
Métodos numéricos en álgebra lineal

PID_00266170

Roberto Casado Vara

Tiempo mínimo de dedicación recomendado: 3 horas



Roberto Casado Vara

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por la profesora: Teresa Sancho (2019)

Primera edición: septiembre 2019
© Roberto Casado Vara
Todos los derechos reservados
© de esta edición, FUOC, 2019
Av. Tibidabo, 39-43, 08035 Barcelona
Realización editorial: FUOC

Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea éste eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares del copyright.

Índice

Introducción	5
1. Error en los métodos numéricos	9
2. Nociones avanzadas de álgebra lineal: repaso	12
2.1. Matrices.....	12
2.2. Valores y vectores propios.....	13
2.3. Normas matriciales	14
3. Métodos de resolución de sistemas de ecuaciones lineales ..	18
3.1. Métodos directos.....	18
3.1.1. Eliminación gaussiana	18
3.1.2. Método de Gauss-Jordan	20
3.1.3. Descomposición LU	20
3.1.4. Descomposición QR	23
3.2. Métodos iterativos	24
3.2.1. Convergencia y estimación del error	25
3.2.2. Método iterativo de Jacobi.....	27
3.2.3. Método iterativo de Gauss-Seidel	30
3.2.4. Aplicación al álgebra lineal: cálculo de valores y vectores propios. Método de la potencia y de la potencia inversa.....	31
Bibliografía	35

Introducción

El objetivo de este material es servir como texto de apoyo en el curso de métodos numéricos en la ciencia de datos. Se incluye una breve introducción teórica que creemos que es suficiente, con referencias a manuales clásicos de métodos numéricos. Los problemas seleccionados aquí no van a cubrir todos los aspectos de los métodos numéricos relacionados con el álgebra lineal, pero motivarán su uso. Además, se os darán las herramientas necesarias para solucionar cualquier problema de álgebra lineal que se os pueda presentar, así como diseñar vuestros propios algoritmos en el lenguaje de programación R.

En disciplinas como la física, la ingeniería o la ciencia de datos, con frecuencia se utilizan los métodos numéricos para resolver problemas. Una vez terminado el proceso de discretización, en el que se sustituye el modelo continuo por una versión simplificada en dimensión finita, generalmente se tiene que resolver un sistema de ecuaciones que en la mayor parte de los casos es lineal. Los métodos de discretización más habituales son la interpolación de funciones y la aproximación con familias de funciones lineales. Aunque no lo trabajaremos en este material, la resolución de ecuaciones en derivadas parciales, tan habitual en sistemas dinámicos, también requiere la resolución de sistemas de ecuaciones lineales. Este es el motivo por el que la resolución numérica de sistemas de ecuaciones lineales es tan importante y nuclear aquí.

En esta primera actividad nos dedicaremos a encontrar los valores y vectores propios de una matriz cuadrada y de sus potencias, y a buscar la solución de un sistema $Ax = b$ en el que el número de ecuaciones es igual al número de incógnitas. Las técnicas que se utilizarán se pueden clasificar en métodos directos y métodos iterativos. En los métodos directos se encuentra la solución en un número finito de operaciones y la solución sería exacta si no cometiéramos errores al efectuar las operaciones. En los métodos iterativos se utiliza una relación de recurrencia que a partir de una estimación inicial permite determinar una sucesión de valores que converge a la solución que se busca.

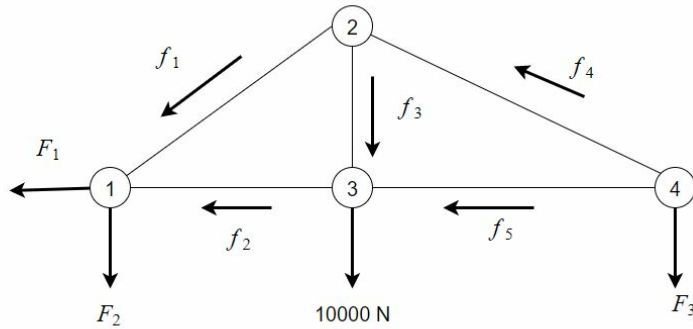
A continuación, vamos a presentar un ejemplo con un buen motivo por el que es importante la resolución de sistemas de ecuaciones lineales mediante métodos numéricos. Suponed que estáis trabajando en una empresa de calidad y os piden que, a partir de unos datos, comprobéis la fiabilidad de un puente. Vamos a presentar el problema en el siguiente ejemplo.

Ejemplo 1

Los armazones son estructuras ligeras capaces de soportar cargas pesadas. En el diseño de puentes, los miembros de la estructura están conectados con juntas rotatorias de pasador que permiten transferir las fuerzas de un miembro a otro. La figura siguiente muestra una estructura sencilla de un puente que se mantiene en equilibrio o estacionaria en el extremo inferior de la izquierda (1), que se desplaza horizontalmente en el extremo inferior derecho (4) y que tiene juntas de pasador en (1),(2),(3) y (4). Se coloca una carga de 10.000 newtons en la junta (5) y las fuerzas que actúan sobre los miembros de la estructura tienen magnitudes dadas por f_1, f_2, f_3, f_4 y f_5 , como se observa en la figura. El miembro de soporte estacionario tiene una fuerza horizontal F_1 y una fuerza vertical F_2 , pero el miembro de soporte móvil tiene una única fuerza vertical F_3 .

Probadlo vosotros mismos

Aunque ahora puede parecer muy complicado, en unas semanas seréis perfectamente capaces de resolverlo de forma muy sencilla; por tanto, os recomendamos que intentéis resolver este sistema con alguno de los métodos aprendidos.



Si la estructura está en equilibrio estático, las fuerzas en cada junta deben agregarse al vector cero, de modo que la suma de las componentes horizontales y verticales de las fuerzas en cada junta deben ser cero. Esto genera el sistema de ecuaciones lineales que expresamos a continuación:

Junta del puente	Componentes horizontales	Componentes verticales
1	$-F_1 + \frac{\sqrt{2}}{2}f_1 + f_2 = 0$	$\frac{\sqrt{2}}{2}f_1 - F_2 = 0$
2	$\frac{\sqrt{2}}{2}f_1 + \frac{\sqrt{3}}{2}f_4 = 0$	$\frac{\sqrt{2}}{2}f_1 - f_3 + \frac{1}{2}f_4 = 0$
3	$-f_2 + f_5 = 0$	$f_3 - 10000 = 0$
4	$-\frac{\sqrt{3}}{2}f_4 - f_5 = 0$	$\frac{1}{2}f_4 - F_3 = 0$

Una vez que se tienen todas estas fuerzas calculadas, es necesario resolver ese sistema de ecuaciones. Con ayuda de algún ingeniero de la empresa en la que estéis, podréis poner ese sistema lineal en forma de matriz (se ha reordenado el sistema para ilustrar mejor el ejemplo):

$$\begin{pmatrix}
 -1 & 0 & 0 & \frac{\sqrt{2}}{2} & 1 & 0 & 0 & 0 \\
 0 & -1 & 0 & \frac{\sqrt{2}}{2} & 0 & 0 & 0 & 0 \\
 0 & 0 & -1 & 0 & 0 & 0 & \frac{1}{2} & 0 \\
 0 & 0 & 0 & -\frac{\sqrt{2}}{2} & 0 & -1 & \frac{1}{2} & 0 \\
 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & -\frac{\sqrt{2}}{2} & 0 & 0 & -\frac{\sqrt{3}}{2} & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{3}}{2} & -1
 \end{pmatrix}
 \begin{pmatrix}
 F_1 \\
 F_2 \\
 F_3 \\
 f_1 \\
 f_2 \\
 f_3 \\
 f_4 \\
 f_5
 \end{pmatrix}
 =
 \begin{pmatrix}
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 10000 \\
 0 \\
 0
 \end{pmatrix}
 \tag{1}$$

Obviamente, intentar resolver este sistema de ecuaciones puede ocupar muchas horas de trabajo. Y aunque alguno de vosotros se animara a hacer las cuentas de forma manual, en la realidad estos sistemas de ecuaciones suelen ser de 100×100 ecuaciones.

Este ejemplo demuestra la necesidad de buscar una forma de resolver este tipo de sistemas de una manera adecuada que no requiera muchas horas de cálculos. Para ello, podemos emplear los métodos numéricos; una vez que se ha decidido el método y se ha programado, bastaría con introducir la matriz y ejecutar el código. En cuestión de minutos tendríamos una solución, mientras que si lo intentásemos calcular a mano podríamos tardar horas, o incluso no podríamos hacerlo.

1. Error en los métodos numéricos

Los cálculos numéricos que se efectúan con una calculadora u ordenador no son iguales a los que se realizan en los cursos de álgebra o de cálculo con papel y lápiz. De nuestras experiencias pasadas, esperamos tener siempre expresiones tan verdaderas como $2 + 2 = 4$ y $(\sqrt{3})^2 = 3$. Pero lo cierto es que en la aritmética computacional estándar (es decir, la aritmética de punto flotante) tendremos algunos casos en los que las operaciones no van a ser “tan verdaderas”. En efecto, el segundo de los ejemplos no es verdad del todo. Para entender qué es lo que ocurre, tendremos que hacer una breve introducción al mundo de la aritmética computacional.

En las matemáticas tradicionales puede haber números con una cantidad infinita de dígitos no periódicos. La aritmética que se emplea define $\sqrt{3}$ como el único número positivo que cuando se multiplica por sí mismo tiene como resultado el número entero 3. Sin embargo, en el mundo computacional, todo número es representado como un número fijo y finito de dígitos. Puesto que $\sqrt{3}$ no tiene una representación de dígitos finitos, en el interior del ordenador se le da una representación aproximada cuyo cuadrado no será exactamente 3, aunque con toda probabilidad estará bastante cerca del 3 y resultará aceptable en la gran mayoría de los casos. Casi siempre, la representación matemática del álgebra o del análisis y la aritmética computacional son satisfactorias y pasan inadvertidas o sin problemas, es decir, dejamos que el ordenador o la calculadora decidan por defecto el nivel de precisión que tendrá el número decimal.

Pero recordemos que no siempre es así, y deberíamos estar alerta ante los problemas que pueda ocasionar en nuestros cálculos. Supongamos que estamos haciendo cálculos para predecir las ventas a nuestros clientes (caso bastante probable en vuestro futuro laboral), y estamos trabajando con cifras muy grandes, dependiendo del volumen del cliente. El “baile” de algunas cifras pueden ocasionar grandes dolores de cabeza en los directivos de la empresa, ya que no cuadrarán las cuentas previstas por vuestros informes con la realidad. Es decir, el error entre vuestras predicciones y la realidad no será aceptable.

Por tanto, el **error de redondeo** surge cuando usamos una calculadora u ordenador para efectuar un cálculo. El error surge porque las operaciones aritméticas realizadas en una máquina incluyen un número finito de dígitos, de manera que los cálculos se llevan a cabo con representaciones aproximadas de los números reales. En un ordenador normal, solo un conjunto relativamente pequeño de números reales se utiliza para representar a todos los números.

Comprobadlo vosotros mismos

En el siguiente enlace podréis comprobar vosotros mismos cómo en algunos casos los errores de precisión pueden afectar a las operaciones matemáticas.
<https://support.microsoft.com/es-es/help/78113/floating-point-arithmetic-may-give-inaccurate-results-in-excel>

Por tanto, el error resultante de reemplazar un número por su representación decimal finita que hacen las calculadoras y los ordenadores recibe el nombre de **error de redondeo**. A continuación, vamos a introducir dos formas de calcular la precisión de una solución calculada con los métodos numéricos. Lo primero que necesitamos saber es que la solución que tenemos calculada es lo que vamos a llamar aproximación (ya que no sabemos si es la solución exacta). Por tanto, si \tilde{p} es una aproximación de un número cualquiera p , el **error absoluto** es $|p - \tilde{p}|$, y el **error relativo** es $\frac{|p - \tilde{p}|}{|p|}$ siempre y cuando $p \neq 0$.

Para ilustrar esta definición vamos a poner un ejemplo.

Ejemplo 2

Tenemos el número $p = 179,015625$, y vamos a considerar dos aproximaciones que ha realizado el ordenador a este número, $\tilde{p}_1 = 179,01560974$ y $\tilde{p}_2 = 179,01564025$. Podríamos preguntarnos cuál de las dos aproximaciones es mejor, pero ¿cómo podríamos medir o validar cuál de ellas es mejor? La solución es sencilla, solamente tenemos que usar los errores relativo y absoluto.

En este caso, el error absoluto es:

$$\begin{aligned} |p - \tilde{p}_1| &= |179,015625 - 179,01560974| = 1,526 \times 10^{-05} \\ |p - \tilde{p}_2| &= |179,015625 - 179,01564025| = 1,525 \times 10^{-05} \end{aligned} \quad (2)$$

y el error relativo es:

$$\begin{aligned} \frac{|p - \tilde{p}_1|}{|p|} &= \frac{|179,015625 - 179,01560974|}{|179,015625|} = 8,524396 \times 10^{-08} \\ \frac{|p - \tilde{p}_2|}{|p|} &= \frac{|179,015625 - 179,01564025|}{|179,015625|} = 8,518809 \times 10^{-08} \end{aligned} \quad (3)$$

Ejemplo 3

En este ejemplo vamos a ver que, al realizar operaciones con dos números cualesquiera, es importante tener en cuenta el nivel de precisión que se quiere, es decir, el número de cifras decimales con el que se desea trabajar. Como se ve en este ejemplo, según el número de cifras decimales que se eligen, el error puede ser muy pequeño o muy relevante, lo que provoca que los cálculos realizados puedan no tener el grado de precisión deseado (es decir, suponed que estos números son millones, un error en esos números puede suponer pérdidas grandes para vuestra empresa).

Si tenemos $p = 0,54617$ y $q = 0,54601$, el valor exacto de restar p y q es: $r = p - q = 0,00016$. Supongamos que la resta se efectúa con cuatro cifras decimales. Al redondear p y q a cuatro dígitos tenemos que $\tilde{p} = 0,5462$ y $\tilde{q} = 0,5460$, respectivamente. Por tanto, $\tilde{r} = \tilde{p} - \tilde{q} = 0,0002$ es la aproximación a 4 dígitos de r .

Ya que

$$\frac{|\tilde{r} - r|}{|\tilde{r}|} = \frac{|0,0002 - 0,00016|}{|0,0002|} = 0,25 \quad (4)$$

el resultado tiene un solo dígito significativo, mientras que \tilde{p} y \tilde{q} eran exactos en los primeros dígitos.

Consejo

Como una medida de la exactitud, el error absoluto no es tan significativo como el error relativo, además puede ser engañoso. Por ejemplo, un error de un gramo es mucho más significativo cuando se calcula la masa de un reactivo químico, que cuando se calcula la masa de un avión.

El error es un elemento muy útil cuando se utilizan métodos numéricos. Muchas veces, lo más importante cuando se realizan cálculos es el error que nos podemos permitir. No es lo mismo tener un error de un metro en la precisión de un GPS que tener un error de un metro midiendo la longitud de una parcela urbana. Por tanto, un buen consejo para vuestro futuro es que seleccionéis el error que “os podéis permitir” al realizar los cálculos o la tolerancia, y uséis este dato como medida de parada en los métodos iterativos para la resolución de los sistemas de ecuaciones lineales.

2. Nociones avanzadas de álgebra lineal: repaso

Para resolver las actividades de esta asignatura es importante dominar los conceptos básicos del álgebra lineal, por lo que recordaremos algunos conceptos básicos que habréis estudiado en asignaturas anteriores, así como la nomenclatura que utilizaremos.

2.1. Matrices

Denotaremos por $\mathbb{M}_{m \times n}(\mathbb{K})$ al espacio vectorial de las matrices de m filas y n columnas cuyos coeficientes del conjunto \mathbb{K} son generalmente números reales (\mathbb{R}) o números complejos (\mathbb{C}). En esta asignatura solo trabajaremos en el conjunto de los reales. Denotaremos por A^T a la matriz $n \times m$ traspuesta de A , por A^* a su matriz conjugada y por A^{ad} a su matriz adjunta (traspuesta de la conjugada).

A continuación, se repasan algunas de las definiciones del álgebra lineal que resultarán útiles durante este material:

- La matriz A es simétrica si $A = A^T$.
- La matriz $A \in \mathbb{M}_{n \times m}(\mathbb{R})$ es unitaria si $A^{-1} = A^{ad}$.
- La matriz $A \in \mathbb{M}_{n \times m}(\mathbb{R})$ es normal si $A^{ad}A = AA^{ad}$.

Por simplicidad de la notación, a partir de este momento vamos a suponer que A es una matriz cuadrada $n \times n$.

Para algunos de los cálculos que vamos a realizar con los métodos de resolución numérica de sistemas de ecuaciones lineales vamos a necesitar algunas suposiciones teóricas que nos faciliten los cálculos. Aunque muchas de estas afirmaciones ya son conocidas por todos vosotros, las recordaremos con el fin de afianzar estos conocimientos. Siendo $A \in \mathbb{M}_{n \times m}(\mathbb{K})$, las siguientes afirmaciones son equivalentes:

- A^{-1} existe
- $\det(A) \neq 0$
- El sistema lineal $Ax = 0$ tiene solamente la solución $x = 0$.
- Para cualquier vector b , el sistema lineal $Ax = b$ tiene solución única.
- Las filas y columnas de A son linealmente independientes.
- El rango de la matriz A es n .

2.2. Valores y vectores propios

Supongamos que tenemos A una matriz cuadrada $n \times n$. Vamos ahora a definir una serie de elementos que serán útiles para trabajar con las matrices. El objetivo de esta sección es dar los detalles matemáticos necesarios para entender el marco teórico que hay detrás de las operaciones que realizan los métodos numéricos cuando calculan valores propios, vectores propios y potencias de matrices, entre otros.

- El espectro de A es el conjunto $Esp(A) \subset \mathbb{R}$ descrito por los valores propios o autovalores de A .
- El radio espectral de A es el número real positivo

$$\rho(A) = \max |\lambda_i| \quad (5)$$

donde λ_i son los valores propios de la matriz A .

- Una pareja (λ, x) es un elemento propio de A si x es un vector propio de A asociado al valor propio λ .

A continuación, se presentan algunos resultados importantes que nos van a aportar la evidencia matemática necesaria para entender cómo funcionan los métodos numéricos para resolver sistemas de ecuaciones lineales. También nos aportarán la justificación matemática de los métodos iterativos para manejar matrices de grandes dimensiones.

Dada una matriz $A \in \mathbb{M}_{n \times n}(\mathbb{R})$, las siguientes afirmaciones son ciertas (pero no equivalentes):

- 1) Los valores propios de una matriz simétrica son reales.
- 2) Si (λ, x) es un elemento propio de A , entonces, para cualquier entero positivo m , (λ^m, x) es un elemento propio de A^m .

De esta afirmación sacamos una importante conclusión: los valores propios de una matriz cuadrada A son los mismos que los valores propios de las potencias de A . Este es un resultado muy potente relacionado con el cálculo de potencias de matrices.

Ahora vamos a definir el concepto de matrices semejantes. Se dice que dos matrices cuadradas A y B son **semejantes** si existe una matriz regular tal que:

$$B = PAP^{-1} \quad (6)$$

Con esta definición podemos enunciar el siguiente resultado, que nos dará las bases teóricas para entender el marco matemático que hay detrás de los cálculos relacionados con las potencias de matrices. Dada una matriz $A \in M_{n \times n}(\mathbb{R})$, las siguientes afirmaciones son ciertas:

- 1) La matriz A es semejante a una matriz diagonal si y solo si tiene n vectores propios linealmente independientes.
- 2) Si la matriz A tiene n autovalores distintos, entonces es semejante a una matriz diagonal.

2.3. Normas matriciales

En el siguiente apartado vamos a explicar los métodos iterativos para obtener las soluciones de sistemas de ecuaciones de la forma $Ax = b$. Antes de empezar con el estudio de dichos métodos, necesitamos contar primero con un medio que nos permita medir la distancia entre los vectores columna de dimensión n . Ello nos permitirá determinar si una serie de estos vectores convergen a una solución del sistema (en el próximo apartado se explicará el concepto de convergencia). Esta "medida" también es muy necesaria cuando la solución se obtiene por los métodos directos que explicaremos en el próximo apartado.

Denotemos \mathbb{R}^n el conjunto de todos los vectores columna de dimensión n , cuyas componentes son números reales. Para definir una distancia en \mathbb{R}^n introduciremos el concepto de norma.

Una *norma vectorial* en \mathbb{R}^n es una función $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ con las siguientes propiedades:

- 1) $\|x\| \geq 0$ para todo $x \in \mathbb{R}^n$
- 2) $\|x\| = 0$ si y solo si $x = (0, 0, \dots, 0)$
- 3) $\|\alpha x\| = |\alpha| \cdot \|x\|$ para todo $\alpha \in \mathbb{R}$ y $x \in \mathbb{R}^n$
- 4) $\|x + y\| \leq \|x\| + \|y\|$ para todo $x, y \in \mathbb{R}^n$

En este material, solo usaremos dos normas específicas en \mathbb{R}^n , que son $l_2 = \|\cdot\|_2$ y $l_\infty = \|\cdot\|_\infty$, que se definen como sigue:

$$\|x\|_2 = \left\{ \sum_{i=1}^n x_i^2 \right\}^{1/2} \quad \text{y} \quad \|x\|_\infty = \max\{|x_i|\} \quad (7)$$

La norma l_2 se llama norma euclidiana del vector x , dado que representa la noción común de distancia respecto del origen en caso de que x estén en \mathbb{R} , \mathbb{R}^2 o en \mathbb{R}^3 . Por ejemplo, el vector $x = (-1, 1, -2)$ en \mathbb{R}^3 tiene las siguientes normas:

Notación

Como los vectores en \mathbb{R}^n son vectores columna, conviene utilizar la notación de la transpuesta que se ha visto en el material de álgebra lineal o se puede consultar en las referencias recomendadas, aunque para facilitar la notación no indicaremos la "t" de transpuesta en los vectores.

$$\begin{aligned} \|x\|_2 &= \sqrt{(-1)^2 + 1^2 + (-2)^2} = \sqrt{6} \\ \|x\|_\infty &= \max\{|-1|, |1|, |-2|\} = 2 \end{aligned} \quad (8)$$

Si $x = (x_1, x_2, \dots, x_n)$ y $y = (y_1, y_2, \dots, y_n)$ son vectores de \mathbb{R}^n , las distancias l_2 y l_∞ entre x e y están definidas por:

$$\|x - y\|_2 = \left\{ \sum_{i=1}^n (x_i - y_i)^2 \right\}^{1/2} \quad \text{y} \quad \|x - y\|_\infty = \max\{|x_i - y_i|\} \quad (9)$$

Ejemplo 4

Dado el sistema lineal

$$\begin{aligned} 3,3330x_1 + 15920x_2 - 10,333x_3 &= 15913 \\ 2,2220x_1 + 1,710x_2 + 8,6120x_3 &= 28,544 \\ 1,5611x_1 + 5,1791x_2 + 1,6852x_3 &= 8,4254 \end{aligned} \quad (10)$$

tiene la solución $(x_1, x_2, x_3) = (1, 1, 1)$. Si efectuamos la eliminación gaussiana en la aritmética de redondeo a cinco dígitos utilizando técnicas clásicas de resolución de sistemas de ecuaciones (por ejemplo, el método de gauss), la solución es:

$$\tilde{x} = (1,2001; 0,99991; 0,92538) \quad (11)$$

Así que, calculando la diferencia entre los vectores x y \tilde{x} en las dos normas que hemos definido se tiene:

$$\begin{aligned} \|x - \tilde{x}\|_\infty &= \max\{|1 - 1,20001|, |1 - 0,99991|, |1 - 0,92538|\} = 0,2001 \\ \|x - \tilde{x}\|_2 &= \sqrt{(1 - 1,20001)^2 + (1 - 0,99991)^2 + (1 - 0,92538)^2} = 0,21356 \end{aligned} \quad (12)$$

En este ejemplo, se puede ver que las aproximaciones de \tilde{x}_2 y \tilde{x}_3 son buenas aproximaciones de x_2 y x_3 , pero que \tilde{x}_1 es una mala aproximación de x_1 .

El concepto de distancia en \mathbb{R}^n también sirve para definir el límite de una sucesión de vectores en ese espacio. Aquí es importante entender que si una sucesión de vectores tiende a un vector, se dice que esa sucesión converge. Es decir, si se tiene una sucesión de vectores que a medida que miramos los términos más grandes de la sucesión se aproximan a un vector x , en otras palabras, los vectores de esa sucesión se aproximan mucho al valor del vector x , se dice que esa sucesión de vectores es convergente. Esto va a resultar muy útil cuando estemos resolviendo sistemas de ecuaciones lineales por métodos iterativos. Estos métodos van a proporcionar una sucesión de soluciones y lo que buscamos es que esa sucesión de soluciones a la larga se aproxime (rápido o despacio) a la solución. Esto es lo mismo que decir que queremos que esa sucesión de soluciones converja a la solución real.

Una vez hecha esta breve introducción al concepto de norma de un vector, estamos en condiciones de aplicar esta misma teoría a las matrices, salvo algunos matices que trataremos con cuidado mientras definimos las normas matriciales y su propiedad.

Primero, es importante decir que puede demostrarse que todas las normas \mathbb{R}^n son equivalentes respecto a la convergencia. Esto es importante, ya que cuando ahora definamos el concepto de normas de matrices, y más adelante el concepto de convergencia de sucesiones de matrices respecto de alguna norma, en algunos casos, será más fácil calcular alguna de las normas que otras. Dado cualquier vector $x \in \mathbb{R}^n$,

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty \quad (13)$$

De este modo, podemos demostrar que cualquier vector que converja respecto de la norma $\|\cdot\|_\infty$ también lo hace respecto de la norma $\|\cdot\|_2$, y esto nos quiere decir que a partir de ahora nos será indiferente qué norma utilizar. En la práctica, se suele usar la norma que más fácil sea de calcular.

Si tenemos en cuenta que queremos resolver sistemas de ecuaciones lineales, y que estos se pueden expresar como matrices, es natural introducir el concepto de normas matriciales.

Una **norma matricial** sobre el conjunto de todas las matrices $n \times n$ es una función de valor real, $\|\cdot\|$, definida en este conjunto y que satisface para todas las matrices A y B de $n \times n$, y todos los números reales α :

- 1) $\|A\| \geq 0$
- 2) $\|A\| = 0$, si y solo si A es 0, la matriz con todas las posiciones cero.
- 3) $\|\alpha A\| = |\alpha| \|A\|$
- 4) $\|A + B\| \leq \|A\| + \|B\|$
- 5) $\|AB\| \leq \|A\| \|B\|$

Una distancia entre las matrices A y B de $n \times n$ respecto a esta norma matricial es $\|A - B\|$.

Aunque las normas matriciales se pueden obtener de varias formas, en este material tenemos interés en las normas que son consecuencia natural de las normas vectoriales, es decir, l_2 y l_∞ . Como ejemplo, se tiene el siguiente resultado:

Si $\|\cdot\|$ es una norma vectorial de \mathbb{R}^n , entonces

$$\|A\| = \max\{\|Ax\|\} \quad (14)$$

es una norma matricial.

Se trata de una norma matricial natural o inducida asociada a una norma vectorial. En este material, vamos a suponer que todas las normas matriciales son naturales, si no especificamos lo contrario en sus definiciones. El siguiente resultado es muy útil cuando se quiere acotar el valor de $\|Ax\|$.

Para todo vector $x \neq 0$, matriz A y cualquier norma natural $\|\cdot\|$, tenemos

$$\|Ax\| \leq \|A\| \cdot \|x\| \quad (15)$$

Este es un muy buen resultado, ya que nos permitirá conocer con cierta exactitud el valor aproximado de $\|Ax\|$.

Con estas definiciones y resultados ya estamos en condiciones de empezar a definir y explicar los métodos de resolución de sistemas de ecuaciones lineales. Definiremos los métodos directos en el siguiente apartado y con los conocimientos que tenéis hasta el momento podemos afrontarlos con total seguridad. Sin embargo, los métodos iterativos son un poco más delicados y requieren el concepto de convergencia de sucesiones de matrices. Este concepto es necesario para definir la prueba de parada y encontrar las soluciones óptimas sin que el método itere sin control creando bucles infinitos.

3. Métodos de resolución de sistemas de ecuaciones lineales

Los sistemas de ecuaciones lineales se utilizan en muchos problemas de ingeniería y de las ciencias, así como en aplicaciones de las matemáticas a las ciencias sociales y al estudio cuantitativo de problemas de administración y economía. En este apartado se describirán los métodos directos e iterativos para resolver el sistema lineal

$$\begin{aligned} E_1 : & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ E_2 : & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ & \vdots \quad \quad \quad \vdots \\ E_n : & a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{aligned} \tag{16}$$

para x_1, \dots, x_n , dadas las a_{ij} con $i, j = 1, 2, \dots, n$ y b_i , para $i = 1, 2, \dots, n$. Estas técnicas son métodos que proporcionan un resultado en un número fijo de pasos, y solo están sujetos a errores de redondeo.

3.1. Métodos directos

Queda claro que, aunque sepamos resolver sistemas lineales, la solución numérica obtenida puede tener errores significativos respecto al resultado esperado. Estudiaremos cómo podemos resolver del modo más rápido y preciso el sistema en cuestión, y comenzaremos con los métodos directos.

3.1.1. Eliminación gaussiana

Entre los métodos directos, el más popular para resolver sistemas lineales es la eliminación gaussiana, que también se utiliza para calcular determinantes e invertir matrices. La idea básica del método, que ilustramos con un ejemplo, consiste en utilizar transformaciones elementales de fila y columna para eliminar sucesivamente las variables empezando por la primera ecuación y la primera variable y continuando con el resto. De esta manera, tras $(n - 1)$ eliminaciones se llega a un sistema equivalente al dado, de matriz triangular superior, que se resuelve directamente por sustitución hacia atrás. Veamos cómo funciona:

Fuente del apartado

Apartado basado en la obra siguiente: *Numerical Analysis: Mathematics of Scientific Computing* de David Kinkaid. Amer Mathematical Society. Standard No. 9781470411152.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & 3 & 4 & 9 \\ -1 & 0 & -1 & -2 \end{array} \right) \quad (17)$$

haciendo transformaciones elementales utilizando la primera fila, hacemos cero todos los elementos de la primera columna excepto el de la diagonal.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -1 & -2 & -3 \\ 0 & 2 & 2 & 4 \end{array} \right) \quad (18)$$

Fijándonos en el elemento diagonal de la segunda fila, hacemos cero todos los elementos de la segunda columna por debajo de la diagonal.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & -2 & -2 \end{array} \right) \quad (19)$$

Con esto tenemos ya el sistema triangular superior equivalente, que resolvemos por sustitución hacia atrás comenzando por la última ecuación y acabando en la primera. De donde se tiene que $x_1 = 1$, $x_2 = 1$ y $x_3 = 1$.

En la eliminación gaussiana tal y como la acabamos de exponer se asume que en la k -ésima eliminación el coeficiente de la variable que se desea eliminar, que se llama pivote y que ocupa la posición (k, k) de la matriz del sistema en ese momento, es distinto de cero. A menudo no sucede así, por lo que se aconseja reordenar las ecuaciones e incluso los términos en cada una de ellas para lograr, por razones de estabilidad numérica, que el pivote sea el mayor posible en valor absoluto.

Existen diferentes estrategias para la elección del pivote, según si la reordenación afecta solo a las filas (pivotación parcial) o tanto a las filas como a las columnas (pivotación total). En principio, podría parecer más conveniente la pivotación total, pero su alto coste numérico, ya que exige en cada eliminación la comparación de todos los elementos de la matriz, hace preferible en la práctica la pivotación parcial con equilibrado de filas y columnas. Este equilibrado tiene como objeto normalizarlas, lo que se consigue multiplicando todos sus elementos por números convenientes.

La eliminación gaussiana es una de las técnicas más simples pero a la vez de las más efectivas. Si os fijáis, una vez puesto el sistema lineal de ecuaciones en su forma matricial, solo tenemos que conseguir transformar esa matriz en una matriz diagonal superior (es decir, que en la parte que está por debajo de la diagonal solo encontremos ceros) y a partir de ahí resolver hacia atrás, esto es, primero tenemos la última incógnita y la sustituimos en el resto de las ecuaciones y así sucesivamente hasta tener el resultado del sistema lineal de ecuaciones.

3.1.2. Método de Gauss-Jordan

El método de Gauss-Jordan es el método directo óptimo para encontrar la inversa de una matriz cuando esta no tiene ninguna estructura particular. En el método de Gauss-Jordan se dispone la matriz cuadrada A de $n \times n$ que se desea invertir a la izquierda y la matriz unidad I_n a su derecha. Se realizan sucesivas eliminaciones gaussianas mediante transformaciones elementales de fila y columna hasta tener a la izquierda la matriz unidad I_n , en cuyo caso, a la derecha tendremos la inversa A^{-1} . Veamos con el mismo ejemplo cómo se calcula la inversa,

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1,5 & 1 & -0,5 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1,5 & -1 & -0,5 \end{array} \right) \quad (20)$$

3.1.3. Descomposición LU

El método de descomposición LU es una variante del método de Gauss y consiste en factorizar la matriz A del sistema lineal que se quiere resolver en dos matrices, una triangular inferior, que llamaremos L , y otra triangular superior, que llamaremos U . Una vez obtenida la descomposición, se encuentra la solución resolviendo el sistema lineal:

$$(LU)x = L(Ux) = b \quad (21)$$

Resolviendo sucesivamente los dos sistemas lineales triangulares siguientes:

$$\begin{aligned} Ly &= b \\ Ux &= y \end{aligned} \quad (22)$$

Existen muchas formas de realizar esta descomposición. Para este material explicaremos una de las más sencillas, el algoritmo de Crout, que funciona siempre que todos los menores principales de A , es decir, las submatrices de A , sean distintos de 0. En esta descomposición se supone que la matriz triangular superior U tiene elementos unidad en la diagonal principal.

La descomposición LU es una técnica de descomposición de matrices muy importante. Y, de hecho, muchos métodos de resolución de ecuaciones lineales y no lineales (no entran en este material) se van a basar en hacer descomposiciones LU, entre otras cosas, para resolverlos. Comprender bien esta técnica os facilitará la comprensión de algunos de los métodos que estudiaremos a continuación.

Ejemplo 5. Descomposición LU

Explicemos este método con un ejemplo práctico.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ -1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix} \quad (23)$$

Factorización LU

En este ejemplo se puede ver de forma práctica cómo calcular las matrices L y U.

Para calcular estos coeficientes procedemos sucesivamente por identificación:

$$\begin{aligned} l_{11} &= 1; & l_{21} &= 2; \\ l_{11}u_{12} &= 2 \rightarrow u_{12} = 2; & l_{21}u_{12} + l_{22} &= 3 \rightarrow u_{12} = 2; \\ l_{11}u_{13} &= 3 \rightarrow u_{13} = 3; & l_{21}u_{13} + l_{22}u_{23} &= 4 \rightarrow u_{23} = 2; \\ l_{31} &= -1 \\ l_{31}u_{12} + l_{32} &= 3; \rightarrow l_{32} &= 2; \\ l_{31}u_{13} + l_{32}u_{23} + l_{33} &= -1 \rightarrow l_{33} &= -2; \end{aligned} \quad (24)$$

Resolviendo este sistema se obtienen los valores de l_{ij} y de u_{ij} . Una vez obtenida la factorización, resolvemos los dos sistemas triangulares.

El primero de ellos $Ly = b$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ -2 \end{pmatrix} \quad (25)$$

de donde se tiene que $y_1 = 6$, $y_2 = 3$ y por último, $y_3 = 1$.

El segundo sistema, $Ux = y$, se resuelve por sustitución hacia atrás.

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix} \quad (26)$$

de donde se tiene que $x_1 = 1$, $x_2 = 1$ y, por último, $x_3 = 1$.

Si la matriz A del sistema no es simétrica, o con muchos ceros, etc., y no converge por otros métodos iterativos, la descomposición LU es el método más aconsejable.

Ejemplo 6. Descomposición LU

Os invitamos a que comprobéis los cálculos de la resolución del siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x_1 + x_2 + 3x_4 &= 4 \\ 2x_1 + x_2 - x_3 + x_4 &= 1 \\ 3x_1 - x_2 - x_3 + 2x_4 &= -3 \\ -x_1 + 2x_2 + 3x_3 - x_4 &= 4 \end{aligned} \quad (27)$$

La secuencia de operaciones $(E_2 - 2E_1) \rightarrow (E_2)$, $(E_3 - 3E_1) \rightarrow (E_3)$, $(E_4 - (-1)E_1) \rightarrow (E_4)$, $(E_3 - 4E_2) \rightarrow (E_3)$, $(E_4 - (-3)E_2) \rightarrow (E_4)$ lo convierte en el sistema triangular:

$$\begin{aligned} x_1 + x_2 + 3x_4 &= 4 \\ -x_2 - x_3 - 5x_4 &= -7 \\ 3x_3 + 13x_4 &= 13 \\ -13x_4 &= -13 \end{aligned} \quad (28)$$

De esta forma, la matriz A se descompone en su factorización LU :

$$A = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} = LU \quad (29)$$

Esta factorización nos permite resolver fácilmente el sistema que contiene la matriz A . Por ejemplo, para resolver:

$$Ax = LUx = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 8 \\ 7 \\ 14 \\ -7 \end{pmatrix} \quad (30)$$

primero introducimos la sustitución $y = Ux$. Luego $Ly = b$, es decir,

$$LUx = Ly = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 8 \\ 7 \\ 14 \\ -7 \end{pmatrix} \quad (31)$$

Este sistema se resuelve para y mediante un simple proceso de sustitución hacia atrás, el resultado es $y_1 = 8$, $y_2 = -9$, $y_3 = 26$ y, por último, $y_4 = -26$. Y entonces resolvemos $Ux = y$ para x , o sea, la solución del sistema original, es decir,

$$LUx = Ly = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 8 \\ -9 \\ 26 \\ -26 \end{pmatrix} \quad (32)$$

y, al emplear la sustitución hacia atrás, obtenemos $x_4 = 2$, $x_3 = 0$, $x_2 = -1$ y $x_1 = 3$.

3.1.4. Descomposición QR

Algunas veces, nos podemos encontrar con que hemos de resolver un sistema de n ecuaciones lineales con m incógnitas, es decir, un sistema $m \times n$. En este caso, y sin entrar en mucho detalle en este material, basta con que se conozca que la técnica usual es buscar un vector x al que llamaremos solución por mínimos cuadrados (se puede consultar en los libros de referencia). En estos casos, se utiliza la técnica conocida como descomposición QR.

La descomposición QR de una matriz $A_{m \times n}$ consiste en expresar la matriz A como producto de dos matrices, Q y R , con $Q_{m \times n}$ ortogonal y R triangular superior. Esta descomposición se usa para resolver sistemas $m \times n$ con $m \leq n$ en el sentido de mínimos cuadrados.

Aclaración

Este método funciona especialmente bien para matrices simétricas tridiagonales. En caso de que la matriz sea simétrica, se puede transformar con operaciones lineales en tridiagonal.

1) Resolved el sistema lineal $Ax = b$ mediante la descomposición que estiméis oportuna:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & -1 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad (33)$$

solución: $x = (1, 1, -2, -1)$

2) Resolved el sistema lineal $Ax = b$ mediante la descomposición que estiméis oportuna:

$$\begin{pmatrix} 4 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 100 \\ 200 \\ 200 \\ 200 \\ 100 \end{pmatrix} \quad (34)$$

solución:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 46,153 \\ 84,615 \\ 92,307 \\ 84,6152 \\ 46,153 \end{pmatrix} \quad (35)$$

3.2. Métodos iterativos

Un método iterativo para resolver el sistema lineal $Ax = b$ comienza con una aproximación inicial $x^{(0)}$ a la solución x y genera una sucesión de vectores $\{x^{(k)}\}_{k=0}^{\infty}$ que converge en x . Los métodos iterativos conllevan un proceso que convierte el sistema $Ax = b$ en otro equivalente de la forma $x = Tx + c$ para alguna matriz fija T y un vector c . Después de seleccionar el valor inicial $x^{(0)}$, la sucesión de vectores de la solución se genera de la manera siguiente:

$$x^{(k)} = Tx^{(k-1)} + c \quad (36)$$

para $k = 1, 2, \dots$. Este resultado es muy potente, ya que permite resolver sistemas complejos con un grado de precisión tan alto como se quiera, es decir, se puede elegir el error que se quiere cometer.

Los métodos iterativos rara vez se usan para resolver sistemas lineales de pequeña dimensión, ya que el tiempo necesario para conseguir una exactitud satisfactoria rebasa el que requieren los métodos directos explicados en el apartado anterior. Sin embargo, en el caso de sistemas grandes con un alto porcentaje de ceros, son eficientes tanto en almacenamiento de computadora como en el tiempo de cómputo. Este tipo de sistema se presenta constantemente en el análisis de circuitos y en la solución numérica de los problemas con valor en la frontera de las ecuaciones diferenciales parciales (no es parte del material, pero se comenta para ilustrar algunos ejemplos de uso).

Todos los métodos iterativos que estudiaremos para resolver el sistema lineal $Ax = b$ se basan en una descomposición de la matriz A del tipo $A = M - N$ con M invertible. El sistema lineal se escribe entonces:

$$Mx = Nx + b \rightarrow x = (M^{-1}N)x + M^{-1}b \quad (37)$$

que escribimos $x = Tx + c$ con $T = M^{-1}N$ y $c = M^{-1}b$.

Se define así el esquema iterativo de punto fijo:

$$Mx^{(k+1)} = Nx^{(k)} + b \rightarrow x^{(k+1)} = (M^{-1}N)x^{(k)} + M^{-1}b \quad (38)$$

Y de acuerdo con los resultados que hemos ido dando sobre la convergencia, este método será convergente si y solo si $\rho(M^{-1}N) < 1$ (recordemos que $\rho(A)$ es el radio espectral de la matriz A).

3.2.1. Convergencia y estimación del error

Cuando se programa un método iterativo para la resolución de algún sistema de ecuaciones lineales, una de las principales preguntas que uno se puede hacer es: ¿cuándo se va a detener este método? Pues la respuesta es: cuando sea tan preciso como queramos. Esta afirmación plantea una nueva pregunta: ¿cuándo la solución va a tener la precisión que queremos? Para poder dar respuesta a esta pregunta, vamos a necesitar introducir en esta sección el concepto de convergencia.

Si suponemos que la sucesión de vectores es convergente a x , tendríamos que

$$x = Tx + c \quad (39)$$

y, por tanto, la convergencia de la matriz depende del radio espectral de la matriz T . De hecho, a partir de esta afirmación se enuncia un teorema, que generaliza este resultado:

Un esquema iterativo para resolver sistemas lineales del tipo

$$x^{(k+1)} = Tx^k + c \quad (40)$$

converge si y solo si el radio espectral de la matriz T es estrictamente menor que la unidad.

Este resultado es muy potente para la resolución de sistemas lineales, ya que nos va a permitir saber cuándo vamos a poder utilizar un método iterativo sin miedo a crear bucles infinitos o a tener soluciones inexactas.

La rapidez de convergencia de un procedimiento depende del radio espectral de la matriz relacionada con el método. Por ello, una forma de seleccionar un procedimiento que acelere la convergencia consiste en seleccionar un método cuya matriz asociada tenga un radio espectral mínimo. Antes de pasar a describir los procedimientos para que podáis elegirlos, vamos a explicar un nuevo modo de medir el grado de aproximación de la solución de un sistema lineal a la verdadera solución.

Si tenemos $\tilde{x} \in \mathbb{R}^n$ una aproximación a la solución del sistema lineal definido por $Ax = b$, el vector residual de \tilde{x} respecto de este sistema es $r = b - A\tilde{x}$. Desde el punto de vista intuitivo, parece razonable que si \tilde{x} es una aproximación a la solución de x de $Ax = b$ y el vector residual $r = b - A\tilde{x}$ tiene la propiedad de que $\|r\|$ es pequeño, entonces $\|x - \tilde{x}\|$ también será pequeño. A menudo este es el caso, pero algunos sistemas no tienen este comportamiento.

Ejemplo 7

En este ejemplo vamos a ver, con un sistema de ecuaciones lineal, que no todas las aproximaciones que consideremos son siempre las óptimas para poder encontrar la solución al sistema de ecuaciones lineales que queremos resolver.

El sistema lineal $Ax = b$ dado por

$$\begin{pmatrix} 1 & 2 \\ 10001 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 30001 \end{pmatrix} \quad (41)$$

tiene la solución única $x = (1,1)$. Vamos a suponer que la aproximación $\tilde{x} = (3,0)$ es la solución del sistema. Calculamos el valor residual y obtenemos:

$$r = b - A\tilde{x} = \begin{pmatrix} 3 \\ 30001 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 10001 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -0,0002 \end{pmatrix} \quad (42)$$

de modo que $\|r\|_\infty = 0,0002$. Aunque la norma del vector residual es pequeña, la aproximación $\tilde{x} = (3,0)$ es evidentemente muy deficiente; de hecho, $\|x - \tilde{x}\|_\infty = 2$.

A partir de este ejemplo, se puede ver la dificultad que tiene interpretar el resultado de un cálculo numérico a partir de la estimación del error. Bajo ciertas condiciones, se puede conseguir información “fiable” sobre el error si se consideran las normas de la matriz A y la de su inversa, A^{-1} . A continuación, vamos a dar un resultado importante para este cálculo. Supongamos que \tilde{x} es una aproximación de la solución de $Ax = b$, que A es una matriz no singular y que r es el vector residual de \tilde{x} . Entonces, para toda norma natural (recordemos que las normas naturales son aquellas normas inducidas por una norma vectorial, por ejemplo, $\|\cdot\|_2$, $\|\cdot\|_\infty$):

$$\begin{aligned} \|x - \tilde{x}\| &\leq \|r\| \cdot \|A^{-1}\| \\ &\text{y} \\ \frac{\|x - \tilde{x}\|}{\|x\|} &\leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|r\|}{\|b\|} \end{aligned} \quad (43)$$

Estas desigualdades nos indican que las cantidades $\|A^{-1}\|$ y $\|A\| \cdot \|A^{-1}\|$ ofrecen un indicio de la conexión entre el vector residual r y la exactitud de la aproximación. En general, el error relativo $\frac{\|x - \tilde{x}\|}{\|x\|}$ es de gran interés, y de acuerdo con el anterior teorema, está acotado por $\|A\| \cdot \|A^{-1}\|$ con el residual relativo de esta aproximación $\frac{\|r\|}{\|b\|}$. En esta aproximación, puede usarse cualquier norma adecuada, siempre que se use la misma durante todo el proceso. El producto $\|A^{-1}\| \cdot \|A\|$ recibe el nombre de condición de la matriz ($K(A)$), y en la práctica si $K(A)$ es menor que 1, la matriz A está bien condicionada. Mientras que si $K(A)$ es mayor que 1, la matriz está mal condicionada. Esto es muy importante en la práctica, ya que no se puede asegurar que un vector residual pequeño implique que una solución aproximada exacta.

3.2.2. Método iterativo de Jacobi

Se basa en una descomposición de la matriz A del tipo $A = M - N$, con $M = D$, parte diagonal de la matriz A cuyos elementos se suponen no nulos, y $N = L + U$, siendo L , U las partes triangulares inferior y superior respectivamente de la matriz A cambiadas de signo. La descomposición propuesta satisface los criterios de selección expuestos. Es posible definir estas matrices D , L y U de la misma forma que en la descomposición LU :

Importante

En una matriz bien condicionada ($K(A) < 1$), un vector residual (r) pequeño implica una seguridad relativa de que una solución aproximada será exacta.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad (44)$$

de modo que la resultante para aplicar este método será:

$$A = M - N = D - (L + U) = \begin{pmatrix} a_{11} & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix} - \begin{pmatrix} 0 & \cdots & 0 \\ -a_{21} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & 0 \end{pmatrix} - \begin{pmatrix} 0 & \cdots & -a_{1n} \\ 0 & \cdots & -a_{2n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad (45)$$

Veamos con un ejemplo cómo funciona el esquema general en la descomposición de Jacobi de la matriz A . Se trata de resolver el sistema lineal:

$$\begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 10 \\ 1 \end{pmatrix} \quad (46)$$

cuya solución es el vector $x = (-1, 5; 3; -0,5)$.

Por tanto, se tiene la siguiente descomposición que hemos descrito en la ecuación (45), es decir, tenemos que calcular las matrices D , L y U para poder resolver el sistema de ecuaciones mediante el método de Jacobi:

$$M = D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad L = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \quad U = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \quad (47)$$

con $N = L + U$.

La elección del estimador inicial no influye en la convergencia, pero sí en el número de iteraciones para llegar a una solución aceptable. Si se conoce una aproximación de la solución, se usará como estimador inicial. En caso de que no sea así, se puede tomar como estimador inicial bien el término independiente, o bien un vector de componentes todas iguales a 1. Nosotros tomaremos aquí como estimador inicial un vector que satisfaga la primera ecuación del sistema lineal $x^{(0)} = (-1, 1, -1)$.

Para comprobar si estamos convergiendo a la solución, es necesario disponer de un buen criterio de convergencia. Aquí, tomamos por comodidad la norma $\|\cdot\|_\infty$ y

$$\|r^{(0)}\| = \|Ax^{(0)} - b\|_\infty = 8 \quad (48)$$

Demos el primer paso del esquema:

$$Dx^{(1)} = Nx^{(0)} + b = (L + U)x^{(0)} + b \quad (49)$$

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} x^{(1)} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} -3 \\ 10 \\ 1 \end{pmatrix} \quad (50)$$

es decir,

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} x^{(1)} = \begin{pmatrix} -4 \\ 10 \\ 0 \end{pmatrix} \rightarrow x^{(1)} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} \quad (51)$$

con

$$\|r^{(1)}\|_\infty = 2 \quad (52)$$

Por tanto, el residuo ha disminuido. Si seguimos iterando:

$$x^{(2)} = \begin{pmatrix} -1,5 \\ 2,75 \\ -0,5 \end{pmatrix}; \quad \|r^{(2)}\|_\infty = 1$$

$$x^{(5)} = \begin{pmatrix} -1,4922 \\ 3,0000 \\ -0,4522 \end{pmatrix}; \quad \|r^{(5)}\|_\infty = 0,0313 \quad (53)$$

$$x^{(9)} = \begin{pmatrix} -1,4999 \\ 3,0000 \\ -0,4999 \end{pmatrix}; \quad \|r^{(9)}\|_\infty = 4,8828 \times 10^{-4}$$

Como consejo, cuando tengáis un sistema de ecuaciones lineales, y estéis pensando en qué método vais a usar, el método de Jacobi debería estar en las primeras posiciones de esa lista.

3.2.3. Método iterativo de Gauss-Seidel

Se basa en una descomposición de la matriz A del tipo $A = M - N$, con $M = D - L$, y $N = U$. En cada iteración se tiene que resolver un sistema triangular por sustitución hacia adelante. La mecánica de cada paso del método de Gauss-Seidel es, por tanto, más complicada que en el método de Jacobi, pero la velocidad de convergencia es superior. Ilustremos el algoritmo con el mismo ejemplo que para el método de Jacobi. Demos el primer paso del esquema:

$$Mx^{(1)} = Nx^{(0)} + b = Ux^{(0)} + b \quad (54)$$

$$\begin{pmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 1 & 4 \end{pmatrix} x^{(1)} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} -3 \\ 10 \\ 1 \end{pmatrix} \quad (55)$$

Y tendremos, por tanto, que resolver el sistema triangular superior:

$$\begin{pmatrix} 4 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 1 & 4 \end{pmatrix} x^{(1)} = \begin{pmatrix} -4 \\ 11 \\ 1 \end{pmatrix} \quad (56)$$

que se resuelve por sustitución hacia adelante:

$$x^{(1)} = \begin{pmatrix} -1,0000 \\ 3,0000 \\ -0,5000 \end{pmatrix} \quad (57)$$

con

$$\|r^{(1)}\|_{\infty} = 2 \quad (58)$$

En el paso siguiente se obtiene el resultado con cuatro decimales:

$$x^{(2)} = \begin{pmatrix} -1,0000 \\ 3,0000 \\ -0,5000 \end{pmatrix} \quad (59)$$

Se constata un comportamiento mucho mejor que el del método de Jacobi. Os proponemos hacer un estudio comparativo entre la convergencia de los dos métodos iterativos presentados para ver que el método de Gauss-Seidel converge más rápido que el método de Jacobi.

3.2.4. Aplicación al álgebra lineal: cálculo de valores y vectores propios. Método de la potencia y de la potencia inversa

La solución de muchos problemas de la física y otros campos requiere calcular, o al menos eliminar, los valores propios y los correspondientes vectores propios de una matriz. Ya hemos visto que una matriz A de $n \times n$ tiene exactamente n valores propios que son raíces del polinomio $p(\lambda) = \det(A - \lambda I)$. En teoría, los valores propios de A se obtienen calculando las n raíces del polinomio $p(\lambda)$; después, se resuelven los sistemas lineales asociados para determinar los vectores propios correspondientes.

En la práctica es difícil obtener el polinomio característico $p(\lambda)$ y, excepto para los valores pequeños de n , también es difícil determinar las raíces del polinomio de grado n . Para obtener los valores y vectores propios, se necesitan los métodos de aproximación.

Método de la potencia

Supongamos que la matriz A es diagonalizable y que tiene un valor propio dominante que denotaremos λ_1 , luego

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \quad (60)$$

Sea $\{x_1, x_2, \dots, x_n\}$ la base de vectores propios asociada, de modo que $Ax_i = \lambda_i x_i$ para $i = 1, \dots, n$. Este método determina λ_1 y un vector propio asociado x_1 . Veamos su fundamento e ilustremos su aplicación con un ejemplo.

Cualquier vector arbitrario $x \in \mathbb{R}^n$ se expresa de una única forma como combinación lineal de los vectores propios de la base.

$$x = C_1 x_1 + C_2 x_2 + \dots + C_n x_n \quad (61)$$

con $C_i \in \mathbb{R}$ para $i = 1, \dots, n$.

Multiplicando los dos miembros de la ecuación anterior por las potencias A^k de A con $k \in \mathbb{N}$

$$\begin{aligned} A^k x &= A^k (C_1 x_1 + C_2 x_2 + \dots + C_n x_n) = C_1 \lambda_1^k x_1 + C_2 \lambda_2^k x_2 + \dots + C_n \lambda_n^k x_n = \\ &= \lambda_1^k \left[C_1 x_1 + C_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + C_n \left(\frac{\lambda_n}{\lambda_1} \right)^k x_n \right] \end{aligned} \quad (62)$$

Puesto que $|\lambda_1| > |\lambda_i|$ para $i \geq 2$, los cocientes $\frac{\lambda_i}{|\lambda_1|}$ son valores más pequeños que 1. Esto quiere decir que a medida que la k se hace mayor, los cocientes $\left(\frac{\lambda_i}{|\lambda_1|}\right)^k$ son cada vez más pequeños, son prácticamente 0.

En consecuencia, podemos decir que cuando k crece $A^{(k \cdot x)}$ es aproximadamente igual a $\lambda_1^k C_1 x_1$:

$$A^k x \approx \lambda_1^k C_1 x_1 \quad (63)$$

Si elegimos una estimadora inicial adecuada $x^{(0)}$, se puede construir la siguiente sucesión:

$$x^{(k+1)} = Ax^{(k)} \longrightarrow x^{(k)} = A^k x^{(0)} = A^k x \quad (64)$$

Por tanto, si en la ecuación anterior sustituimos el valor de $A^k x^{(0)}$ de la ecuación 63, se tiene:

$$x^{(k)} \approx \lambda_1^k C_1 x_1 = \lambda_1^k C_1 \quad (65)$$

y si ahora iteramos un paso más (de k a $k+1$), se tiene:

$$x^{(k+1)} \approx \lambda_1^{(k+1)} C_1 x_1 = \lambda_1^{(k+1)} C_1 \quad (66)$$

Si ahora dividimos las ecuaciones 66 entre la ecuación 65, se obtiene la siguiente expresión, que nos permite calcular λ_1 :

$$\frac{x^{(k+1)}}{x^{(k)}} = \lambda_1 \quad (67)$$

luego si $x_1^{(k)} = 1$, entonces $x_1^{(k+1)} = \lambda_1$. Si a continuación escalamos $x^{(k+1)}$ para que $x_1^{(k+1)} = 1$, entonces $x_1^{(k+2)} = \lambda_1$, etc.

En el siguiente ejemplo vamos a aplicar el algoritmo al cálculo del valor propio λ_1 ; de este modo, el siguiente ejemplo nos va a demostrar que aplicando este método iterativo vamos a ser capaces de calcular el valor propio mayor en

valor absoluto de la matriz A . Fijaos en que en este ejemplo se da el valor de $\lambda_1 = 6$ y se pide comprobarlo. Daos cuenta de que, si en algún momento tuvierais que calcular el valor propio de mayor valor absoluto de una matriz, no tendríais que calcular el polinomio característico y resolverlo, bastaría con usar este método.

Ejemplo 8

Vamos a ver el método de la potencia con un ejemplo para calcular un valor propio de una matriz, de la que previamente vamos a conocer sus valores propios; así, vamos a comprobar que el resultado es correcto. Por tanto, dada la matriz

$$\begin{pmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{pmatrix} \tag{68}$$

Probadlo vosotros mismos

Os invitamos a resolver con detalle el ejercicio y llegar a los resultados que se aportan en él.

que tiene los valores característicos $\lambda_1 = 6$, $\lambda_2 = 3$ y $\lambda_3 = 2$, suponemos $x^{(0)} = (1, 1, 1)$, y entonces:

$$y^{(1)} = Ax^{(0)} = (10, 8, 1) \tag{69}$$

así que

$$\|y^{(1)}\|_\infty = 10; \quad y_1^{(1)} = 10; \quad x^{(1)} = \frac{y^{(1)}}{10} = (1; 0,8; 0,1) \tag{70}$$

Continuando de esta manera se generan los valores de la siguiente tabla, donde λ es la aproximación al valor propio dominante de la matriz, que en este caso es 6. Este método, además, nos dará su vector propio asociado, tal y como se puede observar en la tabla que recoge las doce primeras iteraciones del método de la potencia:

Tabla con los valores de las iteraciones del método de la potencia

Número de iteración	λ	x
0	-	(1,1,1)
1	10	(1; 0,8; 0,1)
2	7,2	(1; 0,75; -0,11)
3	6,5	(1; 0,730769; -0,188803)
4	6,250769	(1; 0,722200; -0,220850)
5	6,111000	(1; 0,718182; -0,235915)
6	6,054546	(1; 0,716216; -0,243095)
7	6,027027	(1; 0,715247; -0,246588)
8	6,013453	(1; 0,714765; -0,248306)
9	6,006711	(1; 0,714525; -0,249157)
10	6,003352	(1; 0,714405; -0,249579)
11	6,001675	(1; 0,714346; -0,249790)
12	6,000837	(1; 0,714316; -0,249895)

Como se puede ver, la aproximación al valor propio dominante (es decir, el valor propio de mayor valor absoluto) es $\lambda_1 = 6,000837 \approx 6$; como se esperaba, además, este método nos da el vector propio asociado a este valor propio.

Método de la potencia inversa

Se puede obtener el valor propio de menor valor absoluto de A y su vector propio asociado aplicando el método de la potencia a A^{-1} , cuyos valores propios son, como sabemos, los recíprocos de los valores de A . El recíproco del menor valor propio de A en valor absoluto es el de mayor absoluto de A^{-1} . En la práctica, se utiliza la descomposición $A = LU$ para resolver este problema en vez de calcular A^{-1} .

Una vez elegido el estimador inicial $x(0)$ se calcula $x^{(1)}$ resolviendo el sistema

$$Ax^{(1)} = (LU)x^{(1)} = x^{(0)} \tag{71}$$

Si A^{-1} no existe, 0 es el valor propio de menor valor absoluto y cualquier vector del núcleo de A se puede tomar como vector propio asociado. El resto de los valores propios y de los vectores propios asociados se pueden obtener aplicando reiteradamente la siguiente idea. Una vez conocido el elemento propio (λ_1, x_1) , se selecciona un estimador inicial que sea ortogonal a x_1 y aplicando el método de las potencias se obtiene λ_2 y un vector propio asociado x_2 . Para obtener λ_3 se elige un vector inicial que sea ortogonal tanto a x_1 como a x_2 y se sigue el proceso hasta tener todos los valores propios calculados y sus respectivos vectores propios asociados a los valores propios.

Hasta el momento se han presentado algunos de los métodos directos o iterativos principales para resolver los sistemas de ecuaciones lineales. Resumiendo, podemos decir que lo primero que debéis hacer es colocar el sistema en forma de matriz, decidir cuál es el error que queréis tener y, por último, decidir qué método es más eficiente para resolverlo.

Después de la introducción a los métodos numéricos para resolver los sistemas de ecuaciones lineales tenéis el nivel suficiente para afrontar el primer ejemplo de esta guía, con el que se motivaba el porqué de la importancia de usar los métodos numéricos para buscar la solución en sistemas de ecuaciones lineales. ¿Habéis vuelto a buscar la solución de ese sistema? ¿Qué método elegiríais?

Convergencia

Cuando definíamos el concepto de convergencia, es posible que no quedase demasiado claro, pero aquí tenemos un claro ejemplo. Si os fijáis en los valores de λ , a medida que el número de iteraciones aumenta, este va tendiendo a 6; se puede decir entonces que λ converge a 6 en este ejemplo.

Pista

Probad a resolverlo con el vector unidad como primera aproximación.

Bibliografía

Atkinson, L. V.; Harley, P. J. (1987). *Introducción a los métodos numéricos con Pascal*. Argentina: Ed. Addison-Wesley.

Douglas Faires, J.; Burden, R. (2004). *Análisis numérico*. México: International Thomson Editores.

Kincaid, D. (2009). *Numerical analysis: mathematics of scientific computing*. Providence, R.I.: American Mathematical Society.

