

Tema 3. Métodos numéricos para la resolución de sistemas lineales y no lineales.

Computación Numérica

Antonio Palacio

Departamento de Matemáticas
Universidad de Oviedo

palacioantonio@uniovi.es

Curso 2021-2022

Contenidos

Métodos directos

1 Métodos directos

- El método de eliminación de Gauss
- La factorización LU
- Eliminación de Gauss con pivoteo
- Factorización de Cholesky
- Factorización QR
- Número de condición de una matriz

2 Métodos iterativos

- Introducción
- El método de Jacobi
- El método de Gauss-Seidel
- Resultados de convergencia

3 Sistemas no lineales

- Introducción
- Métodos de punto fijo
- Método de Newton
- Variaciones sobre el método de Newton

Contenidos I

1 Métodos directos

- El método de eliminación de Gauss
- La factorización LU
- Eliminación de Gauss con pivoteo
- Factorización de Cholesky
- Factorización QR
- Número de condición de una matriz

2 Métodos iterativos

- Introducción
- El método de Jacobi
- El método de Gauss-Seidel
- Resultados de convergencia

3 Sistemas no lineales

- Introducción
- Métodos de punto fijo
- Método de Newton
- Variaciones sobre el método de Newton

Métodos directos

Métodos directos

Introducción

- Dada una matriz real A de orden $n \times n$ y un vector \vec{b} (matriz columna de orden $n \times 1$) se plantea el problema de hallar un vector \vec{x} tal que

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

- A invertible $\implies \vec{x} = A^{-1}\vec{b}$.
- Métodos fáciles de programar.
- Comportarse bien respecto a los errores de redondeo.
- Precisar un número razonable de operaciones.
- Métodos directos y métodos iterativos.

- Notación: $\vec{x} = (x_1, x_2, \dots, x_n) \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

El método de eliminación de Gauss

Método de eliminación de Gauss

- Se considera el sistema lineal de orden $n \times n$

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

- La eliminación Gaussiana consiste en obtener, mediante operaciones elementales, un sistema equivalente al anterior

$$U\vec{x} = \vec{c}$$

siendo U matriz triangular superior.

El método de eliminación de Gauss

Método de eliminación de Gauss

- El método de Gauss transforma $Ax = b$ en un sistema triangular equivalente tras $n - 1$ etapas.
- Secuencia de matrices generadas:

$$A = A_1 \rightarrow A_2 \rightarrow \dots \rightarrow A_n$$

$$b = b^{(1)} \rightarrow b^{(2)} \rightarrow \dots \rightarrow b^{(n)}$$

$$A_k = (a_{ij}^{(k)}), \quad b^{(k)} = (b_i^{(k)})$$

- A_n es una matriz triangular superior.

El método de eliminación de Gauss

Método de eliminación de Gauss

$$A_k = \begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \dots & a_{1k}^{(k)} & \dots & \dots & a_{1n}^{(k)} \\ 0 & a_{22}^{(k)} & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \dots & a_{kn}^{(k)} \\ 0 & 0 & \dots & a_{k+1,k}^{(k)} & a_{k+1,k+1}^{(k)} & \dots & a_{k+1,n}^{(k)} \\ 0 & 0 & \dots & a_{k+2,k}^{(k)} & a_{k+2,k+1}^{(k)} & \dots & a_{k+2,n}^{(k)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{nk}^{(k)} & a_{n,k+1}^{(k)} & \dots & a_{nn}^{(k)} \end{pmatrix}$$

El método de eliminación de Gauss

Método de eliminación de Gauss

- Elemento (i, j) de la matriz A_{k+1} :

$$a_{ij}^{(k+1)} = \begin{cases} a_{ij}^{(k)}, & 1 \leq i \leq k, \quad 1 \leq j \leq n, \\ a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} a_{kj}^{(k)}, & k+1 \leq i \leq n, \quad k+1 \leq j \leq n. \end{cases}$$

- Componente i -ésima del vector $b^{(k+1)}$:

$$b_i^{(k+1)} = \begin{cases} b_i^{(k)}, & 1 \leq i \leq k, \\ b_i^{(k)} - m_{ik} b_k^{(k)}, & k+1 \leq i \leq n. \end{cases}$$

El método de eliminación de Gauss

Método de eliminación de Gauss

- La ecuación k -ésima se denomina **ecuación pivotal**.
- El elemento (k, k) lo denotaremos como el **elemento pivote** del paso k .
- **Multiplicador (i, k)** : $m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$.
- Sea el sistema lineal de n ecuaciones con n incógnitas $Ax = b$, con $|A| \neq 0$. Entonces, el sistema $A_n x = b^{(n)}$, obtenido a través de eliminación gaussiana, es equivalente.

El método de eliminación de Gauss

Método de eliminación de Gauss

Algorithm 1 Algoritmo de Gauss

```

1: for  $k = 1 : n - 1$  do
2:   for  $i = 1 : k + 1$  do
3:      $A(i, k) = A(i, k) / A(k, k)$ 
4:   for  $j = k + 1 : n$  do
5:      $A(i, j) = A(i, j) - A(i, k)A(k, j)$ 
6:   end for
7:    $b(i) = b(i) - A(i, k)b(k)$ 
8: end for
9: end for

```

El método de eliminación de Gauss

Método de eliminación de Gauss

- Expresión de la matriz L :

$$L = \begin{pmatrix} 1 & & & & & \\ m_{21} & 1 & & & & \\ m_{31} & m_{32} & 1 & & & \\ & \dots & \dots & \dots & & \\ m_{n1} & m_{n2} & m_{n3} & \dots & m_{n,n-1} & 1 \end{pmatrix}.$$

- En consecuencia el método de Gauss permite **descomponer** la matriz A como el producto de una matriz **triangular inferior** L por una matriz **triangular superior** U .

El método de eliminación de Gauss

Problema 3.1

Aplique el método de eliminación de Gauss para resolver el sistema $A\vec{x} = \vec{b}$, siendo

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 6 & 15 & 19 \\ 8 & 42 & 60 \end{pmatrix} \text{ y } \vec{b} = \begin{pmatrix} 5 \\ 30 \\ 98 \end{pmatrix}$$

El método de eliminación de Gauss

Cálculos del problema 3.1

Los cálculos realizados pueden resumirse del siguiente modo:

$$\bullet \text{ Primera etapa: } m_{21} = \frac{a_{21}}{a_{11}} = 3; ec_2 \rightarrow ec_2 - m_{21} * ec_1$$

$$m_{31} = \frac{a_{31}}{a_{11}} = 4; ec_3 \rightarrow ec_3 - m_{31} * ec_1$$

\bullet Segunda etapa:

$$m_{32} = \frac{a_{32}}{a_{22}} = 5; ec_3 \rightarrow ec_3 - m_{32} * ec_2$$

Los multiplicadores son $m_{21} = 3$, $m_{31} = 4$ y $m_{32} = 5$ por lo que

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 5 & 1 \end{pmatrix}$$

El método de eliminación de Gauss

Teorema 3.1

Sea A una matriz real de orden $n \times n$. Son equivalentes:

- Todas las submatrices principales son invertibles.
- El método de eliminación de Gauss puede completarse sin encontrar ningún pivote nulo.

Observaciones

- Si la matriz es invertible pero alguno de sus menores principales es nulo, será necesario intercambiar filas a lo largo del proceso para poder obtener una matriz triangular superior. Dicho intercambio se conoce como **pivoteo** o **pivotaje**.
- Desde el punto de vista de la resolución numérica puede ser conveniente aplicar pivoteo cuando los pivotes sean “demasiado pequeños”.
- Si se continua hasta obtener una matriz diagonal el método se denomina **Método de Gauss-Jordan**.

La factorización LU

Teorema 3.2 (Factorización LU)

Si una matriz A admite el proceso de eliminación de Gauss entonces se puede factorizar por $A = LU$, siendo L matriz triangular inferior con unos en la diagonal y U matriz triangular superior:

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \dots & 1 \end{pmatrix} \quad U = \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{nn} \end{pmatrix}$$

tales que:

- $L = (l_{ij})$, siendo l_{ij} , para $i > j$, el multiplicador obtenido durante la eliminación Gaussiana para obtener cero en la posición (i, j) .
- U es la matriz triangular superior obtenida en el método de eliminación de Gauss.

La factorización LU

Resolución mediante la factorización LU

En las condiciones anteriores se dice que L y U son la **factorización LU** de A . La resolución del sistema usando la factorización $A = LU$ se realiza descomponiendo el sistema original en dos sistemas triangulares:

$$A\vec{x} = LU\vec{x} = L(\underbrace{U\vec{x}}_{\vec{y}}) = L\vec{y} = \vec{b}$$

obteniéndose los sistemas:

- $L\vec{y} = \vec{b}$ (fase de sustitución progresiva)
- $U\vec{x} = \vec{y}$ (fase de sustitución regresiva)

Nota 3.1

El sistema $U\vec{x} = \vec{y}$ resultante de la factorización LU es el mismo que el resultante de aplicar el método de Gauss y que se había denotado por $U\vec{x} = \vec{c}$.

La factorización LU

Problema 3.2

Resuelva el sistema $A\vec{x} = \vec{b}$, siendo

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 6 & 15 & 19 \\ 8 & 42 & 60 \end{pmatrix} \text{ y } \vec{b} = \begin{pmatrix} 5 \\ 30 \\ 98 \end{pmatrix}$$

usando la factorización LU de la matriz A.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 5 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 & 4 \\ 0 & 6 & 7 \\ 0 & 0 & 9 \end{pmatrix} = LU$$

La factorización LU

Observaciones

- La factorización LU resultará ventajosa en aquellos problemas que requieran la resolución de muchos sistemas con la misma matriz. Por ejemplo, en el cálculo de la inversa de una matriz deben resolverse los sistemas $A\vec{x} = \vec{e}_1, \dots, A\vec{x} = \vec{e}_n$.
- Como aplicación teórica, tenemos que la descomposición LU permite demostrar que $\det(A) = \det(U)$. Esto demuestra que la eliminación Gaussiana es un método eficiente para el cálculo del determinante de una matriz.

Problema 3.3

Halle la factorización LU de la matriz $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Utilizando dicha factorización, calcule el determinante de la matriz y la matriz inversa.

La factorización LU

Coste operacional

Contando solo multiplicaciones/divisiones, se tiene (el número de sumas/restas es parecido y suele tener menor coste):

- Factorización: $O(n^3/3)$
- Sustitución: $O(n^2/2)$
- Por ejemplo, para un sistema 100×100 , es decir, para $n = 100$, se tiene que:

$$\frac{100^3}{3} \approx 333000 \text{ y } \frac{100^2}{2} = 5000.$$

- La sustitución tiene un coste “pequeño” frente a la factorización.
- En el método de Gauss se realiza una factorización y una sustitución.
- En el método LU se realizan una factorización y dos sustituciones.

La factorización LU

Definición 3.1

Una matriz $A = (a_{ij}) \in \mathcal{M}_{n \times n}(\mathbb{R})$ se dice que es de **diagonal estrictamente dominante** si verifica que

$$|a_{ii}| > |a_{i1}| + |a_{i2}| + \underbrace{\dots}_{\text{falta } |a_{ii}|} + |a_{in}| \text{ con } i = 1, \dots, n$$

Teorema 3.3

Si una matriz es de diagonal estrictamente dominante entonces la factorización LU y la eliminación de Gauss son estables frente a errores de redondeo.

Eliminación de Gauss con pivoteo

Observación

El pivoteo o pivotaje se realiza para elegir filas que tengan pivote no nulo (en términos absolutos o relativos) con el objeto de reducir los errores de la aritmética de precisión finita (errores de redondeo).

Problema 3.4

Compruebe que la eliminación de Gauss para $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$ necesita del pivoteo.

Eliminación de Gauss con pivoteo

Observación

El alto error obtenido se debe a que el pivote utilizado es “pequeño” frente al resto de los coeficientes de su fila.

Pivoteo parcial mediante factor de escala

Este procedimiento permite evitar este problema.

Si $A = (a_{ij})$ es la matriz de la etapa k del proceso de eliminación de Gauss, el método consiste en lo siguiente:

- $s_i = \max\{|a_{ik}|, |a_{ik+1}|, \dots, |a_{in}|\}$, con $i = k, \dots, n$ (factor de escala de la fila i).
- Sea d el índice de fila tal que $\frac{|a_{dk}|}{s_d} = \max\left\{\frac{|a_{kk}|}{s_k}, \frac{|a_{(k+1)k}|}{s_{k+1}}, \dots, \frac{|a_{nk}|}{s_n}\right\}$
- Se intercambian las filas k -ésima y d -ésima, de forma que esta última quede como fila pivotal.

Eliminación de Gauss con pivoteo

Problema 3.5

Resuelva los sistemas $\begin{cases} 10^{-4}x + y = 1 \\ x + y = 2 \end{cases}$ y $\begin{cases} 10x + 10^5y = 10^5 \\ x + y = 2 \end{cases}$, mediante eliminación gaussiana sin pivoteo y con aritmética de 3 dígitos. Calcule el error relativo cometido. La solución exacta es: $x_e = \frac{10000}{9999}$, $y_e = \frac{9998}{9999}$.

Eliminación de Gauss con pivoteo

Ejemplo 3.1

Realiza el proceso de eliminación de Gauss con pivoteo parcial con factor de escala a la matriz

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

Eliminación de Gauss con pivoteo

Problema 3.6

Se considera el sistema del ejemplo anterior
$$\left. \begin{aligned} 10x + 10^5 y &= 10^5 \\ x + y &= 2 \end{aligned} \right\}.$$
 Resuélvalo por Gauss, utilizando aritmética de tres dígitos y mediante pivoteo parcial con factor de escala. Calcule el error relativo cometido. La solución exacta es: $x_e = \frac{10000}{9999}, y_e = \frac{9998}{9999}.$

Eliminación de Gauss con pivoteo

Teorema 3.4

Si A es una matriz invertible, entonces existe una matriz de permutación P tal que PA posee factorización LU y tal que un sistema lineal cuya matriz sea PA puede ser resuelto por eliminación gaussiana sin encontrar pivotes nulos.

Observaciones

- La realización del pivoteo aumenta el coste del método.
- Se llama pivoteo parcial cuando se elige el mayor de los $|a_{ik}|, i = k, k+1, \dots, n$, sin utilizar factor de escala. Tiene menor coste, pero es menos estable. En el Ejercicio 3.6 no habría sido eficaz.
- Se llama pivoteo total cuando se elige el mayor de los $|a_{ij}|, i = k, k+1, \dots, n, j = k, k+1, \dots, n$, realizando intercambio de columnas si es necesario. Tiene mayor coste.
- En MATLAB se tiene la instrucción **[L U P]=lu(A).**

Eliminación de Gauss con pivoteo

Nota 3.2

- El pivoteo puede ser expresado utilizando las llamadas matrices de permutación, que es cualquier matriz obtenida realizando una permutación sobre las columnas de la matriz identidad.
- Cuando a una matriz es multiplicada por la izquierda con una matriz de permutación se obtiene la matriz inicial pero con un intercambio de filas, que se corresponde con la permutación realizada en la matriz identidad.

Por ejemplo,

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}.$$

Factorización de Cholesky

Definición 3.2

Una matriz real y simétrica A se dice definida positiva si $\vec{x}^T A \vec{x} > 0$ para todo vector no nulo $\vec{x} \in \mathbb{R}^n$.

Teorema 3.5

Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. Las siguientes afirmaciones son equivalentes:

- 1 A es simétrica y definida positiva.
- 2 A es simétrica y todos sus menores principales son positivos.
- 3 Existe una matriz triangular inferior $L = (l_{ij})$, con $l_{ii} > 0$ y verificando que $A = LL^T$.

Cuando existe, la factorización anterior es única y se denomina **factorización de Cholesky** de A .

Nota 3.3

La matriz L anterior no coincide, en general, con la matriz triangular inferior definida por la factorización LU.

Factorización de Cholesky

Problema 3.7

Sean $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & \frac{11}{2} \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$. Resuelva el sistema $A\vec{x} = \vec{b}$ utilizando la factorización de Cholesky.

Observaciones

- La resolución de un sistema mediante la factorización de Cholesky es estable frente a los errores de redondeo.
- Requiere aproximadamente la mitad de operaciones que la factorización LU y solo es necesario almacenar la matriz L .
- Puede ser usada para determinar si una matriz simétrica es definida positiva.

Factorización de Cholesky

Algoritmo de Cholesky

$$l_{11} = \sqrt{a_{11}} \quad l_{k1} = \frac{a_{k1}}{l_{11}} \quad k = 2 \dots n$$

Para $j = 2 \dots n - 1$

$$l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2} \quad l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right) \quad i = j+1 \dots n$$

$$l_{nn} = \sqrt{a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2}$$

Factorización QR

Definición 3.3

Se suele denominar **factorización QR** de una matriz A a una descomposición del tipo $A = QR$, siendo Q tal que $Q^t Q = I$ y R matriz triangular superior.

Procedimientos

Existen diferentes procedimientos para obtener tal factorización:

- Procedimientos basados en la ortogonalización de Gram-Schmidt.
 Q es una matriz de orden $m \times n$ cuyas columnas forman una base ortogonal del espacio engendrado por las columnas de A y R es una matriz de orden $n \times n$ invertible y triangular superior.
- Factorización QR de Householder.
 Q es de orden $m \times m$ (ortogonal, por ser cuadrada) y R de orden $m \times n$ (triangular superior). La obtención de esta descomposición es similar a la de la factorización LU .

Factorización QR

Ejemplo 3.2

Resuelve utilizando la factorización QR el sistema de ecuaciones $A\vec{x} = \vec{b}$, siendo:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$

Número de condición de una matriz

Objetivo

Se pretende determinar cuanto cambia la solución del sistema $A\vec{x} = \vec{b}$ al realizar una pequeña modificación en los coeficientes de A o \vec{b} .

Esta modificación ocurre por ejemplo cuando los datos provienen de medidas experimentales (contienen errores) o cuando los números son aproximados por números en punto flotante.

Para dicho estudio se necesitará utilizar algunos conceptos básicos del álgebra.

Número de condición de una matriz

Definición 3.6

En $\mathcal{M}_n(\mathbb{R})$ se definen las **normas matriciales 1, 2 e infinito** como:

- $\|A\|_1 = \sup \{|a_{1j}| + \dots + |a_{nj}|, j = 1, \dots, n\}$
- $\|A\|_2 = \sqrt{\rho(A^T A)}$, siendo $\rho(B)$ el **radio espectral de B**, es decir, el máximo de los módulos de los autovalores de B . Si la matriz A es simétrica, se cumple que $\|A\|_2 = \rho(A)$.
- $\|A\|_\infty = \sup \{|a_{i1}| + \dots + |a_{in}|, i = 1, \dots, n\}$

Nota 3.5

En lo que sigue, $\|A\|$ y $\|\vec{x}\|$ denotarán cualquiera de las normas anteriores.

Problema 3.8

Dada la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, calcule $\|A\|_\infty$, $\|A\|_1$ y $\|A\|_2$.

Número de condición de una matriz

Definición 3.4

Se denominan **autovalores, o valores propios**, de una matriz cuadrada A a las raíces del **polinomio característico** $P(\lambda) = \det(A - \lambda I_d)$.

Así, λ es valor propio de $A \Leftrightarrow P(\lambda) = 0$.

Definición 3.5

En \mathbb{R}^n se definen las **normas 1, 2 e infinito** como:

- $\|\vec{x}\|_1 = |x_1| + \dots + |x_n|$.
- $\|\vec{x}\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$.
- $\|\vec{x}\|_\infty = \max \{|x_1|, \dots, |x_n|\}$.

Nota 3.4

Si una sucesión converge a un punto en una norma, entonces converge a dicho punto con cualquier otra norma.

Número de condición de una matriz

Definición 3.7

Se denomina **error absoluto** al aproximar \vec{x} por \vec{x}^* a $\|\vec{x} - \vec{x}^*\|$ y para $\vec{x} \neq 0$, se denomina **error relativo** al cociente $\frac{\|\vec{x} - \vec{x}^*\|}{\|\vec{x}\|}$.

Problema 3.9

Sea $A\vec{x} = \vec{b}$ el sistema $\begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$.

- Calcule su solución exacta \vec{x}_e .
- Calcule la solución exacta \vec{x}_e^d del sistema perturbado obtenido al sumar $\Delta\vec{b} = (0, 1, -0, 1, 0, 1, -0, 1)$ a \vec{b} ($A\vec{x}_e^d = \vec{b} + \Delta\vec{b}$).
- Calcule los errores relativos de la solución y de la perturbación. Compárelos calculando el cociente de ambos.

Número de condición de una matriz

Definición 3.8

Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz invertible y sea $\|\cdot\|$ una norma matricial. Se denomina **número de condición** de A asociado a la norma $\|\cdot\|$ al número real:

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

Nota 3.6

Se denotará $\text{cond}_1(A) = \|A\|_1 \|A^{-1}\|_1$ y análogamente con las otras normas.

Observaciones:

- El número de condición es una cota de la amplificación que se puede producir por errores en los datos (ver teorema siguiente) y siempre verifica que $\text{cond}(A) \geq 1$. Pero si $\text{cond}(A) \gg 1$ significa que es posible que la solución del problema perturbado sea muy mala.
- Existen técnicas que permiten mejorar el número de condición de una matriz.

Número de condición de una matriz

Proposición

Sea A una matriz simétrica e invertible. Se verifica que $\text{cond}_2(A) = \frac{\max\{|\lambda_1|, \dots, |\lambda_n|\}}{\min\{|\lambda_1|, \dots, |\lambda_n|\}}$, siendo $\lambda_1, \dots, \lambda_n$ los autovalores de A .

Problema 3.10

$$\text{Sea } A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}; \text{ su inversa es } \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}.$$

Calcule el número de condición de A respecto de las tres normas matriciales. Compare este valor con el cociente calculado en el Ejercicio 3.9

Número de condición de una matriz

Teorema 3.6

Sean $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ una matriz invertible, $\vec{b} \in \mathbb{R}^n$ un vector no nulo, \vec{x} la solución del sistema $A\vec{x} = \vec{b}$ y sea $\|\cdot\|$ una norma matricial inducida.

- Para $\Delta\vec{b} \in \mathbb{R}^n$, sea $\Delta\vec{x} \in \mathbb{R}^n$ tal que $A(\vec{x} + \Delta\vec{x}) = \vec{b} + \Delta\vec{b}$. Entonces,

$$\frac{\|\Delta\vec{x}\|}{\|\vec{x}\|} \leq \text{cond}(A) \frac{\|\Delta\vec{b}\|}{\|\vec{b}\|}$$

- Para $\Delta A \in \mathcal{M}_{n \times n}(\mathbb{R})$ sea $\Delta\vec{x} \in \mathbb{R}^n$ tal que $(A + \Delta A)(\vec{x} + \Delta\vec{x}) = \vec{b}$. Entonces,

$$\frac{\|\Delta\vec{x}\|}{\|\vec{x} + \Delta\vec{x}\|} \leq \text{cond}(A) \frac{\|\Delta A\|}{\|A\|}$$

Contenidos

1 Métodos directos

- El método de eliminación de Gauss
- La factorización LU
- Eliminación de Gauss con pivoteo
- Factorización de Cholesky
- Factorización QR
- Número de condición de una matriz

2 Métodos iterativos

- Introducción
- El método de Jacobi
- El método de Gauss-Seidel
- Resultados de convergencia

3 Sistemas no lineales

- Introducción
- Métodos de punto fijo
- Método de Newton
- Variaciones sobre el método de Newton

Introducción

Objetivo

- Encontrar una solución sabiendo que:

$$A\vec{x} = \vec{b} \Leftrightarrow \vec{x} = \vec{G}(\vec{x}) = B\vec{x} + \vec{c}$$

- Se plantea el método de punto fijo siguiente:

$$\left. \begin{array}{l} \vec{x}^{(0)} \in \mathbb{R}^n \\ \vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{c} \end{array} \right\}$$

- La matriz B y el vector \vec{c} , caracterizan los diferentes métodos.
- Resultan útiles para sistemas “grandes”. Son lentos, pero les afectan poco los errores de redondeo. Aprovechan las matrices con gran número de ceros pues permiten almacenar y utilizar solo los elementos no nulos.

Introducción

Notación

Se tiene una matriz $A = (a_{ij})$, y un vector de términos independientes $\vec{b} = (b_i)$, además de las matrices

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} \quad L = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ a_{21} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 0 \end{pmatrix}$$

$$U = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ 0 & 0 & 0 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Esto permite descomponer $A = D + L + U$. Conviene observar que otros autores utilizan la notación de forma que $A = D - L - U$.

El método de Jacobi

Procedimiento

Consiste en despejar la incógnita x_i de la i -ésima ecuación. Suponiendo que a_{ii} es no nulo para todo i , el método puede describirse por:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(\sum_{j=1}^{i-1} -a_{ij}x_j^{(k)} + \sum_{j=i+1}^n -a_{ij}x_j^{(k)} + b_i \right)$$

y en forma matricial:

$$\vec{x}^{(k+1)} = \underbrace{-D^{-1}(L+U)}_{B_J} \vec{x}^{(k)} + \underbrace{D^{-1}\vec{b}}_{\vec{c}_J} = B_J \vec{x}^{(k)} + \vec{c}_J.$$

Esta última expresión puede ser utilizada como definición del **Método de Jacobi**.

El método de Jacobi

Proposición

Sea A una matriz invertible que tiene elementos no nulos en la diagonal. Entonces, la matriz B_J y el vector \vec{c}_J están bien definidos (existe D^{-1}) y se verifica que $A\vec{x} = \vec{b}$ si y solo si $\vec{x} = B_J \vec{x} + \vec{c}_J$

Observación

Si la matriz A es hueca (es decir, con un gran número de coeficientes nulos), también lo es la matriz B_J .

Problema 3.11

$$\text{Se considera el sistema } \left. \begin{array}{rcl} 4x_1 + 3x_2 & = & 24 \\ 3x_1 + 4x_2 - x_3 & = & 30 \\ -x_2 + 4x_3 & = & -24 \end{array} \right\}.$$

- 1 Halle las ecuaciones del método de Jacobi para la resolución de dicho sistema.
- 2 Calcule las matrices B_J y \vec{c}_J de dicho método.
- 3 Realice una iteración con $\vec{x}^{(0)} = (0, 4, 0)$.

El método de Gauss-Seidel

Procedimiento

Este método puede plantearse como una modificación al método de Jacobi.

- Se despeja la incógnita x_i de la i -ésima ecuación.
- Para calcular $x_i^{(k+1)}$ se utilizan:
 - $x_{i+1}^{(k)}, \dots, x_n^{(k)}$ de la etapa anterior.
 - $x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ de la etapa actual.

Suponiendo que a_{ii} es no nulo para todo i , el método puede describirse por:

- $x_i^{(k+1)} = \frac{1}{a_{ii}} \left(\sum_{j=1}^{i-1} -a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^n -a_{ij}x_j^{(k)} + b_i \right).$
- $a_{ii}x_i^{(k+1)} + \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} = -\sum_{j=i+1}^n a_{ij}x_j^{(k)} + b_i.$

Y utilizando la notación matricial:

$$\vec{x}^{(k+1)} = \underbrace{-(D+L)^{-1}U}_{B_{G-S}} \vec{x}^{(k)} + \underbrace{(D+L)^{-1}\vec{b}}_{\vec{c}_{G-S}} = B_{G-S}\vec{x}^{(k)} + \vec{c}_{G-S}.$$

Esta última expresión puede ser utilizada como definición del **Método de Gauss-Seidel**.

El método de Gauss-Seidel

Proposición

Sea A una matriz invertible que tiene elementos no nulos en la diagonal. Entonces, la matriz B_{G-S} y el vector \vec{c}_{G-S} están bien definidos (existe $(D+L)^{-1}$) y se verifica que $A\vec{x} = \vec{b}$ si y solo si $\vec{x} = B_{G-S}\vec{x} + \vec{c}_{G-S}$.

Observación

La iteración del método de Gauss-Seidel puede hacerse resolviendo un sistema cuya matriz es triangular: $(D+L)\vec{x}^{(k+1)} = -U\vec{x}^{(k)} + \vec{b}.$

Problema 3.12

$$\text{Se considera el sistema } \left. \begin{array}{rcl} 4x_1 + 3x_2 & = & 24 \\ 3x_1 + 4x_2 - x_3 & = & 30 \\ -x_2 + 4x_3 & = & -24 \end{array} \right\}.$$

- Halle las ecuaciones del método de Jacobi para la resolución de dicho sistema y deduzca de ellas las del método de Gauss-Seidel.
- Calcule las matrices B_{G-S} y \vec{c}_{G-S} de dicho método.
- Realice una iteración con $\vec{x}^{(0)} = (0, 4, 0).$

Resultados de convergencia

Introducción

- Se pretende estudiar la convergencia de la sucesión generada por

$$\vec{x}^{(0)} \in \mathbb{R}^n \text{ dado}$$

$$\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{c}$$

- Puesto que se trata de un método iterativo de punto fijo, siendo $\vec{g}(\vec{x}) = B\vec{x} + \vec{c}$ la función de iteración, la condición esencial para la convergencia es que $\|B\| < 1$. Es decir, la $\|B\|$ juega el papel del $|g'(r)|$ del teorema de convergencia local del tema anterior.

Resultados de convergencia

Teorema 3.7

Sea $B \in \mathcal{M}_{n \times n}(\mathbb{R})$, $\vec{c} \in \mathbb{R}^n$ y sea $\vec{g}(\vec{x}) = B\vec{x} + \vec{c}$. Son equivalentes:

- Si para alguna norma matricial se verifica que $\|B\| < 1$, la función de iteración $\vec{g}(\vec{x}) = B\vec{x} + \vec{c}$ tiene un único punto fijo y genera una sucesión convergente para todo $\vec{x}^{(0)} \in \mathbb{R}^n$.
- Si $0 \leq \rho(B) < 1$ entonces la función de iteración $\vec{g}(\vec{x})$, tiene un único punto fijo y genera una sucesión convergente para todo $\vec{x}^{(0)} \in \mathbb{R}^n$.
- Si $\vec{c} \neq \vec{0}$ y la función de iteración $\vec{g}(\vec{x})$, tiene un único punto fijo y genera una sucesión convergente para todo $\vec{x}^{(0)} \in \mathbb{R}^n$. Entonces $0 \leq \rho(B) < 1$.

Nota 3.7

Observe que las hipótesis se plantean sobre la matriz B del método iterativo y no sobre la matriz A del sistema.

Resultados de convergencia

Nota 3.8

Si $\rho(B) \geq 1$, puede existir algún $\vec{x}^{(0)} \in \mathbb{R}^n$ tal que el método converja como se puede ver en el siguiente ejercicio.

Problema 3.13

Dado el método iterativo de punto fijo $\vec{x}^{(k+1)} = B\vec{x}^{(k)} + \vec{c}$ con $B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \lambda \end{pmatrix}$ y $\vec{c} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

- 1 Compruebe que si $|\lambda| < 1$ entonces el método converge para cualquier $\vec{x}^{(0)}$ que se considere.
- 2 Para $\lambda = 2$, halle un vector inicial para que el método sea convergente y otro vector inicial para que no converja.

Resultados de convergencia

Teorema 3.8 (Convergencia de Jacobi y de Gauss-Seidel)

Si una matriz cuadrada A es de diagonal estrictamente dominante entonces los métodos de Jacobi y de Gauss-Seidel, aplicados a la resolución del sistema $A\vec{x} = \vec{b}$, son convergentes.

Nota 3.9

Observe que las hipótesis se plantean sobre la matriz A del sistema y no sobre la matriz B del método iterativo.

Problema 3.15

Se considera el sistema lineal $A\vec{x} = \vec{b}$, siendo $A = \begin{pmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ -2 & 1 & 5 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 7 \\ -21 \\ 5 \end{pmatrix}$.

Razone que los métodos de Jacobi y de Gauss-Seidel, aplicados a dicho sistema, son convergentes.

Resultados de convergencia

En la práctica, es razonable considerar la siguiente estimación:

$$\|\vec{x}^{(k)} - \vec{r}\| \approx \rho(B)^k \|\vec{x}^{(0)} - \vec{r}\|$$

Problema 3.14

Dado el sistema

$$\left. \begin{aligned} 4x_1 + 3x_2 &= 24 \\ 3x_1 + 4x_2 - x_3 &= 30 \\ -x_2 + 4x_3 &= -24 \end{aligned} \right\}$$

- 1 Estudie la convergencia del método de Jacobi y la del método de Gauss-Seidel para la resolución de dicho sistema.

Contenidos

- 1 **Métodos directos**
 - El método de eliminación de Gauss
 - La factorización LU
 - Eliminación de Gauss con pivoteo
 - Factorización de Cholesky
 - Factorización QR
 - Número de condición de una matriz
- 2 **Métodos iterativos**
 - Introducción
 - El método de Jacobi
 - El método de Gauss-Seidel
 - Resultados de convergencia
- 3 **Sistemas no lineales**
 - Introducción
 - Métodos de punto fijo
 - Método de Newton
 - Variaciones sobre el método de Newton

Introducción

Introducción

- Se supone planteado un sistema de n ecuaciones y n incógnitas:

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \dots, x_n) &= 0 \end{aligned}$$

- Denotando $\vec{f} = (f_1, \dots, f_n)$ y $\vec{x} = (x_1, \dots, x_n)$, el sistema puede representarse por $\vec{f}(\vec{x}) = \vec{0}$.
- El problema que se plantea entonces es hallar $\vec{r} \in \mathbb{R}^n$ tal $\vec{f}(\vec{r}) = \vec{0}$.

Problema 3.16

En el sistema no lineal $\left. \begin{aligned} x^2 + y^2 &= 5x \\ 2x^4 + y^4 &= 9y \end{aligned} \right\}$, encuentre la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tal que se pueda plantear como las soluciones de $f(x, y) = (0, 0)$

Métodos de punto fijo

Teorema 3.9 (Teorema de convergencia local)

Sean U abierto de \mathbb{R}^n , $\vec{g}: U \rightarrow \mathbb{R}^n$ de clase uno en U y $\vec{r} \in U$ un punto fijo de \vec{g} . Si existe $k \in [0, 1)$ tal que para $i, j = 1, \dots, n$ se verifica $\left| \frac{\partial g_i}{\partial x_j}(\vec{r}) \right| \cdot n \leq k$ entonces se cumple que:

- Existe V entorno de \vec{r} tal que si $\vec{x}^0 \in V$ la sucesión $\vec{x}^{m+1} = \vec{g}(\vec{x}^m)$, $m = 0, 1, \dots$ converge a \vec{r} .
- $\|\vec{r} - \vec{x}^m\| \leq k^m \frac{1}{1-k} \|\vec{x}^1 - \vec{x}^0\|$.
- $\|\vec{r} - \vec{x}^m\| \leq \frac{k}{1-k} \|\vec{x}^m - \vec{x}^{m-1}\|$.

Nota 3.10

La comprobación de las hipótesis del teorema 3.9 suele resultar muy complicada.

Métodos de punto fijo

Procedimiento

- Se pretende hallar $\vec{r} \in \mathbb{R}^n$ tal $\vec{f}(\vec{r}) = \vec{0}$.
- Se busca \vec{g} continua tal que si $\vec{g}(\vec{r}) = \vec{r}$ entonces $\vec{f}(\vec{r}) = \vec{0}$.
- Se construye la sucesión $\vec{x}^{m+1} = \vec{g}(\vec{x}^m)$, que si es convergente, su límite debe ser un punto fijo.

Problema 3.17

Se plantea el sistema no lineal $\begin{aligned} x^2 + y^2 - 5x &= 0 \\ 2x^4 + y^4 - 9y &= 0 \end{aligned}$.

- Halle una función g de forma que la solución del sistema sea equivalente a encontrar un punto fijo de g .
- Aplique la técnica de Gauss-Seidel al método del punto fijo definido por la función del apartado anterior (se calcula secuencialmente cada variable utilizando las variables nuevas).
- Realice una iteración de ambos métodos partiendo de $\vec{x}^0 = (1, 1)$.

Métodos de punto fijo

Problema 3.18

Sea $\vec{g}(x, y) = \left(\frac{x^2 + y^2}{5}, \frac{2x^4 + y^4}{9} \right)$.

- Compruebe que $(0, 0)$ y $(1, 2)$ son puntos fijos de \vec{g} .
- Compruebe que \vec{g} verifica las hipótesis del teorema de convergencia en $(0, 0)$.
- Realice una iteración del método de punto fijo tomando $x^0 = y^0 = \frac{1}{2}$.
- Calcule el error absoluto cometido al aproximar $(0, 0)$ por (x^1, y^1) en la norma dos, en la norma uno y en la norma infinito.

Método de Newton

Procedimiento

- Se busca $\vec{r} \in U$ tal $\vec{f}(\vec{r}) = \vec{0}$ con $U \subset \mathbb{R}^n$ y $\vec{f}: U \rightarrow \mathbb{R}^n$.
- Se toma \vec{x}^0 y se admite que $\vec{f}(\vec{x}) \approx \vec{f}(\vec{x}^0) + d\vec{f}(\vec{x}^0)(\vec{x} - \vec{x}^0)$.
- Se reemplaza entonces el problema original $\vec{f}(\vec{x}) = \vec{0}$ por el sistema lineal: $\vec{f}(\vec{x}^0) + d\vec{f}(\vec{x}^0)(\vec{x} - \vec{x}^0) = \vec{0}$
- Se toma como nueva aproximación de la raíz la solución de ese sistema lineal: $\vec{x}^1 = \vec{x}^0 + [d\vec{f}(\vec{x}^0)]^{-1}(-\vec{f}(\vec{x}^0))$.

Definición 3.9

Matriz jacobiana La **matriz jacobiana o matriz asociada a la diferencial** de una función está definida por:

$$d\vec{f}(\vec{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \frac{\partial f_1}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}_0) & \frac{\partial f_2}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x}_0) & \frac{\partial f_n}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f_n}{\partial x_n}(\vec{x}_0) \end{pmatrix}$$

Método de Newton

Método de Newton

Desde el punto de vista de la programación se plantean las tres etapas siguientes:

- \vec{x}^0 dado
- Se calcula $\Delta\vec{x}^0$ como la solución del sistema lineal

$$[d\vec{f}(\vec{x}^0)](\Delta\vec{x}^0) = -\vec{f}(\vec{x}^0)$$

(es decir, $\Delta\vec{x}^0 = [d\vec{f}(\vec{x}^0)]^{-1}(-\vec{f}(\vec{x}^0))$)

- Se define $\vec{x}^1 = \vec{x}^0 + \Delta\vec{x}^0$.
- 1 El vector $\Delta\vec{x}^0$ se interpreta como la corrección que se debe realizar sobre \vec{x}^0 para aproximar mejor la solución.
- 2 Se realiza test de convergencia sobre la norma de $\Delta\vec{x}^0$.

Método de Newton

Problema 3.19

Realice una iteración del método de Newton para el sistema $\left. \begin{array}{l} x^2 + y^2 = 4 \\ xy = 1 \end{array} \right\}$, tomando como estimación inicial el punto $\left(\frac{1}{2}, 2\right)$.

Teorema 3.10 (Teorema de convergencia local)

Sean $U \subset \mathbb{R}^n$, $\vec{f}: U \rightarrow \mathbb{R}^n$ una función de clase dos y sea $\vec{r} \in U$ tal que $\vec{f}(\vec{r}) = \vec{0}$. Si $d\vec{f}(\vec{r})$ es invertible, existe un entorno V de \vec{r} tal que si $\vec{x}^0 \in V$ la sucesión definida por

$$\vec{x}^{m+1} = \vec{x}^m + [d\vec{f}(\vec{x}^m)]^{-1}(-\vec{f}(\vec{x}^m))$$

para $m = 0, 1, 2, \dots$, converge a \vec{r} . Además, la convergencia es de orden dos, al menos.

Método de Newton

Observaciones

- El método tiene convergencia local rápida.
- Cada iteración del método resulta muy costosa pues se deben realizar las siguientes tareas:
 - 1 Evaluar las n^2 funciones $\frac{\partial f_i}{\partial x_j}$ en \vec{x}^m .
 - 2 Resolver un sistema lineal de orden n .

Problema 3.20

Sabiendo que existe al menos una solución del sistema, estudie si se verifican las hipótesis del teorema de convergencia local del método de Newton para el sistema

$$\left. \begin{array}{l} x^2 + y^2 = 4 \\ xy = 1 \end{array} \right\}$$

Variaciones sobre el método de Newton

Variaciones sobre el método de Newton

El objetivo de estos métodos es evitar o reducir el cálculo de $d\vec{f}(\vec{x}^k)$, a cambio de perder velocidad de convergencia. Las principales opciones en este sentido son las siguientes:

- **Cálculo aproximado de las derivadas parciales:** Se realiza la aproximación

$$\frac{\partial f_i}{\partial x_j}(\vec{x}^k) \approx \frac{f_i(\vec{x}^k + h\vec{e}_j) - f_i(\vec{x}^k)}{h}$$

siendo \vec{e}_j el j -ésimo vector de la base canónica y h un parámetro a elegir.

- **Extensión del método de la secante.** Se realiza la aproximación

$$\frac{\partial f_i}{\partial x_j}(\vec{x}^k) \approx \frac{f_i(\vec{x}^k) - f_i(x_1^k, \dots, x_j^{k-1}, \dots, x_n^k)}{x_j^k - x_j^{k-1}}$$

- **Métodos Cuasi-Newton.** La diferencial $d\vec{f}$ solo se actualiza cada cierto número de iteraciones.

Variaciones sobre el método de Newton

Problema 3.21

Realice una iteración del método obtenido como extensión del método de la secante para el sistema

$$\left. \begin{aligned} x^2 + y^2 &= 4 \\ xy &= 1 \end{aligned} \right\}$$

tomando como estimaciones iniciales los puntos $\left(\frac{1}{2}, 2\right)$ y $\left(\frac{1}{3}, 3\right)$.