.18 Como $A\mathbf{x} = \mathbf{b}$ tiene solución $\forall \mathbf{b}$, A satisface el teorema S. El sistema $A\mathbf{x} = \mathbf{0}$ tiene solución única, luego se satisface el teorema I.

2.39

$$P_{i,j}L_{k,l}(\alpha)P_{i,j} = \begin{cases} L_{k,l}(\alpha) & \text{si } k \neq i \text{ ó } j, \text{ y } l \neq i \text{ ó } j \\ L_{j,l}(\alpha) & \text{si } k = i, \text{ y } l \neq i \text{ ó } j \\ L_{k,j}(\alpha) & \text{si } k \neq i \text{ ó } j, \text{ y } l = i \\ L_{i,l}(\alpha) & \text{si } k = j, \text{ y } l \neq i \text{ ó } j \\ L_{k,i}(\alpha) & \text{si } k \neq i \text{ ó } j, \text{ y } l = j \\ L_{j,i}(\alpha) & \text{si } k = i \text{ y } l = j \\ L_{i,j}(\alpha) & \text{si } k = j \text{ y } l = i \end{cases}$$

2.9. Resumen

Teorema (Propiedades de la suma y multiplicación por escalar de matrices.).

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rA + rB$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

Definición (Producto AB).

$$\begin{aligned} AB = A\left[\, \mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_p \,\right] \\ = \left[\, A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_p \,\right]. \end{aligned}$$

Teorema. *Dadas A de $m \times n$ y B y C:*

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + AB$
4. $c(AB) = (cA)B = A(CB)$
5. $I_m A = A = AI_n$

Teorema (Propiedades de la transpuesta).

1. $(A^T)^T = A,$
2. $(A + B)^T = A^T + B^T,$
3. $(cA)^T = cA^T,$
4. $(AB)^T = B^T A^T.$

Teorema (Inversa de matrices 2×2). *Si $ad - bc \neq 0$, A es invertible y*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ab - cd} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Si $ad - bc = 0$, A es singular.

Teorema (Propiedades de la inversa).

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$
3. $(A^T)^{-1} = (A^{-1})^T$

Teorema. *Una matriz A de $n \times n$ es invertible si y sólo si es equivalente por filas a la matriz identidad I_n. Cualquier sucesión de operaciones elementales de fila que reduce a A a la identidad, aplicada sobre la identidad, produce A^{-1}.*

Algoritmo para calcular A^{-1}.

Dada una matriz A de $n \times n$;

1. formar $\left[\, A \mid I \,\right]$;
2. hallar su forma escalonada reducida;
3. si A es invertible, la matriz obtenida será $\left[\, I \mid A^{-1} \,\right]$.

Teorema (Teorema de la matriz invertible ó B)**.** *Sea A una matriz $n \times n$ (cuadrada). Las siguientes afirmaciones son equivalentes.*

a. *A es invertible.*
b. *A es equivalente por filas a la matriz identidad I_n.*
c. *A tiene n posiciones pivote (es decir, n pivotes no nulos).*
d. *La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ admite sólo la solución trivial.*
e. *Las columnas de A forman un conjunto linealmente independiente.*
f.
g. *La ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para todo $\mathbf{b} \in \mathbf{R}^n$.*
h. *Las columnas de A generan $\mathbf{R}^n$.*
i.
j. *Existe una matriz C $n \times n$ tal que $CA = I$.*
k. *Existe una matriz D $n \times n$ tal que $AD = I$.*
l. *A^T es invertible.*

Teorema (Factorización LU de una matriz)**.** *Toda matriz A de $m \times n$ que se puede reducir por filas a una forma escalonada U sin realizar permutaciones, se puede factorizar de la forma*

$$A = LU.$$

El factor L es una matriz $m \times m$ triangular inferior con unos en la diagonal e invertible, igual al inverso $L = E_1^{-1} E_2^{-1} \cdots E_p^{-1}$ de las matrices E_i correspondientes a operaciones elementales de fila de tipo $L_{ij}(\alpha)$ que reducen A a U. Sus elementos fuera de la diagonal son los multiplicadores utilizados cambiados de signo.