Tema 1. Matrices y sistemas de ecuaciones lineales

Algunos ejemplos

Ejemplo 1. Queremos obtener 1000 litros de un *coupage* 50 % tempranillo, 26 % syrah y 24 % merlot. Disponemos para ello de tres barricas de vino B1, B2 y B3. La composición de B1 es de 3, 1 y 1 partes de tempranillo, syrah y merlot respectivamente. La de B2 es de 1, 2 y 1 y la B3 contiene partes iguales de dichas uvas. ¿Qué cantidad de cada barrica se necesita para obtener el coupage deseado?

Solución. Sean x, y, z los litros que usaremos de las barricas B1, B2 y B3 respectivamente. Debe cumplirse que:

$$\left\{\begin{array}{l} \frac{3}{5}x + \frac{1}{4}y + \frac{1}{3}z = 500 \\ \frac{1}{5}x + \frac{1}{2}y + \frac{1}{3}z = 260 \\ \frac{1}{5}x + \frac{1}{4}y + \frac{1}{3}z = 240 \end{array}\right.$$

Este es un sistema de tres ecuaciones lineales con tres incógnitas cuya solución única es $x = 650$, $y = 80$, $z = 270$.

Ejemplo 2.[1] Considera la siguiente tabla

	Carbón	Electricidad	Acero
Carbón	.0	.4	.6
Electricidad	.6	.1	.2
Acero	.4	.5	.2

La tabla representa un modelo muy sencillo de una economía con tres sectores de producción: carbón, electricidad y acero. Las columnas indican lo que cada sector exporta (*output*) por unidad de producción a otros sectores y permiten calcular los ingresos del sector. Las filas representan lo que cada sector importa (*inputs*) por unidad de producción de otros sectores y permiten calcular los costes de producción del sector[2]. Se supone que la producción de cada sector se mide en unidades de millones de dólares. Queremos calcular los precios por unidad de cada sector x (carbón), y (electricidad), z (acero) de manera que el ingreso de cada sector sea igual a su coste de producción (*precios de equilibrio*). Deberá cumplirse que:

$$\left.\begin{array}{r} .4y + .6z = x \\ .6x + .1y + .2z = y \\ .4x + .5y + .2z = z \end{array}\right\} \Longleftrightarrow \left\{\begin{array}{r} x \;\; - .4y - .6z = 0 \\ .6x - .9y + .2z = 0 \\ .4x + .5y - .8z = 0 \end{array}\right.$$

Se trata de un sistema de tres ecuaciones lineales con tres incógnitas, homogéneo, cuyas soluciones son de la forma $x = .94z$, $y = .85z$ y z es una variable libre a la que podemos asignar valores positivos. Para cada valor asignado a z se obtiene un conjunto de precios de equilibrio.

Ejemplo 3. [3] Cuando se quema gas propano (C_3H_8), éste se combina con oxígeno (O_2) para formar dióxido de carbono (CO_2) y agua (H_2O), de acuerdo con una reacción de la forma:

$$x_1 \mathrm{C_3H_8} + x_2 \mathrm{O_2} \longrightarrow x_3 \mathrm{CO_2} + x_4 \mathrm{H_2O}$$

donde x_1, x_2, x_3 y x_4 son números naturales que se calculan por la condición de que el número total de átomos de hidrógeno (H), oxígeno (O) y carbono (C) al principio y al final de la reacción sean iguales.

[1] Ejemplo tomado del libro *Álgebra Lineal y sus Aplicaciones* de David C. Lay, editado por Addison Wesley.

[2] Wassily Leontief, profesor de Harvard, desarrolló en 1949 un modelo *input-output* para la economía de USA considerando 500 sectores, lo que da lugar a un sistema de 500 ecuaciones lineales con 500 incógnitas. Se le concedió el Premio Nobel de Economía en 1973

[3] Tomado del mismo libro que el ejemplo anterior.

Esto conduce al siguiente sistema de ecuaciones lineales:

$$\left.\begin{array}{rcl} 8x_1 & = & 2x_4 \\ 2x_2 & = & 2x_3 + x_4 \\ 3x_1 & = & x_3 \end{array}\right\} \iff \left\{\begin{array}{l} 8x_1 \qquad\qquad - 2x_4 = 0 \\ \quad 2x_2 - 2x_3 - x_4 = 0 \\ 3x_1 \quad - x_3 \qquad = 0 \end{array}\right.$$

Cuya solución es:

$$x_1 = \frac{1}{4}x_4,\ x_2 = \frac{5}{4}x_4,\ x_3 = \frac{3}{4}x_4$$

donde x_4 puede tomar cualquier valor. Puesto que queremos soluciones que sean números naturales, lo razonable es tomar $x_4 = 4$, con lo que $x_1 = 1$, $x_2 = 5$ y $x_3 = 3$. La reacción ajustada es:

$$\mathrm{C_3H_8} + 5\mathrm{O_2} \longrightarrow 3\mathrm{CO_2} + 4\mathrm{H_2O}$$

Matrices

Las matrices son objetos matemáticos de gran utilidad para representar y resolver sistemas de ecuaciones lineales, trabajar con vectores, representar transformaciones en el espacio y mucho más. Aunque nuestro objetivo ahora es estudiar sistemas de ecuaciones lineales, es decir, dar criterios que permitan saber si tienen o no tienen soluciones y, en caso de que tengan, calcularlas todas, la herramienta fundamental para todo ello serán las matrices.

Un sistema de ecuaciones lineales, SEL en lo que sigue, es un sistema de m ecuaciones con n incógnitas x_i $(i = 1, 2, \dots, n)$ del tipo

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \tag{1}$$

donde se supone que los a_{ij} y los b_i son números reales conocidos, llamados, respectivamente, los *coeficientes* y los *términos independientes* del sistema.

Una solución del sistema (1) es un conjunto de n números reales $s_1, s_2, \dots, s_n$ tales que al sustituir en cada incógnita $x_i = s_i$, $i = 1, 2, \dots, n$, se verifican todas las ecuaciones del mismo. Un SEL se dice **incompatible** si no tienen ninguna solución, se dice **compatible determinado** si tiene una solución única y **compatible indeterminado** si tiene más de una solución (en cuyo caso, de hecho tiene infinitas soluciones).

Dos sistemas de ecuaciones lineales se dicen **equivalentes** si ambos tienen las mismas soluciones. Nuestro objetivo ahora va a ser convertir un SEL en otro SEL equivalente más sencillo y lo vamos a hacer mediante las siguientes tres *operaciones elementales*:

- Intercambiar entre sí dos ecuaciones.
- Multiplicar una ecuación por un número distinto de cero.
- Sustituir una ecuación por su suma con otra multiplicada por un número.

Observa que estas operaciones son reversibles y que al realizar cualquiera de ellas sobre un SEL obtenemos otro SEL equivalente. Por tanto, dos SEL tales que uno de ellos puede obtenerse a partir del otro realizando un número finito de estas operaciones son equivalentes. Para visualizar estas operaciones es muy conveniente usar matrices.

Una **matriz** es un conjunto rectangular de números reales que se acostumbra a encerrar entre paréntesis, es decir, son números dispuestos en filas y columnas de forma que cada fila (y cada columna) tenga el mismo número de elementos. Es una definición algo extraña, pero no importa porque lo que interesa es saber cómo se usan las matrices y los cálculos que se pueden hacer con ellas.

Una matriz que tiene m filas y n columnas se dice que es una matriz de *orden* $m \times n$. Se representa por $\mathcal{M}_{m\times n}$ el conjunto de todas las matrices de orden $m \times n$. El elemento que ocupa la fila i y la columna j de una matriz se representa por a_{ij}. Una matriz $\mathbf{A} \in \mathcal{M}_{m\times n}$ puede representarse en la forma:

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

o, más simplemente, $\mathbf{A} = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$.

Dos matrices $\mathbf{A} = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ y $\mathbf{B} = \big(b_{ij}\big)_{\substack{1\leqslant i\leqslant p\\ 1\leqslant j\leqslant q}}$, son iguales cuando tienen el mismo número de filas, $m = p$, y de columnas, $n = q$, y $a_{ij} = b_{ij}$ para todos $i = 1, 2, \dots m$, $j = 1, 2, \dots, n$.

Una matriz que tiene una sola columna se llama un **vector columna**, y una matriz que tiene una sola fila se llama un **vector fila**.

Si $\mathbf{A} \in \mathcal{M}_{m\times n}$, su matriz traspuesta, que se representa por $\mathbf{A}^t$, es una matriz de orden $n \times m$ que se obtiene intercambiando ordenadamente filas por columnas en $\mathbf{A}$.

Una matriz de orden $n \times n$ (mismo número de filas que de columnas) se llama una **matriz cuadrada** de orden n. La *diagonal principal* de una matriz cuadrada de orden n es la formada por los elementos a_{ii}, $i = 1, 2 \dots, n$. Una matriz cuadrada se llama **diagonal** si todos los elementos que no están en la diagonal principal son cero.

Operaciones con matrices

La suma de dos matrices del mismo orden $\mathbf{A} = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ y $\mathbf{B} = \big(b_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ se define como la matriz $\mathbf{A} + \mathbf{B} = \big(a_{ij} + b_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$, es decir, es la matriz del mismo orden que se obtiene sumando los elementos de $\mathbf{A}$ y $\mathbf{B}$ que ocupan los mismos lugares en ambas matrices. La suma de matrices es conmutativa y asociativa. La matriz $m \times n$ cuyas entradas son todas cero se representa por $\mathbf{O}_{m\times n}$ o, simplemente, $\mathbf{O}$ si no hay lugar a confusión. La matriz que se obtiene sustituyendo en una matriz $\mathbf{A}$ cada elemento por su opuesto se llama matriz opuesta de $\mathbf{A}$ y se representa por $-\mathbf{A}$.

Si λ es un número real y $\mathbf{A}$ es una matriz, la matriz $\lambda\mathbf{A}$ es la que se obtiene multiplicando cada elemento de $\mathbf{A}$ por λ.

Al igual que la suma de matrices del mismo orden, se define como la matriz obtenida sumando los elementos que ocupan el mismo lugar en ambas matrices, podría pensarse que el producto de dos matrices del mismo orden será la matriz obtenida multiplicando los elementos que ocupan el mismo lugar en ambas matrices. Pues quien piense eso se equivoca, porque el producto de matrices se define de una forma muy diferente. Para entender por qué el producto de matrices se define en la forma en que veremos, conviene interpretar las matrices como operadores lineales entre espacios vectoriales.

El espacio vectorial $\mathbb{R}^n$

Para todos $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ y para todo $\lambda \in \mathbb{R}$ definimos:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \qquad \lambda\mathbf{x} = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$$

Con estas operaciones, $\mathbb{R}^n$ es un espacio vectorial cuyos elementos representaremos en negrita y se llaman vectores (en la terminología de matrices son vectores fila). Si $\mathbf{x} = (x_1, x_2, \dots, x_n)$, los números x_i $(1 \leqslant i \leqslant n)$ se llaman las *componentes* del vector $\mathbf{x}$. El vector cuyas componentes son todas nulas lo representaremos por $\mathbf{0}$.

Dado un conjunto de vectores $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ de $\mathbb{R}^n$, cualquier vector $\mathbf{x}$ de la forma

$$\mathbf{x} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_m\mathbf{u}_m$$

donde los λ_i $(1 \leqslant i \leqslant m)$ son números reales, se llama una *combinación lineal* de los vectores $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$. Si el vector $\mathbf{0}$ puede expresarse como combinación lineal de los vectores $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ con coeficientes λ_i *no todos nulos*, se dice que dichos vectores son **linealmente dependientes**, cuando esto no puede hacerse se dice que son **linealmente independientes**. Una **base** de $\mathbb{R}^n$ es cualquier conjunto de n vectores linealmente independientes, $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, con la propiedad de que cualquier vector $\mathbf{x} \in \mathbb{R}^n$ puede expresarse como combinación lineal de ellos:

$$\mathbf{x} = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n$$

En tal caso los números $\lambda_1, \lambda_2 \ldots, \lambda_n$ están determinados de manera única por $\mathbf{x}$ y se llaman las *coordenadas* del vector $\mathbf{x}$ en la base B.

Es habitual usar en $\mathbb{R}^n$ la *base canónica* que es la formada por los *vectores unidad* $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ donde $\mathbf{e}_i$ es el vector que tiene todas sus componentes nulas excepto la que ocupa el lugar i que es igual a 1. En dicha base tenemos que:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \cdots + x_n \mathbf{e}_n$$

es decir, en dicha base las coordenadas coinciden con las componentes del vector.

Producto escalar de dos vectores

El producto escalar de dos vectores $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ en $\mathbb{R}^n$ lo representaremos por $\langle \mathbf{x} | \mathbf{y} \rangle$, y es el número definido por

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i \tag{2}$$

Las siguientes propiedades del producto escalar son consecuencia directa de las propiedades de la suma y del producto de números reales.

a) $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle$.

b) Si λ es un número real se verifica que $\langle \lambda \mathbf{x} | \mathbf{y} \rangle = \lambda \langle \mathbf{x} | \mathbf{y} \rangle$.

c) $\langle \mathbf{x} + \mathbf{y} | \mathbf{z} \rangle = \langle \mathbf{x} | \mathbf{z} \rangle + \langle \mathbf{y} | \mathbf{z} \rangle$.

Producto de matrices

Consideremos dos matrices $\mathbf{A} = (a_{ik})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant k \leqslant p}}$ y $\mathbf{B} = (b_{kj})_{\substack{1 \leqslant k \leqslant p \\ 1 \leqslant j \leqslant n}}$, *tales que el número de columnas de* $\mathbf{A}$ *sea igual al número de filas* de $\mathbf{B}$. Representemos por $\mathbf{A}_i = (a_{i1}, a_{i2}, \ldots, a_{ip})$ el vector formado por los elementos de la fila i de la matriz $\mathbf{A}$, y por $\mathbf{B}^j = (b_{1j}, b_{2j}, \ldots, b_{pj})$ el vector formado por los elementos de la columna j de la matriz $\mathbf{B}$. Observa que dichos vectores tienen el mismo número de componentes por lo que podemos hacer su producto escalar. Se define la matriz producto $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ como la matriz de orden $m \times n$, $\mathbf{C} = (c_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$, cuyos elementos vienen dados por $c_{ij} = \langle \mathbf{A}_i | \mathbf{B}^j \rangle$. Explícitamente

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \quad (1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n) \tag{3}$$

Se define la *matriz identidad* de orden n, $\mathbf{I}_n$, como la matriz diagonal $n \times n$ cuya diagonal principal está formada por unos. Cuando no sea preciso indicar la dimensión, escribiremos simplemente $\mathbf{I}$.

El producto de matrices tiene las propiedades (siempre que los productos puedan hacerse):

Asociativa. $\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C}$.

Distributivas. $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$ y $(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C}$.

Elementos neutros. Si $\mathbf{A} \in \mathcal{M}_{m \times n}$ entonces $\mathbf{I}_m \cdot \mathbf{A} = \mathbf{A}$, $\mathbf{A} \cdot \mathbf{I}_n = \mathbf{A}$.

Traspuesta de un producto. $(\mathbf{A}\cdot\mathbf{B})^t = \mathbf{B}^t\cdot\mathbf{A}^t$.

El producto de matrices no tiene la propiedad conmutativa.

Dada una matriz cuadrada $\mathbf{A}$, se llama **matriz inversa** de $\mathbf{A}$, y se nota $\mathbf{A}^{-1}$, a otra matriz del mismo orden verificando que $\mathbf{A}\cdot\mathbf{A}^{-1} = \mathbf{A}^{-1}\cdot\mathbf{A} = \mathbf{I}$. La existencia y el cálculo de la matriz inversa se estudiarán más adelante.

Producto de una matriz por un vector columna

Consideremos una matriz $\mathbf{A} = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$, de orden $m\times n$. Podemos multiplicar dicha matriz por un vector columna $n\times 1$ y obtenemos un vector columna $m\times 1$. Si $\mathbf{e}_j^t$ es el vector columna $n\times 1$, traspuesto del vector unidad $\mathbf{e}_j$, se tiene que

$$\mathbf{A}\cdot\mathbf{e}_j^t = \begin{pmatrix} a_{1j}\\ a_{2j}\\ \vdots\\ a_{mj}\end{pmatrix}$$

es el vector columna j de $\mathbf{A}$. Representemos dicho vector por $\mathbf{C}^j$. Ahora, dado $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, el producto, $\mathbf{A}\cdot\mathbf{x}^t$, de la matriz $\mathbf{A}$ por el vector columna $\mathbf{x}^t$ es el vector (columna) de $\mathbb{R}^m$ que viene dado por:

$$\mathbf{A}\cdot\mathbf{x}^t = \mathbf{A}\cdot\big(x_1\mathbf{e}_1^t + x_2\mathbf{e}_2^t + \cdots + x_n\mathbf{e}_n^t\big) = \sum_{j=1}^{n}\mathbf{A}\cdot x_j\mathbf{e}_j^t = \sum_{j=1}^{n} x_j\mathbf{C}^j \tag{4}$$

y por tanto es una combinación lineal de los vectores columna de la matriz $\mathbf{A}$.

Matrices asociadas a un SEL. Transformaciones elementales por filas

Al SEL (1) le asociamos la **matriz de los coeficientes** del sistema $\mathbf{A} = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ y la **matriz ampliada** $\widetilde{\mathbf{A}} = \big(\mathbf{A}|\mathbf{b}\big)$ donde $\mathbf{b}$ es el vector columna formado por los términos independientes. Observa que en la matriz ampliada cada fila se corresponde con una ecuación y cada columna, salvo la última, se corresponde con una incógnita. Si representamos por $\mathbf{x}^t$ un vector columna $n\times 1$, podemos escribir el SEL en la forma $\mathbf{A}\cdot\mathbf{x}^t = \mathbf{b}$. En consecuencia, teniendo en cuenta (4), el SEL (1) **es compatible si, y sólo si, el vector** $\mathbf{b}$ **es combinación lineal de los vectores columna de la matriz** $\mathbf{A}$.

Las operaciones elementales sobre ecuaciones del SEL descritas al principio, se traducen en las siguientes **transformaciones elementales por filas** sobre la matriz ampliada:

- Intercambiar dos filas. Representaremos por $F_i \leftrightarrow F_j$ la operación de intercambiar las filas i y j.
- Multiplicar una fila por un número distinto de cero. Representaremos por $F_i \to \alpha F_i$ la operación de multiplicar la fila i por el número α.
- Sustituir una fila por su suma con otra multiplicada por un número. Representaremos por $F_i \to F_i + \alpha F_j$ la operación de sustituir la fila i por su suma con la fila j multiplicada por α.

Observa que todas estas operaciones son reversibles y que al realizar cualquiera de ellas obtenemos una matriz que representa un SEL equivalente al de partida. Dos matrices, $\mathbf{A}$ y $\mathbf{B}$, se dice que son **equivalentes por filas**, y escribiremos $\mathbf{A}\sim\mathbf{B}$, si una se obtiene a partir de otra por un número finito de transformaciones elementales por filas. Realizando sobre la matriz ampliada de un SEL transformaciones elementales por filas obtenemos una matriz equivalente por filas cuyo SEL asociado es equivalente al de partida.

Nuestro objetivo para estudiar el SEL va a ser transformar la matriz ampliada del sistema en otra equivalente por filas lo más sencilla posible.

Reducción por filas. Forma de Hermite

Si una fila de una matriz contiene elementos no nulos, al que está situado más a la izquierda le llamaremos **pivote**. Una matriz se dice que es **escalonada** si las filas nulas (caso de que las haya) se encuentran situadas en la parte inferior y el pivote de cada fila no nula está a la derecha del pivote de la fila anterior. Observa que en un matriz escalonada debajo de un pivote solamente puede haber ceros. Una matriz **escalonada reducida** es una matriz escalonada cuyos pivotes son iguales a 1 y todos los elementos que están por encima de los pivotes son 0.

De las matrices

$$\mathbf{A}=\begin{pmatrix} 2 & 1 & 1 & 1\\ 0 & -1 & 1 & 2\\ 0 & 1 & 0 & 0\end{pmatrix},\mathbf{B}=\begin{pmatrix} 2 & 1 & 0 & 1\\ 0 & -1 & 0 & 2\\ 0 & 0 & 3 & 4\end{pmatrix},\mathbf{C}=\begin{pmatrix} 1 & 0 & 0 & -2\\ 0 & 1 & 0 & 3\\ 0 & 0 & 0 & 0\end{pmatrix},\mathbf{D}=\begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

la $\mathbf{A}$ no es escalonada, la $\mathbf{B}$ sí lo es pero no es reducida, la $\mathbf{C}$ y la $\mathbf{D}$ son escalonadas reducidas.

Se demuestra que, dada una matriz $\mathbf{A}$, hay una *única* matriz *escalonada reducida* H que es *equivalente por filas* a $\mathbf{A}$. Dicha matriz H se llama **forma de Hermite** (por filas) de $\mathbf{A}$ (y también forma de Gauss–Jordan de $\mathbf{A}$ o, simplemente, forma escalonada reducida de $\mathbf{A}$).

$\mathbf{A}$la forma de Hermite de una matriz se puede llegar por distintos caminos; una estrategia básica es conseguir un pivote en la primera fila en la posición $(1,1)$, lo cual es fácil si en la primera columna hay elementos distintos de 0; si no fuera así trataríamos de lograr un pivote en la posición $(1,2)$ y así sucesivamente. Logrado el pivote en la primera fila, se hacen 0 los elementos debajo de él restando a cada fila la primera multiplicada por un número conveniente. Hecho esto, nos fijamos ahora en la submatriz que queda al suprimir la primera fila y las columnas a la izquierda del pivote incluyendo la columna del pivote, y repetimos el proceso anterior. Finalmente, partiendo del último pivote, el que está más a la derecha, se hacen 0 los elementos por encima de él, y esto se hace con los demás pivotes siempre de derecha a izquierda y de abajo arriba. Veamos un ejemplo.

Ejemplo 4.

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5\\ 3 & -7 & 8 & -5 & 8 & 9\\ 3 & -9 & 12 & -9 & 6 & 15\end{pmatrix}\underset{F_1\leftrightarrow F_3}{\sim}\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15\\ 3 & -7 & 8 & -5 & 8 & 9\\ 0 & 3 & -6 & 6 & 4 & -5\end{pmatrix}\underset{F_1\to\frac{1}{3}F_1}{\sim}\begin{pmatrix} 1 & -3 & 4 & -3 & 2 & 5\\ 3 & -7 & 8 & -5 & 8 & 9\\ 0 & 3 & -6 & 6 & 4 & -5\end{pmatrix}\underset{F_2\to F_2-3F_1}{\sim}$$

$$\begin{pmatrix} 1 & -3 & 4 & -3 & 2 & 5\\ 0 & 2 & -4 & 4 & 2 & -6\\ 0 & 3 & -6 & 6 & 4 & -5\end{pmatrix}\underset{F_2\to\frac{1}{2}F_2}{\sim}\begin{pmatrix} 1 & -3 & 4 & -3 & 2 & 5\\ 0 & 1 & -2 & 2 & 1 & -3\\ 0 & 3 & -6 & 6 & 4 & -5\end{pmatrix}\underset{F_3\to F_3-3F_2}{\sim}\begin{pmatrix} 1 & -3 & 4 & -3 & 2 & 5\\ 0 & 1 & -2 & 2 & 1 & -3\\ 0 & 0 & 0 & 0 & 1 & 4\end{pmatrix}$$

Esta matriz ya está en forma escalonada. Lo que queda para llegar a la forma de Hermite es muy fácil.

$$\begin{pmatrix} 1 & -3 & 4 & -3 & 2 & 5\\ 0 & 1 & -2 & 2 & 1 & -3\\ 0 & 0 & 0 & 0 & 1 & 4\end{pmatrix}\underset{\substack{F_2\to F_2-F_3\\ F_1\to F_1-2F_3}}{\sim}\begin{pmatrix} 1 & -3 & 4 & -3 & 0 & -3\\ 0 & 1 & -2 & 2 & 0 & -7\\ 0 & 0 & 0 & 0 & 1 & 4\end{pmatrix}\underset{F_1\to F_1+3F_2}{\sim}\begin{pmatrix} 1 & 0 & -2 & 3 & 0 & -24\\ 0 & 1 & -2 & 2 & 0 & -7\\ 0 & 0 & 0 & 0 & 1 & 4\end{pmatrix}$$

Esta técnica de reducción por filas proporciona un procedimiento sistemático para resolver un SEL. El siguiente SEL

$$\left.\begin{aligned} 3x_2-6x_3+6x_4+4x_5&=-5\\ 3x_1-7x_2+8x_3-5x_4+8x_5&=9\\ 3x_1-9x_2+12x_3-9x_4+6x_5&=15\end{aligned}\right\}\tag{5}$$

Tiene como matriz ampliada la del ejemplo 4

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5\\ 3 & -7 & 8 & -5 & 8 & 9\\ 3 & -9 & 12 & -9 & 6 & 15\end{pmatrix}$$

cuya forma de Hermite, que acabamos de calcular, es

$$\begin{pmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

Por tanto, dicho SEL es equivalente al siguiente:

$$\left.\begin{array}{rl} x_1 \quad - 2x_3 + 3x_4 \quad & = -24 \\ x_2 - 2x_3 + 2x_4 \quad & = \ -7 \\ x_5 & = \quad 4 \end{array}\right\} \tag{6}$$

En este caso tenemos cinco incógnitas y tres ecuaciones, por lo que debemos asignar valores arbitrarios a dos de las incógnitas: a aquellas cuyas columnas correspondientes no tienen pivote, es decir, x_3 y x_4. Entonces, poniendo $x_3 = s$, $x_4 = t$, obtenemos como soluciones del SEL (5) las de la forma:

$$x_1 = -24 + 2s - 3t, \;\; x_2 = -7 + 2s - 2t, \;\; x_3 = s, \;\; x_4 = t, \;\; x_5 = 4$$

donde s y t son números reales cualesquiera que suelen llamarse *parámetros*. En esta situación se dice que $x_3 = s$ y $x_4 = t$ son *variables o parámetros libres* y x_1, x_2, x_5 son *variables dependientes*. Esta manera de representar las soluciones se llama **forma paramétrica**. Dicho SEL es compatible indeterminado. El método que hemos seguido parar resolver este sistema se llama **método de Gauss–Jordan**.

La forma de Hermite de la matriz ampliada de un SEL no solamente permite, cuando el sistema es compatible, calcular muy fácilmente todas las soluciones del mismo, sino que también nos informa del carácter de dicho sistema. Se verifican los siguientes resultados:

1) Si la forma de Hermite de la *matriz ampliada* de un SEL tiene alguna fila cuyos elementos son todos cero menos el último que es distinto de cero, entonces dicho sistema es incompatible. Cuando esto no sucede el sistema es compatible.

2) Si el SEL es compatible y todas las columnas de la forma de Hermite de la *matriz de coeficientes del sistema* tienen un pivote, entonces el sistema es compatible determinado con solución única.

3) Si el SEL es compatible y en la forma de Hermite de la *matriz de coeficientes del sistema* hay columnas sin pivote, entonces las variables asociadas a dichas columnas son variables libres o parámetros y la solución general del sistema se expresa en función de ellas. El sistema es compatible indeterminado.

Observa que para usar estos resultados no es imprescindible (si no se quieren calcular las soluciones) llegar a la forma de Hermite de la matriz ampliada del sistema, sino que basta con llegar a una matriz escalonada equivalente.

Se llama **rango** de una matriz el número de filas distintas de cero de su forma de Hermite[4]. Dicho número también es igual al de columnas de la forma de Hermite que tienen un pivote. El rango de una matriz es un número menor o igual que el número de filas y menor o igual que el número de columnas. El rango de una matriz es cero si, y sólo si, la matriz es la matriz nula.

Teorema de Rouché–Frobenius. Sean $\mathbf{A}$ la matriz de los coeficientes de un SEL y $\widetilde{\mathbf{A}}$ la matriz ampliada. Entonces:

1) Si rango$(\mathbf{A}) <$ rango$(\widetilde{\mathbf{A}})$, el sistema es incompatible.

2) Si rango$(\mathbf{A}) =$ rango$(\widetilde{\mathbf{A}})$ y es igual al número de incógnitas, el sistema es compatible determinado.

3) Si rango$(\mathbf{A}) =$ rango$(\widetilde{\mathbf{A}})$ y es menor que el número de incógnitas, el sistema es compatible indeterminado. En este caso el número de variables libres es igual al número de incógnitas menos rango$(\mathbf{A})$.

[4]Para calcular el rango basta con obtener una matriz escalonada equivalente

Existencia y cálculo de la matriz inversa

El siguiente resultado nos dice qué matrices cuadradas tienen inversa.

Teorema. Sea $\mathbf{A}$ una matriz cuadrada $n \times n$. Las siguientes afirmaciones son equivalentes:

1) $\mathbf{A}$ es invertible, es decir, $\mathbf{A}$ tiene inversa.

2) La forma de Hermite de $\mathbf{A}$ es la identidad $\mathbf{I}_n$.

3) El rango de $\mathbf{A}$ es igual a n.

Un método para calcular la matriz inversa de una matriz $\mathbf{A}$, consiste en escribir una nueva matriz $\mathbf{C} = \big(\mathbf{A} \mid \mathbf{I}\big)$, formada por la matriz $\mathbf{A}$ y a su derecha la matriz identidad $\mathbf{I}$. $\mathbf{A}$continuación en dicha matriz $\mathbf{C}$ se realizan las transformaciones elementales por filas necesarias hasta que a la izquierda nos quede la identidad y de esa forma obtendremos la matriz $\big(\mathbf{I} \mid \mathbf{A}^{-1}\big)$, es decir, $\mathbf{A}^{-1}$ es la matriz que ha quedado a la derecha.

$$\left(\begin{array}{ccc|ccc} 3 & 2 & 3 & 1 & 0 & 0 \\ 3 & 2 & 4 & 0 & 1 & 0 \\ 2 & 1 & 3 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -3 & 2 \\ 0 & 1 & 0 & -1 & 3 & -3 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right)$$

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

Determinantes

$\mathbf{A}$cada matriz cuadrada $\mathbf{A}$ vamos a asociar un número que llamaremos su determinante y representaremos por $\det(\mathbf{A})$ o por $|\mathbf{A}|$. Diremos como se calcula el determinante de matrices 2×2, y daremos una regla que permite calcular el determinante de matrices cuadradas de orden $n > 2$ reduciendo dicho cálculo al de determinantes de orden $n - 1$. Aplicando dicha regla repetidamente se puede reducir el cálculo de cualquier determinante al cálculo de determinantes de orden 2.

El determinante de una matriz 2×2 es:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Si $\mathbf{A}$ es una matriz, representaremos por $\mathbf{A}_{ij}$ la matriz obtenida suprimiendo en $\mathbf{A}$ la fila i y la columna j. Si $\mathbf{A}$ es una matriz cuadrada de orden $n > 2$, $\mathbf{A}_{ij}$ es una matriz cuadrada de orden $n - 1$. Supuesto que $\mathbf{A}$ es una matriz cuadrada de orden $n > 2$ se define:

$$\det(\mathbf{A}) = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det(\mathbf{A}_{ij}) \tag{7}$$

El número $(-1)^{i+j}\det(\mathbf{A}_{ij})$ se llama **adjunto** del elemento a_{ij} de $\mathbf{A}$, y se dice que la expresión anterior es el desarrollo del determinante de $\mathbf{A}$ por los elementos de la fila i.

Veamos que, efectivamente, aplicando la igualdad (7) a una matriz 3×3 obtenemos su determinante.

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

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

Igualdad conocida con el nombre de **regla de Sarrus**.

También podemos calcular el determinante de una matriz desarrollando por los elementos de una columna:

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det(\mathbf{A}_{ij}) \tag{8}$$

Propiedades de los determinantes

Sea $\mathbf{A}$ una matriz cuadrada de orden $n \geqslant 2$. Se verifica que:

1) Si una fila de $\mathbf{A}$ es combinación lineal de otras filas de $\mathbf{A}$, entonces $\det(\mathbf{A}) = 0$.

2) Si $\mathbf{B}$ se obtiene al multiplicar una sola fila de $\mathbf{A}$ por un número α, entonces $\det(\mathbf{B}) = \alpha \det(\mathbf{A})$.

3) Si $\mathbf{B}$ se obtiene al intercambiar dos filas en $\mathbf{A}$, entonces $\det(\mathbf{B}) = -\det(\mathbf{A})$.

4) Si $\mathbf{B}$ se obtiene al sumar a una fila de $\mathbf{A}$ otra multiplicada por un número, entonces $\det(\mathbf{B}) = \det(\mathbf{A})$.

5) $\det(\mathbf{A}) = \det(\mathbf{A}^t)$.

6) Si $\mathbf{B}$ es una matriz cuadrada del mismo orden que $\mathbf{A}$, $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \det(\mathbf{B})$.

7) Una matriz cuadrada se llama **triangular superior** (inferior) si todos los elementos por debajo (por encima) de la diagonal principal son nulos. Si $\mathbf{A}$ es una matriz triangular (superior o inferior), el determinante de $\mathbf{A}$ es igual al producto de los elementos de su diagonal principal.

Como $\det(\mathbf{A}) = \det(\mathbf{A}^t)$, las propiedades anteriores son también válidas si en ellas cambiamos filas por columnas de $\mathbf{A}$.

Estas propiedades permiten calcular el determinante de una matriz transformándola en otra más sencilla mediante transformaciones elementales de filas o de columnas. La estrategia es tratar de hacer en una fila o columna todos los elementos cero salvo uno o dos y desarrollar el determinante por los elementos de dicha fila o columna.

Una importante propiedad de los determinantes es la siguiente.

Teorema. *Una matriz cuadrada es invertible si, y sólo si, su determinante es distinto de cero.*

Las matrices cuadradas con determinante distinto de cero se llaman **matrices regulares**.

Los determinantes pueden usarse para calcular el rango de una matriz cualquiera. Si en una matriz suprimimos algunas filas y columnas la matriz resultante se dice que es una **submatriz** de la primera.

Teorema. *Sea $\mathbf{A} \in \mathcal{M}_{m\times n}$, el rango de $\mathbf{A}$ coincide con el mayor orden de una submatriz cuadrada regular de $\mathbf{A}$.*

Teorema. *Una matriz cuadrada es regular si, y sólo si, sus vectores columna son linealmente independientes, en cuyo caso también son linealmente independientes sus vectores fila.*

Matriz adjunta e inversa

Si $\mathbf{A}$ es una matriz cuadrada su matriz adjunta, representada por $\mathrm{Adj}(\mathbf{A})$ es la formada por los adjuntos de sus elementos. Se verifica el siguiente resultado.

Teorema. Si $\mathbf{A}$ es una matriz cuadrada regular, entonces su inversa viene dada por

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\big(\mathrm{Adj}(\mathbf{A})\big)^t$$

Regla de Cramer

Es una forma de calcular mediante determinantes la solución de un SEL compatible determinado. Si el sistema, escrito en forma matricial, es $\mathbf{A} \cdot \mathbf{x}^t = \mathbf{b}$, donde $\mathbf{A}$ es una matriz regular, el sistema es compatible determinado y su solución viene dada por:

$$x_i = \frac{\det(\mathbf{X}_i)}{\det(\mathbf{A})}, \quad i = 1, 2, \ldots n.$$

donde $\mathbf{X}_i$ es la matriz obtenida al sustituir en la matriz $\mathbf{A}$ la columna i por el vector $\mathbf{b}$.