

## Sistemas Lineales

En el curso de Álgebra Lineal hemos estudiado métodos *exactos* para resolver sistemas de ecuaciones lineales. Sin embargo hay dos razones para analizar numéricamente las soluciones. La primera de ellas es puramente calculística: el número de operaciones y la precisión de los cálculos acumulan errores que debemos tratar de minimizar. La segunda es intrínseca de cada sistema: las soluciones de algunos sistemas lineales varían considerablemente para pequeñas variaciones en los datos. Esta segunda característica tiene gran importancia cuando los datos del sistema proceden de mediciones físicas, y por tanto llevan implícito un cierto grado de incertidumbre. Será por tanto importante el conocer la sensibilidad de la solución respecto de los datos en cada problema.

*Ejemplo 17.* Consideremos el sistema lineal

$$\begin{cases} 99'87x + 12'35y = 2'35 \\ 7'231x + 0'9936y = 1'12 \end{cases}$$

que resolvemos por el método de eliminación de Gauss:

VÉASE EL EJEMPLO 19

$$\begin{pmatrix} 99'87 & 12'35 & \vdots & 2'35 \\ 7'231 & 0'9936 & \vdots & 1'12 \end{pmatrix} \quad \begin{pmatrix} 99'87 & 12'35 & \vdots & 2'35 \\ 0 & 0'0994 & \vdots & 0'9499 \end{pmatrix}$$
$$\begin{pmatrix} 99'87 & 0 & \vdots & -115'7 \\ 0 & 0'0994 & \vdots & 0'9499 \end{pmatrix}$$

Obtenemos como solución:  $x = -1'16$ ,  $y = 9'56$ .

Supongamos que los datos anteriores no se han medido con precisión y se ha obtenido el sistema

$$\begin{cases} 100x + 12y = 2'3 \\ 7x + y = 1'1. \end{cases}$$

La solución es entonces  $x = -0'68$ ,  $y = 5'87$ ; muy diferente de la anterior.

Vemos en este ejemplo los dos problemas con los que nos vamos a enfrentar. Primeramente la inestabilidad de los cálculos. En segundo lugar la dependencia de los datos. Ambos problemas suelen aparecer juntos y dependen principalmente del sistema dado.

*Ejemplo 18.* Veamos ahora el otro problema que puede aparecer en la resolución de sistemas lineales: cómo los errores de redondeo son distintos según el orden de las operaciones. Los sistemas

$$\begin{cases} 57x + 31y = 12 \\ 27x + 16y = 21 \end{cases} \quad y \quad \begin{cases} 0'57x + 0'31y = 0'12 \\ 27x + 16y = 21 \end{cases}$$

son obviamente equivalentes. Si utilizamos el método de Gauss para resolver el segundo sin mantener en las operaciones más de tres cifras decimales obtenemos

$$\begin{pmatrix} 0'57 & 0'31 & \vdots & 0'12 \\ 0 & 1'42 & \vdots & 15'32 \end{pmatrix} \quad \begin{pmatrix} 0'57 & 0 & \vdots & -3'23 \\ 0 & 1'42 & \vdots & 15'32 \end{pmatrix}$$

que da como solución  $x = -5'67$ ,  $y = 10'79$ . La solución correcta del sistema es  $x = -6'12$ ,  $y = 11'64$  que se obtiene por el método de eliminación de Gauss si se opera directamente con la matriz ampliada

$$\begin{pmatrix} 57 & 31 & \vdots & 12 \\ 27 & 16 & \vdots & 21 \end{pmatrix}$$

manteniendo dos cifras decimales durante los cálculos.

## 25 Métodos de resolución ya conocidos

Dado el sistema lineal

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \dots & \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

lo expresaremos en forma matricial como

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}$$

o bien  $A\mathbf{x} = \mathbf{b}$ .

El que el sistema tenga solución y que ésta sea única depende de los rangos de las matrices  $A$  y ampliada  $A|\mathbf{b}$ :

TEOREMA DE ROUCHÉ-FROBENIUS

1. Si  $\text{rango}(A|\mathbf{b}) = \text{rango}(A)$  el sistema es compatible (tiene solución). Si  $\text{rango}(A) = n$  entonces la solución es única (compatible determinado). Si  $\text{rango}(A) < n$  habrá infinitas soluciones (compatible indeterminado).
2. Si  $\text{rango}(A|\mathbf{b}) > \text{rango}(A)$  entonces el sistema es incompatible (no tiene solución).

Un sistema de  $n$  ecuaciones y  $n$  incógnitas que tenga solución única puede resolverse por medio de la Regla de Cramer. Sin embargo este método resulta muy ineficiente. Para utilizarlo hay que calcular

$n+1$  determinantes  $n \times n$ . Cada uno de estos determinantes requiere  $n!$  productos de  $n$  factores cada uno (es decir  $n!(n-1)$  productos binarios) y  $n! - 1$  sumas, y finalmente  $n$  cocientes. En total:  $n!(n-1) + n$  productos y cocientes binarios, y  $n! - 1$  sumas.

Un método más eficiente es el método de eliminación de Gauss que el lector debe conocer (véase en cualquier caso el Ejemplo 19). Para transformar el sistema en uno triangular realizamos el siguiente número de operaciones. En los  $n-1$  primeros pasos hacemos  $n(n-1)$  productos-cocientes y  $(n-1)^2$  sumas; en los siguientes  $n-2$  pasos hacemos  $(n-1)(n-2)$  productos-cocientes y  $(n-1)^2$  sumas; etc.; de forma que el número de sumas y de productos hasta obtener la forma triangular es de

$$\sum_{k=1}^{n-1} (k+1)k = \frac{(n+1)n(n-1)}{3} \quad \sum_{k=1}^{n-1} k^2 = \frac{n(n-1)(2n-1)}{6}$$

es decir, aproximadamente  $\frac{n^3}{3}$  cada una. Si comparamos estos valores para varios  $n$  obtenemos:

$n$	2	3	4	10	20
Cramer	2	12	72	$3'27 \cdot 10^7$	$2'4 \cdot 10^{18}$
Gauss	1	5	14	285	2470

A fin de observar las muy distintas magnitudes de estos números supongamos que realizamos nuestro cálculo con una máquina que efectúa  $10^6$  operaciones cada *tres* segundos; para  $n = 20$  el método de Gauss tardaría  $10^{-3}$  segundos; el método de Cramer 230.000 años.

*Ejemplo 19 (Método de eliminación de Gauss).* Resolver el sistema lineal

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

Utilizando la primera ecuación eliminamos la incógnita  $x$  de las demás ecuaciones. A continuación con la nueva segunda ecuación eliminamos la incógnita  $y$  de la tercera. Finalmente resolvemos por sustitución.

$$\begin{cases} 2x+ & y+ & z = & 1 \\ & -y- & 2z = & -4 \\ & 3y+ & 2z = & 8 \end{cases} \quad \begin{cases} 2x+ & y+ & z = & 1 \\ & -y- & 2z = & -4 \\ & & -4z = & -4 \end{cases}$$

$$\begin{cases} x = & \frac{1}{2} - \frac{1}{2}y - \frac{1}{2}z \\ y = & 4 - 2y \\ z = & 1 \end{cases} \quad \begin{cases} x = & -1 \\ y = & 2 \\ z = & 1. \end{cases}$$

Todo el trabajo anterior puede hacerse con matrices, es decir escribiendo solamente los coeficientes.

Analizando la primera parte del trabajo podemos observar que desde el punto de vista matricial cada paso consiste en el producto por la izquierda por una matriz regular, triangular inferior. Obsérvese que si  $E_{ij} = (a_{kl})$  con  $a_{kl} = \delta_{ki}\delta_{lj}$  entonces  $(I + aE_{ij})A$  es la matriz que se obtiene a partir de  $A$  añadiendo a la fila  $i$ , la fila  $j$  multiplicada por  $a$ . Así que la resolución del sistema consiste en realizar la cadena de operaciones matriciales siguiente

DELTA DE KRONECKER:  $\delta_{ij} = 1$  si  $i = j$ ;  $\delta_{ij} = 0$  si  $i \neq j$ .

$$(I+3E_{32})(I+E_{31})(I-2E_{21})A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (I+3E_{32})(I+E_{31})(I-2E_{21}) \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$$

que podemos escribir en forma abreviada

$$L^{-1}A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = L^{-1} \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$$

donde

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

Si  $U = L^{-1}A$ , esta matriz es triangular superior, por tanto  $A = LU$  es una descomposición de  $A$  como producto de una matriz triangular inferior con unos en la diagonal ( $L$ ) y una matriz triangular superior ( $U$ ) cuyos elementos diagonales son los utilizados en la eliminación que, con abuso de lenguaje, suelen llamarse *pivotes*. La matriz  $U$  puede escribirse como producto de una triangular superior con unos en la diagonal y una matriz diagonal cuyos elementos son los pivotes.

PIVOTES

A veces el proceso anterior no puede llevarse a cabo directamente pues alguno de los pivotes puede ser *cero*. En ese caso debe intercambiarse el orden de las filas (es decir, el orden de las ecuaciones) para conseguir un pivote que no sea *cero*. Desde el punto de vista del álgebra matricial este intercambio de filas corresponde a multiplicar por una matriz que se obtiene a partir de la matriz identidad intercambiando las mismas dos filas.

## 26 Pivote parcial

Si en la aplicación del método de Gauss se utiliza un pivote muy pequeño los errores de redondeo influyen en gran manera en el resultado final.

*Ejemplo 20.* El sistema

$$\begin{aligned} 2x + \frac{1}{3}y + 3z &= 4 \\ x + \frac{1}{6}y - 2z &= 1 \\ 2x + 3y - z &= 6 \end{aligned}$$

tiene solución  $x = 77/56 = 1'375$ ,  $y = 66/56 = 1'179$ ,  $z = 2/7 = 0'286$ . Si utilizamos para resolverlo el método del pivote con aritmética de punto flotante de 4 dígitos obtenemos  $x = 1'57$ ,  $y = 0$ ,  $z = 0'286$ . El problema aparece al redondear  $1/3$  y  $1/6$  que no resultan ser uno exactamente el doble del otro y por tanto un pivote que tendría que ser *cero* (y que nos obligaría a cambiar el orden de las ecuaciones) resulta ser  $0'001$  y se utiliza para eliminar el coeficiente correspondiente de la siguiente ecuación multiplicándolo por un número muy alto. Si antes de realizar esta operación se intercambian las dos últimas ecuaciones forzando a utilizar un pivote mucho más grande el resultado que se obtiene es  $x = 1'37$ ,  $y = 1'18$ ,  $z = 0'286$  (véase el Ejemplo 24).

## 27 Sensibilidad

*Ejemplo 21.*

$$\begin{pmatrix} 1 & 1 \\ 1 & 1'1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \begin{array}{l} x = 1 \\ y = 0 \end{array}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1'1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1'1 \\ 0'9 \end{pmatrix} \quad \begin{array}{l} x = 3'1 \\ y = -2 \end{array}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1'1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0'9 \\ 1'1 \end{pmatrix} \quad \begin{array}{l} x = -1'1 \\ y = 2. \end{array}$$

En el ejemplo anterior vemos que pequeñas variaciones en los datos producen grandes variaciones en la solución. Queremos estudiar métodos que nos permitan detectar *a priori* sistemas que presentan estos problemas. Haremos el estudio cuando las variaciones ocurren en el término independiente, como en el ejemplo.

Si  $A\mathbf{x} = \mathbf{b}$  y  $A(\mathbf{x} + \delta\mathbf{u}) = \mathbf{b} + \varepsilon\mathbf{v}$  entonces  $A\delta\mathbf{u} = \varepsilon\mathbf{v}$  es decir  $\delta\mathbf{u} = A^{-1}\varepsilon\mathbf{v}$ , donde  $\mathbf{u}$ ,  $\mathbf{v}$  son vectores unitarios. Un cambio pequeño de longitud  $\varepsilon$  en el dato  $\mathbf{b}$  puede producir un cambio de longitud  $\delta$  en la solución cuyo tamaño depende de la matriz  $A^{-1}$ . Supongamos que  $A$  tiene autovalores  $\lambda_1$  y  $\lambda_2$  ( $0 < \lambda_1 < \lambda_2$ ), con autovectores  $\mathbf{v}_1, \mathbf{v}_2$ . Entonces  $A^{-1}$  tiene autovalores  $1/\lambda_1, 1/\lambda_2$ , con los mismos autovectores. Por tanto, si  $\mathbf{v} = \mathbf{v}_1$  y  $\delta\mathbf{u}_1 = A^{-1}\varepsilon\mathbf{v}_1 = \varepsilon\frac{1}{\lambda_1}\mathbf{v}_1$ , el cambio en la longitud de la solución será de  $\frac{\varepsilon}{\lambda_1}$ .

Si la matriz  $A$  es simétrica sus autovalores son reales. Supongamos que verifican  $0 < |\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_n|$ . Entonces  $\|A\mathbf{x}\| \leq |\lambda_n|\|\mathbf{x}\|$ , donde  $\|\mathbf{u}\|$  representa la longitud del vector  $\mathbf{u}$ .

Por otra parte, si resolvemos  $A\mathbf{y} = \mathbf{b} + \varepsilon\mathbf{v}$ , obtendremos  $\mathbf{y} = \mathbf{x} + \delta\mathbf{w}$  y por tanto  $A(\delta\mathbf{w}) = \varepsilon\mathbf{v}$  o bien  $\delta\mathbf{w} = \varepsilon A^{-1}\mathbf{v}$ .

Si  $A$  es simétrica, también lo será  $A^{-1}$  que tendrá por autovalores

$$\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$$

por tanto

$$\|\delta \mathbf{w}\| = |\varepsilon| \|A^{-1} \mathbf{v}\| \leq |\varepsilon| \frac{1}{|\lambda|} \|\mathbf{v}\|$$

Entonces el error relativo será

$$\frac{|\delta|}{\|\mathbf{x}\|} \leq \frac{|\varepsilon| |\lambda_n|}{|\lambda_1| \|\mathbf{b}\|} = \frac{|\lambda_n|}{|\lambda_1|} \frac{|\varepsilon|}{\|\mathbf{b}\|}.$$

La acotación del error relativo en  $\mathbf{x}$  por medio del error relativo en  $\mathbf{b}$  viene multiplicada por el factor  $\frac{|\lambda_n|}{|\lambda_1|}$ . Este número se denomina *número de condición* de la matriz  $A$ . Esta desviación en el error relativo se alcanza cuando el vector  $\mathbf{b}$  está en la dirección del autoespacio correspondiente a  $\lambda_n$  y  $\mathbf{v}$  está en la dirección del autoespacio correspondiente a  $\lambda_1$ .

NÚMERO DE CONDICIÓN

Si la matriz  $A$  no es simétrica el problema es más complicado.

*Ejemplo 22.*

$$A = \begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix} \quad \text{tiene autovalor 1 doble sin embargo}$$

$$\begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10 \\ 1 \end{pmatrix} \quad \text{tiene por solución} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix} \quad \text{tiene por solución} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}$$

El error relativo en  $\mathbf{b}$  es

$$\frac{\|(0, 1)\|}{\|(10, 1)\|} = \frac{1}{\sqrt{101}} \approx \frac{1}{10}$$

mientras que el error relativo en  $(x, y)$  es

$$\frac{\|(10, -1)\|}{\|(0, 1)\|} = \sqrt{101} \approx 10$$

es decir, el error relativo se multiplica por 100.

Para analizar este tipo de problemas tenemos que comparar  $\|A\mathbf{x}\|$  con  $\|\mathbf{x}\|$ . Denominamos *norma* de la matriz  $A$  al número

NORMA

$$\|A\| = \sup_{\mathbf{x} \neq \mathbf{o}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}; \quad \text{donde } \mathbf{o} \text{ es el vector de longitud } \textit{cero}.$$

por definición  $\forall \mathbf{x}$ ,  $\|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\|$ . Si  $A\mathbf{x} = \mathbf{b}$  y  $A(\mathbf{x} + \delta \mathbf{v}) = \mathbf{b} + \varepsilon \mathbf{u}$ , entonces  $A(\delta \mathbf{v}) = \varepsilon \mathbf{u}$  y  $\delta \mathbf{v} = A^{-1}(\varepsilon \mathbf{u})$ . Por tanto,  $|\delta| \leq \|A^{-1}\| |\varepsilon|$  y resulta que

$$\frac{|\delta|}{\|\mathbf{x}\|} \leq \frac{\|A^{-1}\| |\varepsilon|}{\|\mathbf{b}\| \frac{1}{\|A\|}} \quad \text{es decir} \quad \frac{|\delta|}{\|\mathbf{x}\|} \leq \|A^{-1}\| \|A\| \frac{|\varepsilon|}{\|\mathbf{b}\|}.$$

El error relativo en  $\mathbf{b}$  se multiplica por un número no mayor que  $\|A\| \|A^{-1}\|$ . La dificultad estriba en calcular  $\|A\|$ .

Veamos un método.

$$\|A\|^2 = \sup \frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2} = \sup \frac{\mathbf{x}^T A^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \lambda_{\max}(A^T A)$$

donde el último número representa el mayor valor propio de la matriz simétrica  $A^T A$ .

Tenemos que calcular también  $\|A^{-1}\|$ . Por el mismo método obtenemos

$$\|A^{-1}\|^2 = \lambda_{\max}((A^{-1})^T A^{-1}) = \lambda_{\max}((AA^T)^{-1}) = \frac{1}{\lambda_{\min}(AA^T)}.$$

Resulta que los autovalores de  $AA^T$  son los mismos que los de  $A^T A$  y por tanto el número de condición buscado es

$$\|A\| \|A^{-1}\| = \sqrt{\frac{\lambda_{\max}(AA^T)}{\lambda_{\min}(AA^T)}}$$

*Ejemplo 23.* En el Ejemplo 22

$$A^T A = \begin{pmatrix} 1 & 0 \\ 10 & 1 \end{pmatrix} \begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 10 \\ 10 & 101 \end{pmatrix}$$

por tanto  $(1-\lambda)(101-\lambda)-100=0$  tiene por soluciones  $\lambda=101'99$  y  $\lambda=1/101'99$ . Como los autovalores de  $A^{-1}$  son los inversos de los anteriores, es decir los mismos, resulta que  $\|A^{-1}\| \|A\| \approx 102$ , que coincide con los resultados obtenidos en el Ejemplo 22.

¿Qué podemos decir acerca del cambio en  $\mathbf{x}$  cuando las variaciones se dan en  $A$ ? Sea  $B$  esta variación; medimos su tamaño por  $\|B\|$  y en términos relativos por  $\frac{\|B\|}{\|A\|}$ .

Si  $A\mathbf{x} = \mathbf{b}$  y  $(A+B)(\mathbf{x}+\delta\mathbf{v}) = \mathbf{b}$  entonces  $A\delta\mathbf{v} = -B(\mathbf{x}+\delta\mathbf{v})$  y por tanto  $\delta\mathbf{v} = -A^{-1}B(\mathbf{x}+\delta\mathbf{v})$ . Así  $\|\delta\mathbf{v}\| \leq \| -A^{-1}\| \|B\| \|\mathbf{x}+\delta\mathbf{v}\|$ ;

$$\frac{\|\delta\mathbf{v}\|}{\|\mathbf{x}+\delta\mathbf{v}\|} \leq \|A^{-1}\| \|B\| = \|A^{-1}\| \|A\| \frac{\|B\|}{\|A\|}.$$

donde  $\frac{\|\delta\mathbf{v}\|}{\|\mathbf{x}+\delta\mathbf{v}\|} \approx \frac{|\delta|}{\|\mathbf{x}\|}$ , error relativo en  $\|\mathbf{x}+\delta\mathbf{v}\|$  (recuérdese que  $\mathbf{u}$  y  $\mathbf{v}$  son unitarios).

¿Cómo puede el cálculo estropear el número de condición de una matriz?

Si tratamos de resolver el problema  $A\mathbf{x} = \mathbf{b}$  por medio de la descomposición  $A = LU$  de la matriz  $A$  en producto de una matriz  $L$ , triangular inferior, y una matriz  $U$ , triangular superior, que se

LA MATRIZ  $A^T A$  NO SOLAMENTE TIENE TODOS SUS VALORES REALES SINO QUE ADEMÁS SON ESTRICTAMENTE POSITIVOS, ES DECIR, ES *definida positiva*: SI  $\lambda$  ES UN AUTOVALOR DE  $A^T A$  Y  $\mathbf{x} \neq \mathbf{0}$  ES UN AUTOVECTOR ASOCIADO A ÉL ENTONCES  $0 < \|A\mathbf{x}\|^2 = (\mathbf{x}^T A^T A) (\mathbf{x}) = \mathbf{x}^T (A^T A \mathbf{x}) = \lambda \|\mathbf{x}\|^2$ .

SI  $\lambda$  ES UN AUTOVALOR DE  $A^T A$  Y  $\mathbf{x}$  ES UN AUTOVECTOR ASOCIADO A ÉL, ENTONCES  $AA^T(A\mathbf{x}) = A(A^T A \mathbf{x}) = A(\lambda\mathbf{x}) = \lambda A\mathbf{x}$ , POR TANTO  $\lambda$  ES UN AUTOVALOR DE  $AA^T$  Y  $A\mathbf{x}$  ES UN AUTOVECTOR ASOCIADO A ÉL.

ha explicado en la página 4, el problema queda dividido en dos: primeramente la resolución de  $L\mathbf{y} = \mathbf{b}$ , y después la resolución del problema  $U\mathbf{x} = L^{-1}\mathbf{b}$ . En cada uno de estos pasos los errores se podrán multiplicar hasta por los números de condición de las matrices  $L$  y  $U$ . Veamos en un ejemplo que el producto de estos dos números de condición es mayor que el número de condición de la matriz  $A$ .

*Ejemplo 24.* Sea

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1'1 \end{pmatrix}.$$

Sus autovalores son 2'051 y 0'049 por tanto el número de condición de la matriz es 41'86. La descomposición  $A = LU$  resulta

$$\begin{pmatrix} 1 & 1 \\ 1 & 1'1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0'1 \end{pmatrix}.$$

Los autovalores de  $L^T L$  son 2'618 y 0'382 y los de  $U^T U$  son 2'005 y 0'005 por tanto  $\|L\| \|L^{-1}\| = 2'62$  y  $\|U\| \|U^{-1}\| = 20$ . El producto es 52'4.

## 28 Métodos iterativos

La dificultad de resolver un sistema lineal radica en que, esencialmente, tenemos que invertir una matriz. Una posible estrategia para simplificar los cálculos consiste en sustituir la matriz a invertir  $A$  por otra de fácil inversión y reescribir un sistema equivalente que se resuelve de forma aproximada basándose en una solución aproximada del sistema. El esquema es el siguiente:

Dado el sistema  $A\mathbf{x} = \mathbf{b}$  escribimos la matriz  $A$  en la forma  $A = S - P$ , donde la matriz  $S$  es de fácil inversión. Entonces podemos reescribir el sistema como  $S\mathbf{x} = \mathbf{b} + P\mathbf{x}$ . Si  $\mathbf{x}^0$  es una solución aproximada buscamos la solución del sistema  $S\mathbf{x} = \mathbf{b} + P\mathbf{x}^0$ , de fácil cálculo:

$$\mathbf{x} = S^{-1}(\mathbf{b} + P\mathbf{x}^0).$$

En realidad el resultado obtenido es una aproximación  $\mathbf{x}^1$  de la solución del sistema  $A\mathbf{x} = \mathbf{b}$ . A partir de ella podemos repetir el procedimiento y construir de forma iterativa la sucesión  $(\mathbf{x}^k)$  que esperamos converja a la solución del sistema. La regla iterativa será

$$\mathbf{x}^{k+1} = S^{-1}(\mathbf{b} + P\mathbf{x}^k).$$

¿Qué hará falta para conseguir la convergencia? Si  $\mathbf{x}$  es la solución buscada, el error en el paso  $k + 1$  se expresa

$$\|\mathbf{x}^{k+1} - \mathbf{x}\| = \|S^{-1}P(\mathbf{x}^k - \mathbf{x})\| \leq \|S^{-1}P\| \|\mathbf{x}^k - \mathbf{x}\|$$



y por tanto

$$\|\mathbf{x}^{k+1} - \mathbf{x}\| \leq \|S^{-1}P\|^{k+1} \|\mathbf{x}^0 - \mathbf{x}\|$$

por lo que si  $\|S^{-1}P\| < 1$  el método converge independientemente de la aproximación de partida  $\mathbf{x}^0$ .

Esta es una condición suficiente para la convergencia, pero no es necesaria. Otra condición suficiente es que el mayor autovalor (en módulo) de la matriz  $S^{-1}P$  sea estrictamente menor que 1.

Veamos dos posibles elecciones de  $S$  y de  $P$  que dan lugar a dos conocidos métodos iterativos

### Método de Gauss–Jacobi

Consiste en despejar de la ecuación  $i$ -ésima la incógnita  $x_i$  ( $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ) lo cual es posible siempre que ningún elemento de la diagonal de  $A$  sea nulo ( $\forall k, a_{kk} \neq 0$ ). En caso de que alguno lo fuera permutaríamos las ecuaciones hasta conseguir una diagonal libre de *ceros*. Con el sistema escrito en esta forma utilizamos las coordenadas de la aproximación  $\mathbf{x}^k$  en el lado derecho del sistema para obtener las coordenadas de  $\mathbf{x}^{k+1}$  en el lado izquierdo. Escrito en coordenadas

$$\begin{aligned} x_1^{k+1} &= \frac{1}{a_{11}}(b_1 - (a_{12}x_2^k + a_{13}x_3^k \cdots + a_{1n}x_n^k)) \\ x_2^{k+1} &= \frac{1}{a_{22}}(b_2 - (a_{21}x_1^k + a_{23}x_3^k + \cdots + a_{2n}x_n^k)) \\ &\dots \quad \dots \\ x_n^{k+1} &= \frac{1}{a_{nn}}(b_n - (a_{n1}x_1^k + a_{n2}x_2^k + \cdots + a_{nn-1}x_{n-1}^k)). \end{aligned} \tag{28.3}$$

La descomposición de  $A$  equivalente a este método consiste en tomar como  $S$  la diagonal de  $A$  y entonces  $P = S - A$  está formada por los elementos que no están sobre la diagonal (cambiados de signo) en su posición y *ceros* en la diagonal.

*Ejemplo 25.* Resolución del sistema

$$\begin{pmatrix} 8 & 2 & 0 \\ 2 & 8 & 2 \\ 0 & 2 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \\ 8 \end{pmatrix}$$

por el método de Gauss-Jacobi:

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \\ 8 \end{pmatrix} - \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} - \begin{pmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Si arrancamos el método con el vector  $\mathbf{x}^0 = (0, 0, 0)$  sucesivas iteraciones nos dan

$$\begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} \quad \begin{pmatrix} \frac{3}{4} \\ 0 \\ \frac{3}{4} \end{pmatrix} \quad \begin{pmatrix} 1 \\ \frac{1}{4} \\ 1 \end{pmatrix} \quad \begin{pmatrix} \frac{7}{8} \\ 0 \\ \frac{7}{8} \end{pmatrix} \quad \begin{pmatrix} \frac{63}{64} \\ 0 \\ \frac{63}{64} \end{pmatrix} \quad \text{que convergen a} \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

### Método de Gauss–Seidel

Si en el método de Gauss–Jacobi al calcular las coordenadas de  $\mathbf{x}^{k+1}$  a partir de las coordenadas de  $\mathbf{x}^k$  según las fórmulas (28.3) utilizadas en el orden en que están escritas cambiamos en cada una de ellas las coordenadas de  $\mathbf{x}^k$  por las coordenadas de  $\mathbf{x}^{k+1}$  ya calculadas hasta ese momento obtenemos el método llamado de Gauss–Seidel. Su expresión en términos de las coordenadas de los dos vectores  $\mathbf{x}^k, \mathbf{x}^{k+1}$  es la siguiente

$$\begin{aligned} x_1^{k+1} &= \frac{1}{a_{11}}(b_1 - (a_{12}x_2^k + a_{13}x_3^k \cdots + a_{1n}x_n^k)) \\ x_2^{k+1} &= \frac{1}{a_{22}}(b_2 - (a_{21}x_1^{k+1} + a_{23}x_3^k + \cdots + a_{2n}x_n^k)) \\ &\dots \\ x_n^{k+1} &= \frac{1}{a_{nn}}(b_n - (a_{n1}x_1^{k+1} + a_{n2}x_2^{k+1} + \cdots + a_{nn-1}x_{n-1}^{k+1})). \end{aligned} \tag{28.4}$$

Para ver cual es su expresión en términos matriciales escribimos en los primeros miembros de las ecuaciones las coordenadas «nuevas» y en los segundos miembros las coordenadas «viejas»

$$\begin{aligned} a_{11}x_1^{k+1} &= b_1 - (a_{12}x_2^k + a_{13}x_3^k \cdots + a_{1n}x_n^k) \\ a_{21}x_1^{k+1} + a_{22}x_2^{k+1} &= b_2 - (a_{23}x_3^k + \cdots + a_{2n}x_n^k) \\ a_{31}x_1^{k+1} + a_{32}x_2^{k+1} + a_{33}x_3^{k+1} &= b_3 - (a_{34}x_4^k + \cdots + a_{3n}x_n^k) \\ &\dots \quad \dots \\ a_{n1}x_1^{k+1} + a_{n2}x_2^{k+1} + \cdots + a_{nn-1}x_{n-1}^{k+1} + a_{nn}x_n^{k+1} &= b_n. \end{aligned}$$

de forma que la descomposición correspondiente de la matriz  $A$  consiste en tomar como  $S$  la parte triangular inferior de la matriz  $A$ , incluida la diagonal, y como matriz  $P$  la parte triangular superior con ceros en la diagonal.

*Ejemplo 26.* Para el mismo sistema que en el Ejemplo 25, la descomposición de Gauss-Seidel es

$$\begin{pmatrix} 8 & 0 & 0 \\ 2 & 8 & 0 \\ 0 & 2 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \\ 8 \end{pmatrix} - \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

en el que tras invertir la primera matriz resulta

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{8} & 0 & 0 \\ -\frac{1}{32} & \frac{1}{8} & 0 \\ \frac{1}{128} & -\frac{1}{32} & \frac{1}{8} \end{pmatrix} \left( \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix} - \begin{pmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right)$$

es decir

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{4} \\ \frac{15}{16} \end{pmatrix} - \begin{pmatrix} 0 & \frac{1}{4} & 0 \\ 0 & -\frac{1}{16} & \frac{1}{4} \\ 0 & \frac{1}{64} & -\frac{1}{16} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

No merece la pena seguir haciendo el cálculo con fracciones. Es mejor aproximar éstas y seguir el procedimiento en coma flotante.

Un pequeño programa en MATLAB proporciona

$\mathbf{x}^0$	$\mathbf{x}^1$	$\mathbf{x}^2$	$\mathbf{x}^3$	$\mathbf{x}^4$	$\mathbf{x}^5$
1'0000	0'9375	0'9922	0'9990	0'9999	1'0000
0'2500	0'0313	0'0039	0'0005	0'0001	0'0000
0'9375	0'9922	0'9990	0'9999	1'0000	1'0000