

Sistemas de ecuaciones lineales

Ingeniería informática Curso: Álgebra

Fernando Chamizo <https://matematicas.uam.es/~fernando.chamizo/>

Comentarios

Ahora veremos los frutos del algoritmo para escalar matrices y su motivación primaria. Es la parte principal de un método eficiente para resolver sistema de ecuaciones lineales y también permite entender la estructura de sus soluciones. Por otro lado, los determinantes tuvieron históricamente su origen en la resolución de estos sistemas y recordaremos una antigua fórmula de interés teórico que, bajo ciertas hipótesis, los utiliza para expresar las soluciones.

1. Sistemas y matrices

Un *sistema de ecuaciones lineales* es una ecuación del tipo

$$A\vec{x} = \vec{b}$$

donde $A \in \mathcal{M}_{m \times n}$ y $\vec{b} \in \mathcal{M}_{m \times 1}$ son una matriz y un vector fijos y $\vec{x} = (x_1, \dots, x_n)^t$ es un vector de incógnitas. Si desarrollas el producto $A\vec{x}$ verás que simplemente son m ecuaciones del tipo $a_{i1}x_1 + \dots + a_{in}x_n = b_i$. La tabla de los coeficientes a_{ij} viene dada por la matriz A y si la extendemos con el vector \vec{b} como última columna, decimos que es la *matriz ampliada* del sistema que se denota con A^+ . Esto es, $A^+ = (A|\vec{b})$, donde la barra vertical no indica nada, es solo para separar.

Resulta que un sistema de ecuaciones lineales con matriz es escalonada es sencillo de resolver (dando por hecho que esto sea posible) porque la solución de la ecuación correspondiente al último pivote se puede sustituir en las anteriores para obtener un sistema con una incógnita menos y repetir el proceso tantas veces como pivotes haya. Esto es lo que se llama *sustitución regresiva*. Por ejemplo, consideremos el siguiente sistema con matriz escalonada:

$$\begin{array}{rcl} 2x_1 - x_2 - 2x_3 & = & 6 \\ x_2 - x_3 & = & -1 \\ 3x_3 & = & -3 \end{array} \quad A = \begin{pmatrix} 2 & -1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{pmatrix}.$$

La última ecuación implica $x_3 = -1$ y al sustituir en las otras se obtiene

$$\begin{array}{rcl} 2x_1 - x_2 & = & 4 \\ x_2 & = & -2 \end{array} \quad \text{con matriz} \quad \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}.$$

Sustituyendo $x_2 = -2$ se sigue $x_1 = 1$ con lo que la solución es $x_1 = 1$, $x_2 = -2$, $x_3 = -1$.

En la sustitución regresiva, como el nombre sugiere, siempre se despejan unas variables en términos de variables con índice superior. Si alguna variable no se pueda despejar, significa que puede tomar valores arbitrarios, actúa como un parámetro dando lugar a infinidad de soluciones. Es habitual denominar tales parámetros con λ , μ o λ_j . Un momento de reflexión lleva a convencerse de que *las variables que despejamos son justamente las que corresponden a las columnas en las que hay pivotes*. El siguiente ejemplo ilustra la situación.

Ejemplo. Resolver por sustitución regresiva el sistema $A\vec{x} = \vec{b}$ con

$$A = \begin{pmatrix} 1 & 1 & -5 & 3 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{y} \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}.$$

Escrito explícitamente el sistema es

$$\left. \begin{array}{rcl} x_1 + x_2 - 5x_3 + 3x_4 & = & 1 \\ -x_2 + 2x_3 - 2x_4 & = & 0 \\ 2x_4 & = & 4 \end{array} \right\}$$

La última ecuación da $x_4 = 2$ y al sustituir en la anterior se obtiene $-x_2 + 2x_3 = 4$. No podemos despejar x_2 y x_3 por separado, solo una en función de la otra. Siguiendo la política de la sustitución regresiva, tomamos x_3 (la de índice mayor) como parámetro $x_3 = \lambda$ y despejamos $x_2 = 2\lambda - 4$. Finalmente, al sustituir en la primera ecuación $x_1 + (2\lambda - 4) - 5\lambda + 6 = 1$ que da $x_1 = -1 + 3\lambda$. En definitiva, el conjunto de soluciones es:

$$\{\vec{x} \in \mathbb{R}^4 : A\vec{x} = \vec{b}\} = \{(-1 + 3\lambda, 2\lambda - 4, \lambda, 2)^t : \lambda \in \mathbb{R}\}.$$

Si A termina con cierta cantidad de filas de ceros y \vec{b} no, evidentemente el sistema no se puede resolver, pues incluye ecuaciones imposibles del tipo $0 = b_j \neq 0$.

Es fácil anticiparse al siguiente paso: Conocemos un algoritmo, la eliminación de Gauss, que reduce una matriz a forma escalonada, entonces si lo aplicamos a la matriz ampliada de un sistema y lo combinamos con la sustitución regresiva, tenemos un algoritmo para resolver sistemas. Con un ligero abuso de notación es común referirse a tal algoritmo como *eliminación de Gauss*, ya que esta es su parte principal. Es importante notar que como las transformaciones T1, T2 y T3 se pueden deshacer, no se pierden ni añaden soluciones al aplicar el algoritmo.

Comentario. En cálculo numérico (el área relacionada con la implementación de métodos matemáticos en los ordenadores), la eliminación de Gauss para resolver sistemas va acompañado de ciertos refinamientos. Se suele representar como la llamada *descomposición LU*, una expresión de la matriz del sistema como el producto LU donde L y U son una matriz triangular inferior ($l_{ij} = 0$ si $j > i$) y una matriz triangular superior ($u_{ij} = 0$ si $j < i$). En realidad, no siempre existe tal descomposición y se hacen ciertas permutaciones de las incógnitas para que no surjan problemas y para que los pivotes sean grandes con el motivo de reducir errores de redondeo. A la técnica de permutación para que el pivote tenga el mayor valor absoluto de cada fila se le llama *pivote parcial*. La relación entre la descomposición LU y la eliminación de Gauss es muy sencilla y, si tienes interés, la encontrarás en casi cualquier fuente que se ocupe de aspectos numéricos.

Ejemplos que habíamos visto antes con el único propósito inmotivado de escalonar matrices, ahora se vuelven más tangibles porque permiten resolver sistemas.

Ejemplo. Resolver con eliminación de Gauss el sistema

$$\begin{array}{rcl} 2x_1 - x_2 - 2x_3 & = & 6 \\ 2x_1 + x_2 - 4x_3 & = & 4 \\ -2x_1 + 3x_2 + 3x_3 & = & -11 \\ 10x_1 - 2x_2 - 10x_3 & = & 24 \end{array} \quad \text{que tiene matriz ampliada } A^+ = \left(\begin{array}{ccc|c} 2 & -1 & -2 & 6 \\ 2 & 1 & -4 & 4 \\ -2 & 3 & 3 & -11 \\ 10 & -2 & -10 & 24 \end{array} \right).$$

Creamos los ceros en la primera columna para el primer escalón:

$$\left(\begin{array}{ccc|c} 2 & -1 & -2 & 6 \\ 2 & 1 & -4 & 4 \\ -2 & 3 & 3 & -11 \\ 10 & -2 & -10 & 24 \end{array} \right) \xrightarrow[\substack{f_2 \mapsto f_2 - f_1 \\ f_3 \mapsto f_3 + f_1 \\ f_4 \mapsto f_4 - 5f_1}]{} \left(\begin{array}{ccc|c} 2 & -1 & -2 & 6 \\ 0 & 2 & -2 & -2 \\ 0 & 2 & 1 & -5 \\ 0 & 3 & 0 & -6 \end{array} \right).$$

A pesar de que T2 no es estrictamente necesaria, la usamos para simplificar y creamos los ceros que faltan:

$$\xrightarrow[\substack{f_2 \mapsto f_2/2 \\ f_4 \mapsto f_4/3}]{} \left(\begin{array}{ccc|c} 2 & -1 & -2 & 6 \\ 0 & 1 & -1 & -1 \\ 0 & 2 & 1 & -5 \\ 0 & 1 & 0 & -2 \end{array} \right) \xrightarrow[\substack{f_3 \mapsto f_3 - 2f_2 \\ f_4 \mapsto f_4 - f_2}]{} \left(\begin{array}{ccc|c} 2 & -1 & -2 & 6 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 1 & -1 \end{array} \right) \xrightarrow{f_4 \mapsto f_4 - \frac{1}{3}f_3} \left(\begin{array}{ccc|c} 2 & -1 & -2 & 6 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 1 & -1 \end{array} \right).$$

Con esto se llega a una forma escalonada que coincide con la que vimos al comienzo, así que la solución es como allí $x_1 = 1$, $x_2 = -2$, $x_3 = -1$.

Si se emplea la eliminación de Gauss-Jordan, entonces la sustitución regresiva es superflua, se obtiene directamente la solución.

Ejemplo. Resolver $2x_1 + x_2 = 4$, $x_1 + x_2 = 3$ utilizando reducción de Gauss-Jordan.

El algoritmo es

$$\left(\begin{array}{cc|c} 2 & 1 & 4 \\ 1 & 1 & 3 \end{array}\right) \xrightarrow{f_2 \mapsto f_2 - \frac{1}{2}f_1} \left(\begin{array}{cc|c} 2 & 1 & 4 \\ 0 & 1/2 & 1 \end{array}\right) \xrightarrow{f_1 \mapsto f_1/2, f_2 \mapsto 2f_2} \left(\begin{array}{cc|c} 1 & 1/2 & 2 \\ 0 & 1 & 2 \end{array}\right) \xrightarrow{f_1 \mapsto f_1 - f_2/2} \left(\begin{array}{cc|c} 1 & 0 & \mathbf{1} \\ 0 & 1 & \mathbf{2} \end{array}\right).$$

La solución se lee directamente en la última columna, marcada en negrita.

A pesar de que la eliminación de Gauss-Jordan evita la sustitución regresiva en el caso de solución única (también lo hace con pequeñas variaciones en el caso de infinitas soluciones), desde el punto de vista computacional lo que uno gasta en extender el algoritmo de Gauss al de Gauss-Jordan es comparable a lo que gasta en sustitución regresiva. Más interesante es que la eliminación de Gauss-Jordan permite resolver simultáneamente varios sistemas con la misma matriz de coeficientes, una situación común en algunos problemas de ingeniería. En breve, para resolver

$$A\vec{x} = \vec{b}_1, \quad A\vec{x} = \vec{b}_2, \quad \dots \quad A\vec{x} = \vec{b}_k,$$

se aplica eliminación de Gauss-Jordan a $(A|\vec{b}_1\vec{b}_2\ldots\vec{b}_k)$.

Ejemplo. Resolver simultáneamente los siguientes sistemas mediante eliminación de Gauss-Jordan:

$$\begin{cases} 2x_1 + 3x_2 = 7, \\ 3x_1 - x_2 = 5, \end{cases} \quad \begin{cases} 2x_1 + 3x_2 = 11, \\ 3x_1 - x_2 = 0, \end{cases} \quad \begin{cases} 2x_1 + 3x_2 = 2, \\ 3x_1 - x_2 = 3. \end{cases}$$

La reducción de Gauss de $(A|\vec{b}_1\vec{b}_2\vec{b}_3)$ es

$$\left(\begin{array}{cc|ccc} 2 & 3 & 7 & 11 & 2 \\ 3 & -1 & 5 & 0 & 3 \end{array}\right) \xrightarrow{f_2 \mapsto f_2 - \frac{3}{2}f_1} \left(\begin{array}{cc|ccc} 2 & 3 & 7 & 11 & 2 \\ 0 & -\frac{11}{2} & -\frac{11}{2} & -\frac{33}{2} & 0 \end{array}\right).$$

Ahora debemos multiplicar la primera fila por $1/2$ y la segunda por $-2/11$ para que los pivotes sean uno y después crear un cero encima del segundo pivote:

$$\xrightarrow{f_1 \mapsto f_1/2, f_2 \mapsto -2f_2/11} \left(\begin{array}{cc|ccc} 1 & 3/2 & 7/2 & 11/2 & 1 \\ 0 & 1 & 1 & 3 & 0 \end{array}\right) \xrightarrow{f_1 \mapsto f_1 - 3f_2/2} \left(\begin{array}{cc|ccc} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 3 & 0 \end{array}\right).$$

Por tanto, las soluciones de los sistemas son $(x_1, x_2) = (2, 1), (1, 3), (1, 0)$.

2. Estructura del conjunto de soluciones

El objetivo aquí es dar el siguiente enunciado teórico acerca de la estructura de las soluciones de los sistemas lineales que será útil a lo largo del curso:

Un sistema de ecuaciones lineales $A\vec{x} = \vec{b}$ con $\text{rg}(A) = r$ y $A \in \mathcal{M}_{m \times n}$ tiene solución si y solo si $r = \text{rg}(A^+)$. En ese caso, eligiendo una solución \vec{x}_0 , todas las soluciones vienen dadas por

$$\vec{x}_0 + \sum_{j=1}^{n-r} \lambda_j \vec{v}_j$$

con λ_j arbitrarios y \vec{v}_j ciertos vectores. Además elecciones distintas de los λ_j dan lugar a soluciones distintas. Para $n - r = 0$ se entiende que el sumatorio (que es vacío) no aparece, esto es, que la solución es única.

Si tienes alma de matemático, entenderás que este enunciado una forma sintética de recoger las conclusiones obtenidas al aplicar la reducción de Gauss. Nota que siempre $n \geq r$ porque con n columnas caben a lo más n escalones.

Una observación importante para este curso es que los sistemas de la forma $A\vec{x} = \vec{0}$, llamados *sistemas homogéneos*, siempre tienen la solución $\vec{x}_0 = \vec{0}$ y así, cuando esta no es la única solución, todas las soluciones vienen dadas por el sumatorio.

Comentario. Ahora ya podemos entender por qué el rango de una matriz está bien definido. Supongamos que al aplicar eliminación de Gauss a M alguien obtuviera n pivotes y otro, aplicándola de diferente modo, obtuviera $n' < n$ pivotes. Si consideramos la matriz A formada por las columnas en las que hay pivotes con el primer procedimiento, entonces $A\vec{x} = \vec{0}$ tendría $\vec{x} = \vec{0}$ como única solución (podemos despejar todas las variables), mientras que con el segundo procedimiento aparecerían parámetros λ_j y habría infinitas soluciones, lo cual es una contradicción.

Una consecuencia del resultado anterior que casi siempre sirve para resolver un problema de las pruebas de acceso a la universidad se llama *teorema de Rouché-Frobenius* en los países de habla hispana:

Dado un sistema de ecuaciones lineales $A\vec{x} = \vec{b}$ con $A \in \mathcal{M}_{m \times n}$ se cumple:

- 1) Tiene solución única, se dice que es compatible determinado, si y solo si $n = \text{rg}(A) = \text{rg}(A^+)$.
- 2) Tiene infinitas soluciones, se dice que es compatible indeterminado, si y solo si $n > \text{rg}(A) = \text{rg}(A^+)$.
- 3) No tiene solución, se dice que es incompatible, si y solo si $\text{rg}(A) \neq \text{rg}(A^+)$.

Comentario. El matemático F.G. Frobenius no tiene mucho que ver con este resultado. Parece que el nombre fue introducido por el matemático hispanoargentino J. Rey Pastor. En el mundo anglosajón se llama *teorema de Rouché-Capelli*, lo que es mucho más lógico porque tanto E. Rouché como A. Capelli lo enunciaron y demostraron, aunque no fueran los primeros en hacerlo.

El hecho de que el resultado principal sea teórico, no impide considerar un ejemplo.

Ejemplo. Hallar posibles vectores v_j en el resultado para el sistema homogéneo $A\vec{x} = \vec{0}$ con

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

La forma escalonada se consigue con los siguientes pasos:

$$A \xrightarrow[\substack{f_2 \mapsto f_2 + 3f_1 \\ f_3 \mapsto f_3 - 2f_1}]{\substack{f_2 \mapsto f_2 + 3f_1 \\ f_3 \mapsto f_3 - 2f_1}} \begin{pmatrix} 1 & 1 & -1 & 2 & -3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 3 & 0 & -1 \end{pmatrix} \xrightarrow{f_3 \mapsto f_3 - 3f_2} \begin{pmatrix} 1 & 1 & -1 & 2 & -3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Las columnas que no contienen pivotes son la segunda y la cuarta y corresponden a parámetros libres, así $x_2 = \lambda_1$, $x_4 = \lambda_2$, mientras que despejando x_1 , x_3 y x_5 utilizando las ecuaciones de abajo a arriba (sustitución regresiva), $x_5 = 0$, $x_3 = 0$, $x_1 = -\lambda_1 - 2\lambda_2$. Con ello

$$\vec{x} = \begin{pmatrix} -\lambda_1 - 2\lambda_2 \\ \lambda_1 \\ 0 \\ \lambda_2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

El paso intermedio es solo para practicar con el producto de matrices, está claro que las coordenadas de \vec{v}_1 y \vec{v}_2 , los dos últimos vectores columna, vienen dadas por los coeficientes de λ_1 y λ_2 en la expresión inicial para \vec{x} .

3. Regla de Cramer

Es posible dar una expresión compacta y elegante en términos de determinantes para sistemas de ecuaciones lineales compatibles determinados con el mismo número de ecuaciones e incógnitas. Esta expresión es la llamada *regla de Cramer* y seguro que la conoces de cursos anteriores. Su enunciado es:

Sea $A\vec{x} = \vec{b}$ con $A \in \mathcal{M}_n(K)$ un sistema compatible determinado y sea A_j la matriz A reemplazando la columna j -ésima por \vec{b} , entonces su solución es

$$x_j = \frac{|A_j|}{|A|}, \quad j = 1, 2, \dots, n.$$

A diferencia de lo que ocurre con la eliminación de Gauss, permite hallar las incógnitas independientemente: se puede calcular una sin conocer las otras.

Ejemplo. Escribir la solución general, suponiéndola única, de un sistema con dos ecuaciones lineales con dos incógnitas.

Un sistema genérico con dos ecuaciones y dos incógnitas es

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \quad \text{y la regla de Cramer da} \quad x_1 = \frac{1}{|A|} \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \quad x_2 = \frac{1}{|A|} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

Desarrollando los determinantes,

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}}, \quad x_2 = \frac{a_{11} b_2 - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}}.$$

Si el sistema tuviera infinitas soluciones o no tuviera ninguna, como sabemos, $|A| = 0$ y estas fórmulas perderían su sentido porque llevarían a dividir por cero.

Comentario. La regla de Cramer es muy antigua, anterior a la teoría de matrices y vectores. Se remonta al siglo XVIII cuando todavía no había una definición de determinante. Lo que hizo Cramer es describir precisamente los términos que aparecerían en el denominador y en el numerador cuando se generaliza el ejemplo anterior a cualquier número de variables.

Aunque la regla de Cramer sea compacta y elegante, tiene poca utilidad práctica en ejemplos concretos cuando la dimensión crece. El desarrollo de un determinante 4×4 con la definición contiene 24 productos de 4 factores, lo que conduce a $24 \cdot 3 + 23 \approx 100$ operaciones simples. Entonces, si nos negamos a aplicar ninguna forma de eliminación de Gauss, la regla de Cramer para un sistema genérico con $A \in \mathcal{M}_4$ requerirá que hagamos unas 500 operaciones simples, lo cual es demasiado.

Ejemplo. Aplicando la regla de Cramer, resolver el sistema $A\vec{x} = \vec{b}$ con

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 1 & -2 & -3 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{y} \quad \vec{b} = \begin{pmatrix} -2 \\ 5 \\ 10 \end{pmatrix}.$$

Tras unos cálculos obtenemos que el determinante de A es 16. Más cálculos muestran

$$\begin{vmatrix} -2 & 2 & 1 \\ 5 & 1 & 2 \\ 10 & -2 & -3 \end{vmatrix} = 48, \quad \begin{vmatrix} 1 & -2 & 1 \\ 3 & 5 & 2 \\ 1 & 10 & -3 \end{vmatrix} = -32 \quad \text{y} \quad \begin{vmatrix} 1 & 2 & -2 \\ 3 & 1 & 5 \\ 1 & -2 & 10 \end{vmatrix} = -16.$$

Por tanto la solución es $x = 48/16 = 3$, $y = -32/16 = -2$, $z = -16/16 = -1$.

Comentario. A pesar de que la regla de Cramer es bastante inútil en términos prácticos y, en general, los algoritmos numéricos rara vez emplean determinantes, el caso de dimensión tres participa en el *algoritmo de intersección de Möller-Trumbore*, ampliamente usado en los motores de videojuegos y en

animación. En pocas palabras, los objetos virtuales 3D se representan habitualmente como mallas de triángulos con texturas y para detectar colisiones hay que resolver muchas veces el problema de si una recta en \mathbb{R}^3 corta a un triángulo. Este problema conduce de manera bastante trivial a un sistema de ecuaciones lineales y el algoritmo mencionado es un pequeño atajo cuya efectividad se multiplica por la gran cantidad de triángulos a considerar en la práctica.