Sistemas de ecuaciones lineales

Escuela de Ingeniería Informática de Oviedo

Contenidos

Problema

Objetivo

Resolver sistemas de ecuaciones lineales con el mismo número de ecuaciones que de incógnitas

Problema

Dados los números a_{ij} y b_j para $i, j = 1, 2, \ldots, n$ se trata de hallar los números $x_1, x_2, \ldots, x_n$ que verifican las n ecuaciones lineales

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n & = & b_n \end{array}$$

Problema

- Definimos:
    - la matriz de coeficientes: $A = (a_{ij})_{i,j=1}^{n}$
    - el vector del segundo miembro: $\mathbf{b} = (b_i)_{i=1}^{n}$
    - el vector de incógnitas: $\mathbf{x} = (x_i)_{i=1}^{n}$
- Usando la notación matricial, el sistema se escribe

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

- En lo que sigue, suponemos que el sistema tiene una única solución

Soluciones numéricas

Los métodos pueden ser:

1. Métodos directos
   - La solución se calcula en un número finito de pasos conocidos *a priori.*
   - Sólo están sujetos a errores de redondeo.
   - Son adecuados para resolver
     - sistemas pequeños
     - sistemas grandes con matriz llena.

   Veremos los métodos de: **Gauss, Gauss con pivote y Descomposición LU**

2. Métodos iterativos
   - Construyen una sucesión que converge a la solución del sistema.
   - Además de los errores de redondeo., exite un error de truncamiento.

   Veremos los métodos de: **Jacobi y Gauss-Seidel**

Método de Gauss: triangularización

- La idea del método de Gauss es transforma el sistema de ecuaciones lineales original en un sistema de matriz triangular superior con las mismas soluciones.
- Para ello se realizan las operaciones:
  - $f_i \to f_i + \lambda f_j, \quad j \neq i$
  - $f_i \leftrightarrow f_j$ (si usamos la estrategia del pivote)
- En estos cálculos sólo intervienen la matriz de coeficientes y el vector del segundo miembro

Método de Gauss: sustitución regresiva

Una vez hemos triangularizado la matriz de coeficientes tenemos un sistema del tipo

$$U\mathbf{x} = \mathbf{b}'$$

donde U es una matriz triangular superior:

$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & u_{nn} \end{pmatrix}$$

Método de Gauss: sustitución regresiva

- La ecuación i-ésima del sistema, para $i = 1, 2, \ldots, n$ es:

$$u_{ii}x_i + u_{ii+1}x_{i+1} + \cdots + u_{in}x_n = b_i$$

- Por lo tanto, para $i = n, n-1, \ldots, 1$, si $u_{ii} \neq 0$ entonces

$$x_i = \frac{b_i - u_{ii+1}x_{i+1} - \cdots - u_{in}x_n}{u_{ii}} = \frac{1}{u_{ii}}\left(b_i - \sum_{j=i+1}^{n} u_{ij}x_j\right)$$

Método de Gauss: sustitución regresiva

- **Algoritmo sustitución hacia atrás**:

  1. Si $u_{nn} = 0$, parar (no existe solución única).
  2. Hacer $x_n = \dfrac{b_n}{u_{nn}}$
  3. Para $i = n-1, \ldots, 1$
     1. Si $u_{ii} = 0$, parar (no tiene solución)
     2. Hacer

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

Un ejemplo de Gauss

Resolver un sistema lineal por Gauss

$$\begin{array}{rrrcr} 2x & +3y & -z & = & 5 \\ 4x & +4y & -3z & = & 3 \\ -2x & +3y & -z & = & 1 \end{array}$$

TRIANGULARIZACIÓN

Hacemos ceros por debajo del pivote 2 en la primera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \begin{array}{lcl} f_1' & = & f_1 \\ f_2' & = & f_2 - \frac{4}{2} f_1 \\ f_3' & = & f_3 - \frac{-2}{2} f_1 \end{array}$$

Un ejemplo de Gauss

Resolver un sistema lineal por Gauss

$$\begin{array}{rrrcr} 2x & +3y & -z & = & 5 \\ 4x & +4y & -3z & = & 3 \\ -2x & +3y & -z & = & 1 \end{array}$$

TRIANGULARIZACIÓN

Hacemos ceros por debajo del pivote **2** en la primera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{rrrr} \mathbf{2} & 3 & -1 & 5 \\ \mathbf{4} & 4 & -3 & 3 \\ -\mathbf{2} & 3 & -1 & 1 \end{array} \right) \begin{array}{lcl} f_1' & = & f_1 \\ f_2' & = & f_2 - \frac{\mathbf{4}}{\mathbf{2}} f_1 \\ f_3' & = & f_3 - \frac{-\mathbf{2}}{\mathbf{2}} f_1 \end{array}$$

Un ejemplo de Gauss

TRIANGULARIZACIÓN

Hacemos ceros por debajo del pivote -2 en la segunda columna.

$$\begin{array}{c} f_1' \\ f_2' \\ f_3' \end{array} \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix} \begin{array}{rcl} f_1'' & = & f_1' \\ f_2'' & = & f_2' \\ f_3'' & = & f_3 - \frac{6}{-2} f_2' \end{array}$$

Y ya tenemos una matriz triangular superior (con ceros por debajo de la diagonal principal).

$$\begin{array}{c} f_1'' \\ f_2'' \\ f_3'' \end{array} \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 0 & -5 & -15 \end{pmatrix}$$

Un ejemplo de Gauss

TRIANGULARIZACIÓN

Hacemos ceros por debajo del pivote $-\mathbf{2}$ en la segunda columna.

$$\begin{array}{c} f_1' \\ f_2' \\ f_3' \end{array} \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -\mathbf{2} & -1 & -7 \\ 0 & \mathbf{6} & -2 & 6 \end{pmatrix} \begin{array}{rcl} f_1'' & = & f_1' \\ f_2'' & = & f_2' \\ f_3'' & = & f_3 - \frac{\mathbf{6}}{-\mathbf{2}} f_2' \end{array}$$

Y ya tenemos una matriz triangular superior (con ceros por debajo de la diagonal principal).

$$\begin{array}{c} f_1'' \\ f_2'' \\ f_3'' \end{array} \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 0 & -5 & -15 \end{pmatrix}$$

Un ejemplo de Gauss

SUSTITUCIÓN REGRESIVA

Despejamos las incógnitas empezando por la ecuación de abajo y progresamos hacia arriba.

$$\begin{array}{rrrcr} 2x & +3y & -z & = & 5 \\ & -2y & -z & = & -7 \\ & & -5z & = & -15 \end{array}$$

Empezamos por la z

$$\begin{array}{rcl} z & = & -15/(-5) = 3 \\ y & = & (-7+z)/(-2) = (-7+3)/(-2) = 2 \\ x & = & (5-3y+z)/2 = (5-3(2)+(3))/2 = 1 \end{array}$$

Un ejemplo de Gauss

SUSTITUCIÓN REGRESIVA

Despejamos las incógnitas empezando por la ecuación de abajo y progresamos hacia arriba.

$$\begin{array}{rrrcr} 2x & +3y & -z & = & 5 \\ & -2y & -z & = & -7 \\ & & -5z & = & -15 \end{array}$$

Empezamos por la z

$$\begin{array}{rcl} z & = & -15/(-5) = 3 \\ y & = & (-7 + z)/(-2) = (-7 + 3)/(-2) = 2 \\ x & = & (5 - 3y + z)/2 = (5 - 3(2) + (3))/2 = 1 \end{array}$$

La estrategia del pivote

En la triangularización, en la primera transformación de la matriz, obtenemos ceros por debajo de a_{11}, en el segundo paso hacemos ceros por debajo de a'_{22} y así sucesivamente. Estos elemetos a_{ii} son los pivotes. Hay casos donde se cambian:

- Si es pivote parcial, **intercambiando filas** y tomando como pivote el elemento de mayor valor absoluto a_{ij} que está en la columna del pivote por debajo del pivote.
- Si es pivote total, **intercambiando filas y columnas**.

Ejemplos que requieren pivote:

- cuando $a_{ii} = 0$
- cuando a_{ii} es muy pequeño y se pueden producir grandes errores de redondeo.

Un ejemplo de Gauss con pivote

Resolver un sistema lineal por Gauss con pivote

$$\begin{array}{rrrcr} x & +y & -z & = & 0 \\ 2x & +y & +z & = & 7 \\ 3x & -2y & -z & = & -4 \end{array}$$

TRIANGULARIZACIÓN

En este primer paso buscamos el pivote en la primera columna. Cogemos como pivote el elemento de mayor valor absoluto. Hacemos ceros por debajo del pivote.

$$\begin{pmatrix} \mathbf{1} & 1 & -1 & 0 \\ \mathbf{2} & 1 & 1 & 7 \\ \mathbf{3} & -2 & -1 & -4 \end{pmatrix} \Leftrightarrow \begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} \mathbf{3} & -2 & -1 & -4 \\ \mathbf{2} & 1 & 1 & 7 \\ \mathbf{1} & 1 & -1 & 0 \end{pmatrix} \begin{array}{rcl} f_1' & = & f_1 \\ f_2' & = & f_2 - \frac{2}{3} f_1 \\ f_3' & = & f_3 - \frac{1}{3} f_1 \end{array}$$

Un ejemplo de Gauss con pivote

Resolver un sistema lineal por Gauss con pivote

$$\begin{array}{rrrcr} x & +y & -z & = & 0 \\ 2x & +y & +z & = & 7 \\ 3x & -2y & -z & = & -4 \end{array}$$

TRIANGULARIZACIÓN

En este primer paso buscamos el pivote en la primera columna. Cogemos como pivote el elemento de mayor valor absoluto. Hacemos ceros por debajo del pivote.

$$\left(\begin{array}{rrrr} \mathbf{1} & 1 & -1 & 0 \\ \mathbf{2} & 1 & 1 & 7 \\ \mathbf{3} & -2 & -1 & -4 \end{array}\right) \Leftrightarrow \begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{rrrr} \mathbf{3} & -2 & -1 & -4 \\ \mathbf{2} & 1 & 1 & 7 \\ \mathbf{1} & 1 & -1 & 0 \end{array}\right) \begin{array}{lcl} f_1' & = & f_1 \\ f_2' & = & f_2 - \frac{\mathbf{2}}{\mathbf{3}} f_1 \\ f_3' & = & f_3 - \frac{\mathbf{1}}{\mathbf{3}} f_1 \end{array}$$

TRIANGULARIZACIÓN

Ahora el máximo valor, el pivote **7/3** está en la segunda columna por lo que no hace falta intercambiar filas.

$$\begin{array}{c} f_1' \\ f_2' \\ f_3' \end{array} \left(\begin{array}{cccc} 3 & -2 & -1 & -4 \\ 0 & \mathbf{\frac{7}{3}} & \frac{5}{3} & \frac{29}{3} \\ 0 & \mathbf{\frac{5}{3}} & -\frac{2}{3} & \frac{4}{3} \end{array} \right) \quad \begin{array}{ccl} f_1'' & = & f_1' \\ f_2'' & = & f_2' \\ f_3'' & = & f_3 - \frac{\mathbf{5/3}}{\mathbf{7/3}} f_2' \end{array}$$

Y ya tenemos una matriz triangular superior (con ceros por debajo de la diagonal principal)

$$\begin{array}{c} f_1'' \\ f_2'' \\ f_3'' \end{array} \left(\begin{array}{cccc} 3 & -2 & -1 & -4 \\ 0 & \frac{7}{3} & \frac{5}{3} & \frac{29}{3} \\ 0 & 0 & -\frac{13}{7} & -\frac{39}{7} \end{array} \right)$$

TRIANGULARIZACIÓN

Ahora el máximo valor, el pivote **7/3** está en la segunda columna por lo que no hace falta intercambiar filas.

$$\begin{array}{c} f_1' \\ f_2' \\ f_3' \end{array} \begin{pmatrix} 3 & -2 & -1 & -4 \\ 0 & \frac{7}{3} & \frac{5}{3} & \frac{29}{3} \\ 0 & \frac{5}{3} & -\frac{2}{3} & \frac{4}{3} \end{pmatrix} \quad \begin{array}{rcl} f_1'' & = & f_1' \\ f_2'' & = & f_2' \\ f_3'' & = & f_3 - \frac{5/3}{7/3} f_2' \end{array}$$

Y ya tenemos una matriz triangular superior (con ceros por debajo de la diagonal principal)

$$\begin{array}{c} f_1'' \\ f_2'' \\ f_3'' \end{array} \begin{pmatrix} 3 & -2 & -1 & -4 \\ 0 & \frac{7}{3} & \frac{5}{3} & \frac{29}{3} \\ 0 & 0 & -\frac{13}{7} & -\frac{39}{7} \end{pmatrix}$$

SUSTITUCIÓN REGRESIVA

Despejamos las incógnitas empezando por la ecuación de abajo y progresamos hacia arriba.

$$\begin{array}{rrrcr} 3x & -2y & -z & = & -4 \\ & \frac{7}{3}y & +\frac{5}{3}z & = & \frac{29}{3} \\ & & -\frac{13}{7}z & = & -\frac{39}{7} \end{array}$$

Empezamos con la z

$$\begin{aligned} z &= -(39/7)/(-13/7) = 3 \\ y &= ((29/3) - (5/3)z)/(7/3) = 2 \\ x &= (-4 + 2y + z)/3 = 1 \end{aligned}$$

SUSTITUCIÓN REGRESIVA

Despejamos las incógnitas empezando por la ecuación de abajo y progresamos hacia arriba.

$$\begin{array}{rrrcr} 3x & -2y & -z & = & -4 \\ & \frac{7}{3}y & +\frac{5}{3}z & = & \frac{29}{3} \\ & & -\frac{13}{7}z & = & -\frac{39}{7} \end{array}$$

Empezamos con la z

$$\begin{aligned} z &= -(39/7)/(-13/7) = 3 \\ y &= ((29/3) - (5/3)z)/(7/3) = 2 \\ x &= (-4 + 2y + z)/3 = 1 \end{aligned}$$

Método de Gauss-Jordan

- La idea del método de Gauss-Jordan es transformar el sistema de ecuaciones lineales original en un sistema de diagonal con las mismas soluciones.

- Para ello se realizan las mismas operaciones que en el método de Gauss:
  - $f_i \to f_i + \lambda f_j, \quad j \neq i$
  - $f_i \leftrightarrow f_j$ (si usamos la estrategia del pivote)

- En estos cálculos intervienen la matriz de coeficientes y el vector del segundo miembro.

Un ejemplo de Gauss-Jordan

Resolver un sistema lineal por Gauss-Jordan

$$\begin{array}{rrrcr} 2x & +3y & -z & = & 5 \\ 4x & +4y & -3z & = & 3 \\ -2x & +3y & -z & = & 1 \end{array}$$

Dividimos la primera fila por el pivote 2.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \quad f_1 = f_1/2$$

Un ejemplo de Gauss-Jordan

Resolver un sistema lineal por Gauss-Jordan

$$\begin{array}{rrrcr} 2x & +3y & -z & = & 5 \\ 4x & +4y & -3z & = & 3 \\ -2x & +3y & -z & = & 1 \end{array}$$

Dividimos la primera fila por el pivote **2**.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} \mathbf{2} & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \quad f_1 = f_1/\mathbf{2}$$

Un ejemplo de Gauss-Jordan

Hacemos ceros por debajo del pivote en la primera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \begin{array}{lcl} f_1 & & \\ f_2 & = & f_2 - 4f_1 \\ f_3 & = & f_3 - (-2)f_1 \end{array}$$

Dividimos la segunda fila por el pivote $-\mathbf{2}$.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & -\mathbf{2} & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix} \begin{array}{lcl} f_2 & = & f_2/(-\mathbf{2}) \end{array}$$

Un ejemplo de Gauss-Jordan

Hacemos ceros por debajo del pivote en la primera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix} \begin{array}{lcl} f_1 & & \\ f_2 & = & f_2 - 4f_1 \\ f_3 & = & f_3 - (-2)f_1 \end{array}$$

Dividimos la segunda fila por el pivote $\mathbf{-2}$.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & \mathbf{-2} & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix} \quad f_2 = f_2/(\mathbf{-2})$$

Un ejemplo de Gauss-Jordan

Hacemos ceros por encima y por debajo del pivote en la segunda columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 6 & -2 & 6 \end{pmatrix} \begin{array}{rcl} f_1 & = & f_1 - (3/2)f_2 \\ f_2 & & \\ f_3 & = & f_3 - 6f_2 \end{array}$$

Dividimos la tercera fila por el pivote $-\mathbf{5}$.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & -\mathbf{5} & -15 \end{pmatrix} \begin{array}{rcl} & & \\ & & \\ f_3 & = & f_3/(-\mathbf{5}) \end{array}$$

Un ejemplo de Gauss-Jordan

Hacemos ceros por encima y por debajo del pivote en la segunda columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 6 & -2 & 6 \end{pmatrix} \begin{array}{lcl} f_1 & = & f_1 - (3/2)f_2 \\ f_2 & & \\ f_3 & = & f_3 - 6f_2 \end{array}$$

Dividimos la tercera fila por el pivote $\mathbf{-5}$.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & \mathbf{-5} & -15 \end{pmatrix} \begin{array}{lcl} & & \\ & & \\ f_3 & = & f_3/(\mathbf{-5}) \end{array}$$

Un ejemplo de Gauss-Jordan

Hacemos ceros por encima del pivote en la tercera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & 1 & 3 \end{pmatrix} \begin{array}{lcl} f_1 & = & f_1 - (-5/4) f_3 \\ f_2 & = & f_2 - (1/2) f_3 \\ f_3 & & \end{array}$$

Y el sistema equivalente resultante es

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix} \begin{array}{l} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{array}$$

Es decir, la solución del sistema la da la columna de términos independientes.

Un ejemplo de Gauss-Jordan

Hacemos ceros por encima del pivote en la tercera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccc} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & 1 & 3 \end{array} \right) \begin{array}{lcl} f_1 & = & f_1 - (-5/4) f_3 \\ f_2 & = & f_2 - (1/2) f_3 \\ f_3 & & \end{array}$$

Y el sistema equivalente resultante es

$$\left(\begin{array}{cccc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right) \begin{array}{l} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{array}$$

Es decir, la solución del sistema la da la columna de términos independientes.

Cálculo de la matriz inversa con Gauss-Jordan

La matriz inversa de una matriz A $n \times n$, caso de existir, es una matriz $n \times n$ que llamaremos A^{-1} que verifica $AA^{-1} = I = A^{-1}A$.

Si consideramos que las columnas de A^{-1} e I son

$$A^{-1} = (c_1 c_2 \ldots c_n) \text{ y } I = (e_1 e_2 \ldots e_n)$$

como $AA^{-1} = I$ también se verifica

$$Ac_1 = e_1, \; Ac_2 = e_2, \ldots, \; Ac_n = e_n$$

Es decir, las columnas de la matriz A^{-1} son las soluciones de n sistemas que tienen como matriz de coeficientes la matriz A y como término independiente las columnas de I.

Es decir, si resolvemos simultaneamente estos n sistemas, las respectivas soluciones serán las columnas de A^{-1}.

Un método muy adecuado para resolver simultáneamente sistemas con la misma matriz de coeficientes es **Gauss-Jordan**.

Cálculo de la matriz inversa con Gauss-Jordan

- Escribir una matriz $n \times 2n$ que consiste en la matriz dada A a la izquierda y la matriz identidad I de dimensión $n \times n$ a la derecha $[A|I]$.
- Mediante operaciones por filas, transformar la matriz A en la matriz I, de forma que obenemos a partir de $[A|I]$ obtenemos $\left[I|A^{-1}\right]$.
- Comprobar que $AA^{-1} = I = A^{-1}A$.

Cálculo de la matriz inversa con Gauss-Jordan

Calcular la matriz inversa de A

$$A = \begin{pmatrix} 3 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

Escribimos la matriz $[A|I]$. Dividimos la primera fila por el pivote **3**.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 3 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \quad f_1 = f_1/3$$

Cálculo de la matriz inversa con Gauss-Jordan

Calcular la matriz inversa de A

$$A = \begin{pmatrix} 3 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

Escribimos la matriz$[A|I]$. Dividimos la primera fila por el pivote **3**.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} \mathbf{3} & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \quad f_1 = f_1/\mathbf{3}$$

Cálculo de la matriz inversa con Gauss-Jordan

Hacemos ceros por debajo del pivote en la primera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & \frac{2}{3} & 1 & \frac{1}{3} & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{lcl} f_1 & & \\ f_2 & = & f_2 - 2f_1 \\ f3 & = & f_3 - 3f_1 \end{array}$$

Dividimos la segunda fila por el pivote $-\frac{1}{3}$.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & \frac{2}{3} & 1 & \frac{1}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & -1 & -\frac{2}{3} & 1 & 0 \\ 0 & -1 & -2 & -1 & 0 & 1 \end{array} \right) \begin{array}{lcl} & & \\ f_2 & = & f_2/(-\frac{1}{3}) \\ & & \end{array}$$

Cálculo de la matriz inversa con Gauss-Jordan

Hacemos ceros por debajo del pivote en la primera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & \frac{2}{3} & 1 & \frac{1}{3} & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{ccc} f_1 & & \\ f_2 & = & f_2 - 2f_1 \\ f3 & = & f_3 - 3f_1 \end{array}$$

Dividimos la segunda fila por el pivote $-\frac{1}{3}$.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & \frac{2}{3} & 1 & \frac{1}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & -1 & -\frac{2}{3} & 1 & 0 \\ 0 & -1 & -2 & -1 & 0 & 1 \end{array} \right) \quad f_2 = f_2/(-\frac{1}{3})$$

Cálculo de la matriz inversa con Gauss-Jordan

Hacemos ceros por encima y por debajo del pivote en la segunda columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & \frac{2}{3} & 1 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 3 & 2 & -3 & 0 \\ 0 & -1 & -2 & -1 & 0 & 1 \end{pmatrix} \begin{array}{lcl} f_1 & = & f_1 - (2/3)f_2 \\ f_2 & & \\ f_3 & = & f_3 - (-1)f_2 \end{array}$$

Como el pivote ya es 1, hacemos ceros por encima del pivote en la tercera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & -1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 & -3 & 0 \\ 0 & 0 & 1 & 1 & -3 & 1 \end{pmatrix} \begin{array}{lcl} f_1 & = & f_1 - (-1)f_3 \\ f_2 & = & f_2 - 3f_3 \\ f_3 & & \end{array}$$

Cálculo de la matriz inversa con Gauss-Jordan

Hacemos ceros por encima y por debajo del pivote en la segunda columna.

$$\begin{array}{c} f_1 \\ \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & \dfrac{2}{3} & 1 & \dfrac{1}{3} & 0 & 0 \\ 0 & 1 & 3 & 2 & -3 & 0 \\ 0 & -1 & -2 & -1 & 0 & 1 \end{array} \right) \begin{array}{lcl} f_1 & = & f_1 - (2/3) f_2 \\ f_2 & & \\ f_3 & = & f_3 - (-1) f_2 \end{array}$$

Como el pivote ya es 1, hacemos ceros por encima del pivote en la tercera columna.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & 0 & -1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 & -3 & 0 \\ 0 & 0 & 1 & 1 & -3 & 1 \end{array} \right) \begin{array}{lcl} f_1 & = & f_1 - (-1) f_3 \\ f_2 & = & f_2 - 3 f_3 \\ f_3 & & \end{array}$$

Cálculo de la matriz inversa con Gauss-Jordan

Como a la izquierda ya tenemos la matriz I la matriz de la derecha será A^{-1}.

$$[I|A^{-1}] = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & 6 & -3 \\ 0 & 0 & 1 & 1 & -3 & 1 \end{pmatrix}$$

Comprobamos que es la matriz inversa $AA^{-1} = I = A^{-1}A$

$$AA^{-1} = \begin{pmatrix} 3 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ -1 & 6 & -3 \\ 1 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Cálculo de la matriz inversa con Gauss-Jordan

Como a la izquierda ya tenemos la matriz I la matriz de la derecha será A^{-1}.

$$[I|A^{-1}] = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & 6 & -3 \\ 0 & 0 & 1 & 1 & -3 & 1 \end{array}\right)$$

Comprobamos que es la matriz inversa $AA^{-1} = I = A^{-1}A$

$$AA^{-1} = \begin{pmatrix} 3 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ -1 & 6 & -3 \\ 1 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Factorización LU

- Queremos descomponer la matriz de coeficientes A en una matriz triangular superior U y una matriz triangular inferior L de forma que

$$A = LU$$

- Matrices que admiten factorización LU son, por ejemplo:
  1. Matrices estrictamente diagonal dominantes:
     - por filas: $|a_{ii}| > \sum_{\substack{j=1 \\ j \neq 1}}^{n} |a_{ij}|$ para $i = 1, 2, \dots, n$
     - por columnas $|a_{ii}| > \sum_{\substack{j=1 \\ j \neq 1}}^{n} |a_{ji}|$ para $i = 1, 2, \dots, n$
  2. Matrices simétricas y definidas positivas:

$$A = A^T \quad \text{y} \quad \mathrm{x}^{\mathrm{T}} A \mathrm{x} > 0 \quad \text{para todo } \mathrm{x} \neq 0$$

Factorización LU

- La factorización LU de una matriz no es única
- Habitualmente se fija

$$l_{ii} = 1 \quad \text{para } i = 1, 2, \ldots, n$$

En este caso:

- $U = A^{(n)}$, la matriz triangular superior que se obtiene al triangularizar A como en Gauss
- Para $i = 2, \ldots, n$ y $j = 1, 2, \ldots, i-1$

$$l_{ij} = m_{ij}$$

siendo m_{ij} las λ de $f_i \to f_i + \lambda f_j$,

Factorización LU

- Consideramos el sistema de ecuaciones lineales

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

donde la matriz A admite la factorización LU

- Entonces

$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow LU\mathbf{x} = \mathbf{b}$$

- Para resolver el sistema, hay que
  1. resolver $L\mathbf{y} = \mathbf{b}$
  2. resolver $U\mathbf{x} = \mathbf{y}$

Factorización LU: sustitución progresiva

Si tenemos un sistema

$$\mathit{L}\mathbf{x} = \mathbf{b}$$

donde L es una matriz triangular inferior:

$$L = \begin{pmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{pmatrix}$$

Factorización LU: sustitución progresiva

- La ecuación i-ésima del sistema, para $i = 1, 2, \ldots, n$ es:

$$l_{i1}x_1 + l_{i2}x_2 + \cdots + l_{ii}x_i = b_i$$

- Por lo tanto, para $i = 1, 2, \ldots, n$, si $l_{ii} \neq 0$ entonces

$$x_i = \frac{b_i - l_{i1}x_1 - \cdots - l_{ii-1}x_{i-1}}{l_{ii}} = \frac{1}{l_{ii}}\left(b_i - \sum_{j=1}^{i-1} l_{ij}x_j\right)$$

Factorización LU: sustitución progresiva

- Algoritmo sustitución hacia adelante:
  1. Si $l_{11} = 0$, parar (no existe solución única).
  2. Hacer $x_1 = \dfrac{b_1}{l_{11}}$
  3. Para $i = 2, \ldots, n$
     1. Si $l_{ii} = 0$, parar (no tiene solución)
     2. Hacer

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

Resolver un sistema lineal por factorización LU

$$\begin{array}{rrrcr} x & +y & +z & = & 1 \\ -x & +y & & = & 0 \\ & -2y & +2z & = & -4 \end{array}$$

FACTORIZACIÓN

En el primer paso hacemos ceros por debajo del elemento a_{11}

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ 0 & -2 & 2 \end{pmatrix} \quad \begin{array}{lclcrl} f_1' & = & f_1 & & & \\ f_2' & = & f_2 & - & (-1)/1 & f_1 \\ f_3' & = & f_3 & - & 0/1 & f_1 \end{array}$$

Los multiplicadores, que aparecen en rojo, son los elementos con los que construímos la matriz L. Los insertamos en la matriz, en lugar de los ceros creados.

Resolver un sistema lineal por factorización LU

$$\begin{array}{rrrcr} x & +y & +z & = & 1 \\ -x & +y & & = & 0 \\ & -2y & +2z & = & -4 \end{array}$$

FACTORIZACIÓN

En el primer paso hacemos ceros por debajo del elemento a_{11}

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} \mathbf{1} & 1 & 1 \\ -1 & 1 & 0 \\ 0 & -2 & 2 \end{pmatrix} \qquad \begin{array}{rclcrl} f_1' & = & f_1 & & & \\ f_2' & = & f_2 & - & (-1)/1 & f_1 \\ f_3' & = & f_3 & - & 0/1 & f_1 \end{array}$$

Los multiplicadores, que aparecen en rojo, son los elementos con los que construímos la matriz L. Los insertamos en la matriz, en lugar de los ceros creados.

FACTORIZACIÓN

Repetimos el proceso creando ceros por debajo de a'_{22}.

$$\begin{array}{c} f'_1 \\ f'_2 \\ f'_3 \end{array} \begin{pmatrix} 1 & 1 & 1 \\ -1 & \mathbf{2} & 1 \\ 0 & -2 & 2 \end{pmatrix} \quad \begin{array}{lcl} f''_1 & = & f'_1 \\ f''_2 & = & f'_2 \\ f''_3 & = & f'_3 - (-2/2)\, f'_2 \end{array}$$

FACTORIZACIÓN

Y llegamos a la matriz que almacena simultáneamente L y U.

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

FACTORIZACIÓN

Repetimos el proceso creando ceros por debajo de a'_{22}.

$$\begin{array}{c} f'_1 \\ f'_2 \\ f'_3 \end{array} \begin{pmatrix} 1 & 1 & 1 \\ -1 & \mathbf{2} & 1 \\ 0 & -2 & 2 \end{pmatrix} \qquad \begin{array}{ccccc} f''_1 & = & f'_1 & & \\ f''_2 & = & f'_2 & & \\ f''_3 & = & f'_3 & - & (-2/2) \quad f'_2 \end{array}$$

FACTORIZACIÓN

Y llegamos a la matriz que almacena simultáneamente L y U.

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

FACTORIZACIÓN

La matriz que almacena simultáneamente L y U.

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

FACTORIZACIÓN

Y las matrices L y U son:

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

FACTORIZACIÓN

La matriz que almacena simultáneamente L y U.

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

FACTORIZACIÓN

Y las matrices L y U son:

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

- Tenemos

$$A = LU$$

- Entonces

$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow LU\mathbf{x} = \mathbf{b}$$

- Para resolver el sistema, hay que

1. resolver $L\mathbf{y} = \mathbf{b}$ por **sustitución progresiva**
2. resolver $U\mathbf{x} = \mathbf{y}$ por **sustitución regresiva**

Sustitución Progresiva

Resolvemos el sistema triangular inferior $Ly = b$ y obtenemos y.

$$\begin{array}{rrrcr} y_1 & & & = & 1 \\ -y_1 & +y_2 & & = & 0 \\ & -y_2 & +y_3 & = & -4 \end{array} \qquad \begin{array}{rcrrcr} y_1 & = & 1 & & & \\ y_2 & = & y_1 & & = & 1 \\ y_3 & = & -4 & +y_2 & = & -3 \end{array}$$

Sustitución Regresiva

Resolvemos el sistema triangular superior $Ux = y$ y obtenemos x, que era lo que buscábamos.

$$\begin{array}{rrrcr} x_1 & +x_2 & +x_3 & = & 1 \\ & 2x_2 & +x_3 & = & 1 \\ & & 3x_3 & = & -3 \end{array} \qquad \begin{array}{rcrcr} x_3 & = & -3/3 & = & -1 \\ x_2 & = & (1 - x_3)/2 & = & 1 \\ x_1 & = & 1 - x_2 - x_3 & = & 1 \end{array}$$

Sustitución Progresiva

Resolvemos el sistema triangular inferior $Ly = b$ y obtenemos y.

$$\begin{array}{rrrcr} y_1 & & & = & 1 \\ -y_1 & +y_2 & & = & 0 \\ & -y_2 & +y_3 & = & -4 \end{array} \qquad \begin{array}{rcrrcr} y_1 & = & 1 & & & \\ y_2 & = & {\color{red}y_1} & & = & 1 \\ y_3 & = & -4 & +{\color{red}y_2} & = & -3 \end{array}$$

Sustitución Regresiva

Resolvemos el sistema triangular superior $Ux = y$ y obtenemos x, que era lo que buscábamos.

$$\begin{array}{rrrcr} x_1 & +x_2 & +x_3 & = & 1 \\ & 2x_2 & +x_3 & = & 1 \\ & & 3x_3 & = & -3 \end{array} \qquad \begin{array}{rcrcr} x_3 & = & -3/3 & = & -1 \\ x_2 & = & (1 - {\color{red}x_3})/2 & = & 1 \\ x_1 & = & 1 - {\color{red}x_2} - {\color{red}x_3} & = & 1 \end{array}$$

Métodos iterativos: ventajas

- Los métodos iterativos son, en general, **más eficientes** que los métodos directos para resolver sistemas de ecuaciones lineales **grandes y de matriz hueca** ya que se basan en la operación “multiplicación de matriz por vector”.
- Como consecuencia:
  - Sólo es necesario almacenar los coeficientes no nulos de la matriz del sistema.
  - Si no se exige mucha precisión, se puede obtener una aproximación aceptable en un número pequeño de iteraciones.
  - Son menos sensibles a los errores de redondeo.

Métodos iterativos: desventajas

- En general, **no es posible predecir el número de operaciones** que se requieren para obtener una aproximación a la solución con una precisión determinada.
- El tiempo de cálculo y la precisión del resultado pueden depender de la elección de ciertos **parámetros**.
- Generalmente no se gana tiempo por iteración si la matriz de coeficientes es simétrica (como en el método directo de Cholesky). En este caso, un método directo genérico puede reducir el tiempo de cálculo a la mitad.

Métodos iterativos

- Dada una aproximación inicial $\mathbf{x}^{(0)}$, un método iterativo **genera una sucesión de aproximaciones** $\mathbf{x}^{(k)}$, para $k = 1, 2, \ldots$, que converge a la solución del sistema.

- Para generar esta sucesión, se repite el mismo esquema de operaciones hasta que:

  - se obtiene una aproximación a la solución con una precisión especificada de antemano, o
  - se rebasa un número máximo de iteraciones

Métodos iterativos

- Los métodos iterativos clásicos (lineales) se basan en reescribir el problema
$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow \mathbf{x} = G\mathbf{x} + \mathbf{c}$$
donde G es una matriz $n \times n$ y $\mathbf{c}$ es un vector columna de dimensión n.

- **Algoritmo**:
  - Sea $\mathbf{x}^{(0)}$ una aproximación inicial a la solución
  - Para $k = 1, 2, \ldots$
$$\mathbf{x}^{(k)} = G\mathbf{x}^{(k-1)} + \mathbf{c}$$

- La matriz G se llama matriz de iteración y el vector $\mathbf{c}$ se llama vector de iteración.

Método de Jacobi

Se basa en la descomposición

$$A = L + D + U$$

donde

$$L = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{pmatrix} \; D = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \; U = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Método de Jacobi

- Tenemos que

$$
\begin{aligned}
&A\mathbf{x} = \mathbf{b} \Rightarrow (L + D + U)\,\mathbf{x} = \mathbf{b} \Rightarrow \\
&D\mathbf{x} = -\,(L + U)\,\mathbf{x} + \mathbf{b} \Rightarrow \mathbf{x} = -D^{-1}\,(L + U)\,\mathbf{x} + D^{-1}\mathbf{b}
\end{aligned}
$$

- Y por lo tanto

$$
\mathbf{x}^{(k+1)} = -D^{-1}\,(L + D)\,\mathbf{x}^{(k)} + D^{-1}\mathbf{b}
$$

Método de Jacobi

- **Algoritmo de Jacobi**

  1. Elegir una aproximación inicial $\mathbf{x}^{(0)}$
  2. Para $k = 1, 2, \ldots, MaxIter$
     1. Para $i = 1, 2, \ldots, n$, calcular

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

     2. Si se cumple el criterio de parada, tomar $\mathbf{x}^{(k)}$ como aproximación a la solución

Resolver un sistema lineal por Jacobi

$$\begin{array}{rrrrcr} 10x_1 & -x_2 & +2x_3 & & = & 6 \\ -x_1 & +11x_2 & -x_3 & +3x_4 & = & 6 \\ 2x_1 & -x_2 & +10x_3 & -x_4 & = & 11 \\ 3x_2 & -x_3 & & +8x_4 & = & 15 \end{array}$$

Despejamos las incógnitas

Primero despejamos x_1 en la primera ecuación, x_2 en la segunda, ...

$$\begin{aligned} x_1 &= (6 + x_2 - 2x_3)/10 \\ x_2 &= (6 + x_1 + x_3 - 3x_4)/11 \\ x_3 &= (11 - 2x_1 + x_2 + x_4)/10 \\ x_4 &= (15 - 3x_2 + x_3)/8 \end{aligned}$$

Resolver un sistema lineal por Jacobi

$$
\begin{array}{rrrrcr}
10x_1 & -x_2 & +2x_3 & & = & 6 \\
-x_1 & +11x_2 & -x_3 & +3x_4 & = & 6 \\
2x_1 & -x_2 & +10x_3 & -x_4 & = & 11 \\
3x_2 & -x_3 & & +8x_4 & = & 15
\end{array}
$$

Despejamos las incógnitas

Primero despejamos x_1 en la primera ecuación, x_2 en la segunda, ...

$$
\begin{array}{rcl}
x_1 & = & (6 + x_2 - 2x_3)/10 \\
x_2 & = & (6 + x_1 + x_3 - 3x_4)/11 \\
x_3 & = & (11 - 2x_1 + x_2 + x_4)/10 \\
x_4 & = & (15 - 3x_2 + x_3)/8
\end{array}
$$

Despejamos las incógnitas

Teniendo en cuenta las ecuaciones

$$\begin{aligned}
x_1 &= (6 + x_2 - 2x_3)/10 \\
x_2 &= (6 + x_1 + x_3 - 3x_4)/11 \\
x_3 &= (11 - 2x_1 + x_2 + x_4)/10 \\
x_4 &= (15 - 3x_2 + x_3)/8
\end{aligned}$$

Definimos la suceción

$$\begin{aligned}
x_1^{(k+1)} &= (6 + x_2^{(k)} - 2x_3^{(k)})/10 \\
x_2^{(k+1)} &= (6 + x_1^{(k)} + x_3^{(k)} - 3x_4^{(k)})/11 \\
x_3^{(k+1)} &= (11 - 2x_1^{(k)} + x_2^{(k)} + x_4^{(k)})/10 \\
x_4^{(k+1)} &= (15 - 3x_2^{(k)} + x_3^{(k)})/8
\end{aligned}$$

Despejamos las incógnitas

Teniendo en cuenta las ecuaciones

$$\begin{aligned} x_1 &= (6 + x_2 - 2x_3)/10 \\ x_2 &= (6 + x_1 + x_3 - 3x_4)/11 \\ x_3 &= (11 - 2x_1 + x_2 + x_4)/10 \\ x_4 &= (15 - 3x_2 + x_3)/8 \end{aligned}$$

Definimos la suceción

$$\begin{aligned} x_1^{(k+1)} &= (6 + x_2^{(k)} - 2x_3^{(k)})/10 \\ x_2^{(k+1)} &= (6 + x_1^{(k)} + x_3^{(k)} - 3x_4^{(k)})/11 \\ x_3^{(k+1)} &= (11 - 2x_1^{(k)} + x_2^{(k)} + x_4^{(k)})/10 \\ x_4^{(k+1)} &= (15 - 3x_2^{(k)} + x_3^{(k)})/8 \end{aligned}$$

Primera iteración

Punto inicial $\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}\right)^t = (0,0,0,0)^t$.

$$
\begin{aligned}
x_1^{(1)} &= (6 + x_2^{(0)} - 2x_3^{(0)})/10 = 0.6 \\
x_2^{(1)} &= (6 + x_1^{(0)} + x_3^{(0)} - 3x_4^{(0)})/11 = 0.545 \\
x_3^{(1)} &= (11 - 2x_1^{(0)} + x_2^{(0)} + x_4^{(0)})/10 = 1.1 \\
x_4^{(1)} &= (15 - 3x_2^{(0)} + x_3^{(0)})/8 = 1.875
\end{aligned}
$$

Criterio de parada

Impondremos que la distancia entre dos puntos consecutivos sea menor que una 0.01.

$$\left\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\right\|_\infty = \max_{1\le i\le 4}\left(\left|\mathbf{x}_i^{(1)} - \mathbf{x}_i^{(0)}\right|\right) = \max(0.6, 0.545, 1.1, 1.875) = 1.875$$

Primera iteración

Punto inicial $\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}\right)^t = (0,0,0,0)^t$.

$$
\begin{aligned}
x_1^{(1)} &= (6 + x_2^{(0)} - 2x_3^{(0)})/10 = 0.6 \\
x_2^{(1)} &= (6 + x_1^{(0)} + x_3^{(0)} - 3x_4^{(0)})/11 = 0.545 \\
x_3^{(1)} &= (11 - 2x_1^{(0)} + x_2^{(0)} + x_4^{(0)})/10 = 1.1 \\
x_4^{(1)} &= (15 - 3x_2^{(0)} + x_3^{(0)})/8 = 1.875
\end{aligned}
$$

Criterio de parada

Impondremos que la distancia entre dos puntos consecutivos sea menor que una 0.01.

$$\left\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\right\|_\infty = \max_{1\le i\le 4}\left(\left|\mathbf{x}_i^{(1)} - \mathbf{x}_i^{(0)}\right|\right) = \max(0.6, 0.545, 1.1, 1.875) = 1.875$$

Segunda iteración

$$
\begin{aligned}
x_1^{(2)} &= (6 + x_2^{(1)} - 2x_3^{(1)})/10 && = (6 + 0.545 - 2(1.1))/10 = 0.435 \\
x_2^{(2)} &= (6 + x_1^{(1)} + x_3^{(1)} - 3x_4^{(1)})/11 && = (6 + 0.6 + 1.1 - 3(1.875))/11 = 1.886 \\
x_3^{(2)} &= (11 - 2x_1^{(1)} + x_2^{(1)} + x_4^{(1)})/10 && = (11 - 2(0.6) + 0.545 + (1.875))/10 = 1.22 \\
x_4^{(2)} &= (15 - 3x_2^{(1)} + x_3^{(1)})/8 && = (15 - 3(0.545) + 1.1)/8 = 1.808
\end{aligned}
$$

Criterio de parada

¿Se cumple? No

$$\left\| \mathbf{x}^{(2)} - \mathbf{x}^{(1)} \right\|_\infty = 0.357 > 0.01$$

Segunda iteración

$$
\begin{aligned}
x_1^{(2)} &= (6 + x_2^{(1)} - 2x_3^{(1)})/10 && = (6 + 0.545 - 2(1.1))/10 = 0.435 \\
x_2^{(2)} &= (6 + x_1^{(1)} + x_3^{(1)} - 3x_4^{(1)})/11 && = (6 + 0.6 + 1.1 - 3(1.875))/11 = 1.886 \\
x_3^{(2)} &= (11 - 2x_1^{(1)} + x_2^{(1)} + x_4^{(1)})/10 && = (11 - 2(0.6) + 0.545 + (1.875))/10 = 1.22 \\
x_4^{(2)} &= (15 - 3x_2^{(1)} + x_3^{(1)})/8 && = (15 - 3(0.545) + 1.1)/8 = 1.808
\end{aligned}
$$

Criterio de parada

¿Se cumple? No

$$\left\|\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\right\|_\infty = 0.357 > 0.01$$

Más iteraciones

Hasta que se cumpla la condición de parada

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{x}^{(1)} = \begin{pmatrix} 0.6 \\ 0.545 \\ 1.1 \\ 1.875 \end{pmatrix} \quad x^{(2)} = \begin{pmatrix} 0.435 \\ 0.189 \\ 1.222 \\ 1.808 \end{pmatrix} \quad x^{(3)} = \begin{pmatrix} 0.374 \\ 0.203 \\ 1.213 \\ 1.957 \end{pmatrix}$$

$$x^{(4)} = \begin{pmatrix} 0.378 \\ 0.156 \\ 1.241 \\ 1.950 \end{pmatrix} \dots \; x^{(6)} = \begin{pmatrix} 0.369 \\ 0.153 \\ 1.240 \\ 1.979 \end{pmatrix} \qquad \text{con} \qquad x = \begin{pmatrix} 0.368 \\ 0.154 \\ 1.239 \\ 1.972 \end{pmatrix}$$

Criterio de parada

Se cumple en la iteración 6.

$$\left\| \mathbf{x}^{(6)} - \mathbf{x}^{(5)} \right\|_\infty = 0.007 < 0.01$$

Más iteraciones

Hasta que se cumpla la condición de parada

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{x}^{(1)} = \begin{pmatrix} 0.6 \\ 0.545 \\ 1.1 \\ 1.875 \end{pmatrix} \quad x^{(2)} = \begin{pmatrix} 0.435 \\ 0.189 \\ 1.222 \\ 1.808 \end{pmatrix} \quad x^{(3)} = \begin{pmatrix} 0.374 \\ 0.203 \\ 1.213 \\ 1.957 \end{pmatrix}$$

$$x^{(4)} = \begin{pmatrix} 0.378 \\ 0.156 \\ 1.241 \\ 1.950 \end{pmatrix} \dots \; x^{(6)} = \begin{pmatrix} 0.369 \\ 0.153 \\ 1.240 \\ 1.979 \end{pmatrix} \qquad \text{con} \qquad x = \begin{pmatrix} 0.368 \\ 0.154 \\ 1.239 \\ 1.972 \end{pmatrix}$$

Criterio de parada

Se cumple en la iteración 6.

$$\left\| \mathbf{x}^{(6)} - \mathbf{x}^{(5)} \right\|_\infty = 0.007 < 0.01$$

Método de Gauss-Seidel

- Tenemos que

$$A\mathbf{x} = \mathbf{b} \Rightarrow (L + D + U)\,\mathbf{x} = \mathbf{b} \Rightarrow$$
$$(L + D)\mathbf{x} = -U\mathbf{x} + \mathbf{b} \Rightarrow \mathbf{x} = -(L + D)^{-1}U\mathbf{x} + (L + D)^{-1}\mathbf{b}$$

- Y por lo tanto

$$\mathbf{x}^{(k+1)} = -(L + D)^{-1}U\mathbf{x}^{(k)} + (L + D)^{-1}\mathbf{b}$$

Método de Gauss-Seidel

- **Algoritmo de Jacobi**

  1. Elegir una aproximación inicial $\mathbf{x}^{(0)}$
  2. Para $k = 1, 2, \ldots, MaxIter$
     1. Para $i = 1, 2, \ldots, n$, calcular

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

     2. Si se cumple el criterio de parada, tomar $\mathbf{x}^{(k)}$ como aproximación a la solución

Método de Gauss-Seidel

Observar que:

- **Jacobi**

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

- **Gauss-Seidel**

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

Resolver un sistema lineal por Gauss-Seidel

$$
\begin{array}{rrrrcr}
10x_1 & -x_2 & +2x_3 & & = & 6 \\
-x_1 & +11x_2 & -x_3 & +3x_4 & = & 6 \\
2x_1 & -x_2 & +10x_3 & -x_4 & = & 11 \\
3x_2 & -x_3 & & +8x_4 & = & 15
\end{array}
$$

Despejamos las incógnitas

Como en Jacobi. Primero despejamos x_1 en la primera ecuación, x_2 en la segunda, ...

$$
\begin{array}{rcl}
x_1 & = & (6 + x_2 - 2x_3)/10 \\
x_2 & = & (6 + x_1 + x_3 - 3x_4)/11 \\
x_3 & = & (11 - 2x_1 + x_2 + x_4)/10 \\
x_4 & = & (15 - 3x_2 + x_3)/8
\end{array}
$$

Resolver un sistema lineal por Gauss-Seidel

$$
\begin{array}{rrrrcr}
10x_1 & -x_2 & +2x_3 & & = & 6 \\
-x_1 & +11x_2 & -x_3 & +3x_4 & = & 6 \\
2x_1 & -x_2 & +10x_3 & -x_4 & = & 11 \\
3x_2 & -x_3 & & +8x_4 & = & 15
\end{array}
$$

Despejamos las incógnitas

Como en Jacobi. Primero despejamos x_1 en la primera ecuación, x_2 en la segunda, ...

$$
\begin{aligned}
x_1 &= (6 + x_2 - 2x_3)/10 \\
x_2 &= (6 + x_1 + x_3 - 3x_4)/11 \\
x_3 &= (11 - 2x_1 + x_2 + x_4)/10 \\
x_4 &= (15 - 3x_2 + x_3)/8
\end{aligned}
$$

Teniendo en cuenta las ecuaciones

$$\begin{aligned}
x_1 &= (6 + x_2 - 2x_3)/10\\
x_2 &= (6 + x_1 + x_3 - 3x_4)/11\\
x_3 &= (11 - 2x_1 + x_2 + x_4)/10\\
x_4 &= (15 - 3x_2 + x_3)/8
\end{aligned}$$

Definimos la sucesión

Pero ahora, en cuanto calculamos una aproximación de una de las incógnitas ya la usamos

$$\begin{aligned}
x_1^{(k+1)} &= (6 + x_2^{(k)} - 2x_3^{(k)})/10\\
x_2^{(k+1)} &= (6 + x_1^{(k+1)} + x_3^{(k)} - 3x_4^{(k)})/11\\
x_3^{(k+1)} &= (11 - 2x_1^{(k+1)} + x_2^{(k+1)} + x_4^{(k)})/10\\
x_4^{(k+1)} &= (15 - 3x_2^{(k+1)} + x_3^{(k+1)})/8
\end{aligned}$$

Teniendo en cuenta las ecuaciones

$$\begin{aligned} x_1 &= (6 + x_2 - 2x_3)/10 \\ x_2 &= (6 + x_1 + x_3 - 3x_4)/11 \\ x_3 &= (11 - 2x_1 + x_2 + x_4)/10 \\ x_4 &= (15 - 3x_2 + x_3)/8 \end{aligned}$$

Definimos la sucesión

Pero ahora, en cuanto calculamos una aproximación de una de las incógnitas ya la usamos

$$\begin{aligned} x_1^{(k+1)} &= (6 + x_2^{(k)} - 2x_3^{(k)})/10 \\ x_2^{(k+1)} &= (6 + x_1^{(k+1)} + x_3^{(k)} - 3x_4^{(k)})/11 \\ x_3^{(k+1)} &= (11 - 2x_1^{(k+1)} + x_2^{(k+1)} + x_4^{(k)})/10 \\ x_4^{(k+1)} &= (15 - 3x_2^{(k+1)} + x_3^{(k+1)})/8 \end{aligned}$$

Primera iteración

Y comenzamos también con $\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}\right)^t = (0, 0, 0, 0)^t$.

$$
\begin{aligned}
x_1^{(1)} &= = (6 + x_2^{(0)} - 2x_3^{(0)})/10 & &= 0.6(6 + 0 - 0)/10 = 0.6 \\
x_2^{(1)} &= (6 + x_1^{(1)} + x_3^{(0)} - 3x_4^{(0)})/11 & &= (6 + 0.6 + 0 - 0)/11 = 0.6 \\
x_3^{(1)} &= = (11 - 2x_1^{(1)} + x_2^{(1)} + x_4^{(0)})/10 & &= 1.1(11 - 2(0.6) + (0.6) + 0)/10 = 1.04 \\
x_4^{(1)} &= (15 - 3x_2^{(1)} + x_3^{(1)})/8 & &= (15 - 3(0.6) + (1.04))/8 = 1.78
\end{aligned}
$$

Criterio de parada

Impondremos que la distancia entre dos puntos consecutivos sea menor que una 0.01.

$$\left\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\right\|_\infty = \max_{1 \le i \le 4} \left(\left|\mathbf{x}_i^{(1)} - \mathbf{x}_i^{(0)}\right|\right) = \max(0.6, 0.6, 1.04, 1.78) = 1.78$$

Primera iteración

Y comenzamos también con $\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}\right)^t = (0, 0, 0, 0)^t$.

$$
\begin{aligned}
x_1^{(1)} &= = (6 + x_2^{(0)} - 2x_3^{(0)})/10 & &= 0.6(6 + 0 - 0)/10 = 0.6 \\
x_2^{(1)} &= (6 + x_1^{(1)} + x_3^{(0)} - 3x_4^{(0)})/11 & &= (6 + 0.6 + 0 - 0)/11 = 0.6 \\
x_3^{(1)} &= = (11 - 2x_1^{(1)} + x_2^{(1)} + x_4^{(0)})/10 & &= 1.1(11 - 2(0.6) + (0.6) + 0)/10 = 1.04 \\
x_4^{(1)} &= (15 - 3x_2^{(1)} + x_3^{(1)})/8 & &= (15 - 3(0.6) + (1.04))/8 = 1.78
\end{aligned}
$$

Criterio de parada

Impondremos que la distancia entre dos puntos consecutivos sea menor que una 0.01.

$$
\left\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\right\|_\infty = \max_{1\le i\le 4}\left(\left|\mathbf{x}_i^{(1)} - \mathbf{x}_{\mathrm{i}}^{(0)}\right|\right) = \max(0.6, 0.6, 1.04, 1.78) = 1.78
$$

Segunda iteración

$$
\begin{aligned}
x_1^{(2)} &= (6 + x_2^{(1)} - 2x_3^{(1)})/10 && = (6 + 0.6 - 2(1.04))/10 = 0.452 \\
x_2^{(2)} &= (6 + x_1^{(2)} + x_3^{(1)} - 3x_4^{(1)})/11 && = (6 + 0.452 + 1.04 - 3(1.78))/11 = 0.196 \\
x_3^{(2)} &= (11 - 2x_1^{(2)} + x_2^{(2)} + x_4^{(1)})/10 && = (11 - 2(0.452) + 0.196 + (1.78))/10 = 1.20 \\
x_4^{(2)} &= (15 - 3x_2^{(2)} + x_3^{(2)})/8 && = (15 - 3(0.196) + 1.207)/8 = 1.953
\end{aligned}
$$

Criterio de parada

¿Se cumple? No

$$\left\| \mathbf{x}^{(2)} - \mathbf{x}^{(1)} \right\|_\infty = 0.404 > 0.01$$

Segunda iteración

$$
\begin{aligned}
x_1^{(2)} &= (6 + x_2^{(1)} - 2x_3^{(1)})/10 && = (6 + 0.6 - 2(1.04))/10 = 0.452 \\
x_2^{(2)} &= (6 + x_1^{(2)} + x_3^{(1)} - 3x_4^{(1)})/11 && = (6 + 0.452 + 1.04 - 3(1.78))/11 = 0.196 \\
x_3^{(2)} &= (11 - 2x_1^{(2)} + x_2^{(2)} + x_4^{(1)})/10 && = (11 - 2(0.452) + 0.196 + (1.78))/10 = 1.2 \\
x_4^{(2)} &= (15 - 3x_2^{(2)} + x_3^{(2)})/8 && = (15 - 3(0.196) + 1.207)/8 = 1.953
\end{aligned}
$$

Criterio de parada

¿Se cumple? No

$$\left\| \mathbf{x}^{(2)} - \mathbf{x}^{(1)} \right\|_\infty = 0.404 > 0.01$$

Más iteraciones

Hasta que se cumpla la condición de parada

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{x}^{(1)} = \begin{pmatrix} 0.6 \\ 0.6 \\ 1.04 \\ 1.78 \end{pmatrix} \quad \mathbf{x}^{(2)} = \begin{pmatrix} 0.452 \\ 0.196 \\ 1.207 \\ 1.953 \end{pmatrix} \quad \mathbf{x}^{(3)} = \begin{pmatrix} 0.378 \\ 0.157 \\ 1.235 \\ 1.971 \end{pmatrix}$$

$$\mathbf{x}^{(4)} = \begin{pmatrix} 0.369 \\ 0.154 \\ 1.239 \\ 1.972 \end{pmatrix} \qquad \text{con} \qquad \mathbf{x} = \begin{pmatrix} 0.368 \\ 0.154 \\ 1.239 \\ 1.972 \end{pmatrix}$$

Criterio de parada

Se cumple en la iteración 4.

$$\left\| \mathbf{x}^{(4)} - \mathbf{x}^{(3)} \right\|_\infty = 0.009 < 0.01$$

Más iteraciones

Hasta que se cumpla la condición de parada

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{x}^{(1)} = \begin{pmatrix} 0.6 \\ 0.6 \\ 1.04 \\ 1.78 \end{pmatrix} \quad \mathbf{x}^{(2)} = \begin{pmatrix} 0.452 \\ 0.196 \\ 1.207 \\ 1.953 \end{pmatrix} \quad \mathbf{x}^{(3)} = \begin{pmatrix} 0.378 \\ 0.157 \\ 1.235 \\ 1.971 \end{pmatrix}$$

$$\mathbf{x}^{(4)} = \begin{pmatrix} 0.369 \\ 0.154 \\ 1.239 \\ 1.972 \end{pmatrix} \qquad \text{con} \qquad \mathbf{x} = \begin{pmatrix} 0.368 \\ 0.154 \\ 1.239 \\ 1.972 \end{pmatrix}$$

Criterio de parada

Se cumple en la iteración 4.

$$\left\| \mathbf{x}^{(4)} - \mathbf{x}^{(3)} \right\|_\infty = 0.009 < 0.01$$

Convergencia de los métodos iterativos

Los métodos iterativos clásicos (lineales) se basan en reescribir el problema

$$A\mathbf{x} = \mathbf{b} \Longleftrightarrow \mathbf{x} = G\mathbf{x} + \mathbf{c}$$

Si tomamos un $\mathbf{x}^{(0)}$ como aproximación inicial a la solución, para $k = 1, 2, \ldots$ podemos definir el algoritmo como las sucesión recursiva

$$\mathbf{x}^{(k)} = G\mathbf{x}^{(k-1)} + \mathbf{c}$$

Estudio de la convergencia

Definición

Una matriz A se dice estrictamente diagonal dominante si

$$\sum_{j=1 (i \neq j)}^{n} |a_{ij}| < |a_{ii}| \quad \text{para} \quad i = 1, 2, \ldots, n$$

Teorema

Una condición suficiente de convergencia de los métodos de Jacobi y Gauss-Seidel es que la **matriz de coeficientes** del sistema lineal sea estrictamente **diagonal dominante**.

Estudio de la convergencia

Definición

Una matriz A se dice estrictamente diagonal dominante si

$$\sum_{j=1 (i \neq j)}^{n} |a_{ij}| < |a_{ii}| \quad \text{para} \quad i = 1, 2, \ldots, n$$

Teorema

Una condición suficiente de convergencia de los métodos de Jacobi y Gauss-Seidel es que la **matriz de coeficientes** del sistema lineal sea estrictamente **diagonal dominante**.

Estudio de la convergencia

Ejemplo

¿Podemos garantizar la convergencia del método de Jacobi y Gauss-Seidel utilizando el criterio del teorema en el sistema $Ax = b$

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 3 \end{pmatrix}?$$

Estudiamos si A es diagonal dominante

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 3 \end{pmatrix} \quad \begin{array}{l} |2| > |1| + |0| \\ |5| > |2| + |-1| \\ |3| > |0| + |-1| \end{array}$$

Y A es diagonal dominate. Por lo tanto, usando este criterio podemos garantizar la convergencia tanto de ambos métodos.

Estudio de la convergencia

Ejemplo

¿Podemos garantizar la convergencia del método de Jacobi y Gauss-Seidel utilizando el criterio del teorema en el sistema $Ax = b$

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 3 \end{pmatrix}?$$

Estudiamos si A es diagonal dominante

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 5 & -1 \\ 0 & -1 & 3 \end{pmatrix} \quad \begin{array}{l} |2| > |1| + |0| \\ |5| > |2| + |-1| \\ |3| > |0| + |-1| \end{array}$$

Y A es diagonal dominate. Por lo tanto, usando este criterio podemos garantizar la convergencia tanto de ambos métodos.

Estudio de la convergencia

Definición

El radio espectral de una matriz G es el módulo (o valor absoluto) máximo $|\lambda_i|$ de los autovalores de G.

Teorema

Dado un sistema lineal $\mathbf{x} = G\mathbf{x} + \mathbf{c}$, donde $I - G$ es inversible, las siguientes afirmaciones son equivalentes:

- El método iterativo asociado $\mathbf{x}^{(k)} = G\mathbf{x}^{(k-1)} + \mathbf{c} \quad \text{para} \quad k = 1, 2, \ldots, n$ es convergente.
- El radio espectral de G es menor que 1.
- Alguna norma de G es menor que 1.

Estudio de la convergencia

Definición

El radio espectral de una matriz G es el módulo (o valor absoluto) máximo $|\lambda_i|$ de los autovalores de G.

Teorema

Dado un sistema lineal $\mathbf{x} = G\mathbf{x} + \mathbf{c}$, donde $I - G$ es inversible, las siguientes afirmaciones son equivalentes:

- El método iterativo asociado $\mathbf{x}^{(k)} = G\mathbf{x}^{(k-1)} + \mathbf{c} \quad \text{para} \quad k = 1, 2, \ldots, n$ es convergente.
- El radio espectral de G es menor que 1.
- Alguna norma de G es menor que 1.

Estudio de convergencia

Ejemplo

Estudiar si el método de Gauss-Seidel sería convergente para el sistema lineal $Ax = b$

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & -1 \\ 3 & 1 & 3 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

La matriz $B_{G-S} = -(L + D)^{-1} U$ con

$$A = L + D + U = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 3 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Estudio de convergencia

Ejemplo

Estudiar si el método de Gauss-Seidel sería convergente para el sistema lineal $Ax = b$

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & -1 \\ 3 & 1 & 3 \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

La matriz $B_{G-S} = -(L + D)^{-1} U$ con

$$A = L + D + U = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 3 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Estudio de convergencia

Por lo que

$$L + D = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 3 & 1 & 3 \end{pmatrix} \quad y \quad U = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Calculemos por Gauss-Jordan $(L+D)^{-1}$. Dividimos la primera fila por el pivote

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 3 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 1 & 3 & 0 & 0 & 1 \end{pmatrix} \qquad f_1 \rightarrow f_1/3$$

Estudio de convergencia

Por lo que

$$L + D = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 3 & 1 & 3 \end{pmatrix} \quad y \quad U = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Calculemos por Gauss-Jordan $(L+D)^{-1}$. Dividimos la primera fila por el pivote

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 3 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 1 & 3 & 0 & 0 & 1 \end{pmatrix} \qquad f_1 \;\rightarrow\; f_1/3$$

Estudio de convergencia

Hacemos ceros por debajo del pivote.

$$\begin{array}{c} f_1 \\ \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 1 & 3 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{array}{ccl} f_1 & \to & f_1 \\ f_2 & \to & f_2 - f_1 \\ f_3 & \to & f_3 - 3f_1 \end{array}$$

Dividimos la segunda fila por el pivote.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 2 & 0 & -\frac{1}{3} & 1 & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{pmatrix} \qquad f_2 \to f_2/2$$

Estudio de convergencia

Hacemos ceros por debajo del pivote.

$$\begin{array}{c} f_1 \\ \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 1 & 3 & 0 & 0 & 1 \end{array} \right) \quad \begin{array}{ccc} f_1 & \rightarrow & f_1 \\ f_2 & \rightarrow & f_2 - f_1 \\ f_3 & \rightarrow & f_3 - 3f_1 \end{array}$$

Dividimos la segunda fila por el pivote.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \left(\begin{array}{cccccc} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 2 & 0 & -\frac{1}{3} & 1 & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{array} \right) \quad f_2 \rightarrow f_2/2$$

Estudio de convergencia

Hacemos cero por debajo del pivote (por encima ya tenemos ceros).

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{pmatrix} \quad f_3 \to f_3 - f_2$$

Dividimos la tercera fila por el pivote.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 3 & -\frac{5}{6} & -\frac{1}{2} & 1 \end{pmatrix} \quad f_3 \to f_3/3$$

Estudio de convergencia

Hacemos cero por debajo del pivote (por encima ya tenemos ceros).

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{pmatrix} \qquad f_3 \quad \rightarrow \quad f_3 - f_2$$

Dividimos la tercera fila por el pivote.

$$\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 3 & -\frac{5}{6} & -\frac{1}{2} & 1 \end{pmatrix} \qquad f_3 \quad \rightarrow \quad f_3/3$$

Estudio de convergencia

Ya tenemos a la izquierda la matriz identidad.

$$\begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -\frac{5}{18} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

Por lo tanto

$$(L+D)^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{6} & \frac{1}{2} & 0 \\ -\frac{5}{18} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

Estudio de convergencia

Ya tenemos a la izquierda la matriz identidad.

$$\begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -\frac{5}{18} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

Por lo tanto

$$(L+D)^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{6} & \frac{1}{2} & 0 \\ -\frac{5}{18} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

Estudio de convergencia

Ya podemos calcular la matriz de iteración $B_{G-S} = -(L+D)^{-1}U$

$$B_{G-S} = -\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ -\frac{1}{6} & \frac{1}{2} & 0 \\ -\frac{5}{18} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{6} & \frac{2}{3} \\ 0 & \frac{5}{18} & \frac{1}{9} \end{pmatrix}$$

Estudio de convergencia

Calculemos la norma 1 de esta matriz.

$$B_{G-S}^{T} = \begin{pmatrix} 0 & 0 & 0 \\ -\frac{1}{3} & \frac{1}{6} & \frac{5}{18} \\ -\frac{1}{3} & \frac{2}{3} & \frac{1}{9} \end{pmatrix} \quad \begin{array}{l} 0+0+0=0 \\ \left|-\frac{1}{3}\right| + \left|\frac{1}{6}\right| + \left|\frac{5}{18}\right| = \frac{7}{9} \\ \left|-\frac{1}{3}\right| + \left|\frac{2}{3}\right| + \left|\frac{1}{9}\right| = \frac{10}{9} \end{array}$$

La norma 1 viene dada por

$$||B_{G-S}||_1 = \mathrm{Max}\left(0, \frac{7}{9}, \frac{10}{9}\right) = \frac{10}{9} > 1$$

Como $||B_{G-S}||_1 < 1$ es condición suficiente, no necesaria, de convergencia, no podemos asegurar ni que converge ni que no converge.

Estudio de convergencia

Calculemos la norma 1 de esta matriz.

$$B_{G-S}^T = \begin{pmatrix} 0 & 0 & 0 \\ -\dfrac{1}{3} & \dfrac{1}{6} & \dfrac{5}{18} \\ -\dfrac{1}{3} & \dfrac{2}{3} & \dfrac{1}{9} \end{pmatrix} \quad \begin{array}{l} 0+0+0=0 \\ \left|-\dfrac{1}{3}\right| + \left|\dfrac{1}{6}\right| + \left|\dfrac{5}{18}\right| = \dfrac{7}{9} \\ \left|-\dfrac{1}{3}\right| + \left|\dfrac{2}{3}\right| + \left|\dfrac{1}{9}\right| = \dfrac{10}{9} \end{array}$$

La norma 1 viene dada por

$$||B_{G-S}||_1 = \mathrm{Max}\left(0, \frac{7}{9}, \frac{10}{9}\right) = \frac{10}{9} > 1$$

Como $||B_{G-S}||_1 < 1$ es condición suficiente, no necesaria, de convergencia, no podemos asegurar ni que converge ni que no converge.

Estudio de convergencia

Calculemos la norma infinito.

$$B_{G-S} = \begin{pmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{6} & \frac{2}{3} \\ 0 & \frac{5}{18} & \frac{1}{9} \end{pmatrix} \quad \begin{array}{l} 0 + \left|-\frac{1}{3}\right| + \left|-\frac{1}{3}\right| = \frac{2}{3} \\ 0 + \left|\frac{1}{6}\right| + \left|\frac{2}{3}\right| = \frac{5}{6} \\ 0 + \left|\frac{5}{18}\right| + \left|\frac{1}{9}\right| = \frac{7}{18} \end{array}$$

La norma infinito es

$$||B_{G-S}||_\infty = \mathrm{Max}\left(\frac{2}{3}, \frac{5}{6}, \frac{7}{18}\right) = \frac{5}{6} < 1$$

Como $||B_{G-S}||_\infty < 1$ podemos asegurar que converge.

Estudio de convergencia

Calculemos la norma infinito.

$$B_{G-S} = \begin{pmatrix} 0 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{6} & \frac{2}{3} \\ 0 & \frac{5}{18} & \frac{1}{9} \end{pmatrix} \quad \begin{array}{l} 0 + \left|-\frac{1}{3}\right| + \left|-\frac{1}{3}\right| = \frac{2}{3} \\ 0 + \left|\frac{1}{6}\right| + \left|\frac{2}{3}\right| = \frac{5}{6} \\ 0 + \left|\frac{5}{18}\right| + \left|\frac{1}{9}\right| = \frac{7}{18} \end{array}$$

La norma infinito es

$$||B_{G-S}||_\infty = \mathrm{Max}\left(\frac{2}{3}, \frac{5}{6}, \frac{7}{18}\right) = \frac{5}{6} < 1$$

Como $||B_{G-S}||_\infty < 1$ podemos asegurar que converge.

Estudio de convergencia

Además, si estudiamos los autovalores de la matriz B_{G-S}

$$\begin{vmatrix} 0-\lambda & -\frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{1}{6}-\lambda & \frac{2}{3} \\ 0 & \frac{5}{18} & \frac{1}{9}-\lambda \end{vmatrix} = 0$$

$$-\lambda\left(\frac{1}{6}-\lambda\right)\left(\frac{1}{9}-\lambda\right)-\frac{2}{3}\frac{5}{18}(-\lambda)=\frac{1}{6}\lambda+\frac{5}{18}\lambda^2-\lambda^3=(\lambda)\left(\frac{1}{6}+\frac{5}{18}\lambda-\lambda^2\right)=0$$

Y los autovalores son

$$\lambda_1 = 0, \quad \lambda_2 = 0.57, \quad \lambda_3 = 0.29$$

que son todos menores que 1 en valor absoluto.

Estudio de convergencia

Además, si estudiamos los autovalores de la matriz B_{G-S}

$$\begin{vmatrix} 0-\lambda & -\dfrac{1}{3} & -\dfrac{1}{3} \\ 0 & \dfrac{1}{6}-\lambda & \dfrac{2}{3} \\ 0 & \dfrac{5}{18} & \dfrac{1}{9}-\lambda \end{vmatrix} = 0$$

$$-\lambda\left(\frac{1}{6}-\lambda\right)\left(\frac{1}{9}-\lambda\right)-\frac{2}{3}\frac{5}{18}(-\lambda)=\frac{1}{6}\lambda+\frac{5}{18}\lambda^2-\lambda^3=(\lambda)\left(\frac{1}{6}+\frac{5}{18}\lambda-\lambda^2\right)=0$$

Y los autovalores son

$$\lambda_1 = 0, \quad \lambda_2 = 0.57, \quad \lambda_3 = 0.29$$

que son todos menores que 1 en valor absoluto.

Recordamos: para factorizar una matriz...

Matriz diagonalizable

Una *matriz cuadrada* A se dice que **es diagonalizable** si puede descomponerse de la forma

$$A = PDP^{-1}$$

donde

- P es una *matriz invertible* cuyos vectores columna son vectores propios de A.
- D es una *matriz diagonal* formada por los valores propios de A.

Inconvenientes

- A tiene que ser cuadrada.
- P tiene que ser invertible.
- Por lo tanto no todas las matrices se pueden factorizar de esta manera.

Recordamos: para factorizar una matriz...

Matriz diagonalizable

Una *matriz cuadrada* A se dice que **es diagonalizable** si puede descomponerse de la forma

$$A = PDP^{-1}$$

donde

- P es una *matriz invertible* cuyos vectores columna son vectores propios de A.
- D es una *matriz diagonal* formada por los valores propios de A.

Inconvenientes

- A tiene que ser cuadrada.
- P tiene que ser invertible.
- Por lo tanto no todas las matrices se pueden factorizar de esta manera.

Pero...

Matriz real simétrica

Cualquier *matriz cuadrada simétrica con coeficientes reales es diagonalizable*

$$A = PDP^T$$

y

- P está formada por una base ortonormal de vectores propios y y los vectores filas de P^{-1} son los vectores columnas de P. Por lo tanto $P^{-1} = P^T$
- D los valores propios son reales.

$MM^T \in \mathcal{M}_{m\times m}(\mathbb{R})$ $M^TM \in \mathcal{M}_{n\times n}(\mathbb{R})$ **Son matrices cuadradas simétricas con coeficientes reales** para cualquier matriz rectangular $M \in \mathcal{M}_{m\times n}(\mathbb{R})$

Pero...

Matriz real simétrica

Cualquier *matriz cuadrada simétrica con coeficientes reales es diagonalizable*

$$A = PDP^T$$

y

- P está formada por una base ortonormal de vectores propios y y los vectores filas de P^{-1} son los vectores columnas de P. Por lo tanto $P^{-1} = P^T$
- D los valores propios son reales.

$MM^T \in \mathcal{M}_{m\times m}(\mathbb{R})$ $M^T M \in \mathcal{M}_{n\times n}(\mathbb{R})$ **Son matrices cuadradas simétricas con coeficientes reales** para cualquier matriz rectangular $M \in \mathcal{M}_{m\times n}(\mathbb{R})$

Matriz real simétrica

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Diagonalizamos

Factorizamos $A = PDP^{-1}$ pero $P^{-1} \neq P^T$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Matriz real simétrica

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Diagonalizamos

Factorizamos $A = PDP^{-1}$ pero $P^{-1} \neq P^T$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Las columnas de P son vectores ortogonales

$$
\begin{aligned}
v_1 \cdot v_2 &= (1,2,1)\cdot(-1,0,1) = -1+0+1 = 0\\
v_1 \cdot v_3 &= (1,2,1)\cdot(1,-1,1) = 1-2+1 = 0\\
v_3 \cdot v_2 &= (1,-1,1)\cdot(-1,0,1) = -1+0+1 = 0
\end{aligned}
$$

Si normalizamos los vectores

$$
v_1 = \frac{v_1}{\|v_1\|} = \frac{(1,2,1)}{\sqrt{1^2+2^2+1^2}} = \left(\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, \frac{1}{\sqrt{6}}\right) \qquad v_2 = \frac{v_2}{\|v_2\|} \qquad v_3 = \frac{v_3}{\|v_3\|}
$$

Las columnas de P son vectores ortogonales

$$
\begin{aligned}
v_1 \cdot v_2 &= (1,2,1) \cdot (-1,0,1) = -1+0+1 = 0 \\
v_1 \cdot v_3 &= (1,2,1) \cdot (1,-1,1) = 1-2+1 = 0 \\
v_3 \cdot v_2 &= (1,-1,1) \cdot (-1,0,1) = -1+0+1 = 0
\end{aligned}
$$

Si normalizamos los vectores

$$
v_1 = \frac{v_1}{\|v_1\|} = \frac{(1,2,1)}{\sqrt{1^2+2^2+1^2}} = \left(\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, \frac{1}{\sqrt{6}}\right) \qquad v_2 = \frac{v_2}{\|v_2\|} \qquad v_3 = \frac{v_3}{\|v_3\|}
$$

Base normalizada Ahora sí $A = PDP^T$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

Descomposición en valores singulares (SVD)

Factorización SVD

Cualquier matriz rectangular $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ puede ser descompuesta de forma única

$$A = U\Sigma V^T$$

- $U \in \mathcal{M}_{m\times m}(\mathbb{R})$ sus columnas son los autovectores de AA^T

$$AA^T = U\Sigma V^T \left(U\Sigma V^T\right)^T = U\Sigma V^T V \Sigma^T U^T = U(\Sigma\Sigma^T)U^T$$

- $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ sus columnas son los autovectores de A^TA

$$A^TA = \left(U\Sigma V^T\right)^T U\Sigma V^T = V\Sigma^T U^T U \Sigma V^T = V(\Sigma^T\Sigma)V^T$$

- $\Sigma \in \mathcal{M}_{m\times n}(\mathbb{R})$ los elementos de su diagonal son no negativos y se llaman Valores Singulares y son las raíces de los autovalores de U y V

Descomposición en valores singulares (SVD)

Factorización SVD

Cualquier matriz rectangular $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ puede ser descompuesta de forma única

$$A = U\Sigma V^T$$

- $U \in \mathcal{M}_{m\times m}(\mathbb{R})$ sus columnas son los autovectores de AA^T

$$AA^T = U\Sigma V^T \left(U\Sigma V^T\right)^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$$

- $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ sus columnas son los autovectores de $A^T A$

$$A^T A = \left(U\Sigma V^T\right)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T\Sigma)V^T$$

- $\Sigma \in \mathcal{M}_{m\times n}(\mathbb{R})$ los elementos de su diagonal son no negativos y se llaman Valores Singulares y son las raíces de los autovalores de U y V

Descomposición en valores singulares (SVD)

Factorización SVD

Cualquier matriz rectangular $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ puede ser descompuesta de forma única

$$A = U\Sigma V^T$$

- $U \in \mathcal{M}_{m\times m}(\mathbb{R})$ sus columnas son los autovectores de AA^T

$$AA^T = U\Sigma V^T \left(U\Sigma V^T\right)^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$$

- $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ sus columnas son los autovectores de $A^T A$

$$A^T A = \left(U\Sigma V^T\right)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T\Sigma)V^T$$

- $\Sigma \in \mathcal{M}_{m\times n}(\mathbb{R})$ los elementos de su diagonal son no negativos y se llaman Valores Singulares y son las raíces de los autovalores de U y V

Descomposición en valores singulares (SVD)

Factorización SVD

Cualquier matriz rectangular $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ puede ser descompuesta de forma única

$$A = U\Sigma V^T$$

- $U \in \mathcal{M}_{m\times m}(\mathbb{R})$ sus columnas son los autovectores de AA^T

$$AA^T = U\Sigma V^T \left(U\Sigma V^T\right)^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$$

- $V \in \mathcal{M}_{n\times n}(\mathbb{R})$ sus columnas son los autovectores de $A^T A$

$$A^T A = \left(U\Sigma V^T\right)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T\Sigma)V^T$$

- $\Sigma \in \mathcal{M}_{m\times n}(\mathbb{R})$ los elementos de su diagonal son no negativos y se llaman Valores Singulares y son las raíces de los autovalores de U y V

Descomposición

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = U\Sigma V^T$$

Cálculo de U

$$AA^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = U(\Sigma\Sigma^T)U^T$$

Cálculo de V

$$A^TA = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = V(\Sigma^T\Sigma)V^T$$

Cálculo de U

$$AA^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = U(\Sigma\Sigma^T)U^T$$

Cálculo de V

$$A^T A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = V(\Sigma^T\Sigma)V^T$$

Interpretación geométrica

Ejemplo

$$A = U\Sigma V^T = \begin{pmatrix} -\cos(\phi) & \sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

con $\phi = \frac{\pi}{6}$, $\theta = \frac{\pi}{4}$, $\sigma_1 = 2$, $\sigma_2 = \frac{1}{2}$

A es una transformación lineal

$A = U\Sigma V^T$ está compuesta de:

- Un giro V^T (conserva longitudes y ángulos)
- Un escalamiento a los largo de los ejes Σ
- Otro giro U

Interpretación geométrica

Ejemplo

$$A = U\Sigma V^T = \begin{pmatrix} -\cos(\phi) & \sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} -\cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

con $\phi = \frac{\pi}{6}$, $\theta = \frac{\pi}{4}$, $\sigma_1 = 2$, $\sigma_2 = \frac{1}{2}$

A es una transformación lineal

$A = U\Sigma V^T$ está compuesta de:

- Un giro V^T (conserva longitudes y ángulos)
- Un escalamiento a los largo de los ejes Σ
- Otro giro U

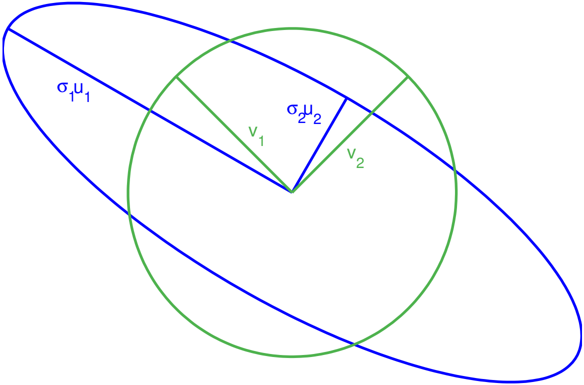
Interpretación geométrica

Interpretación del dibujo

- El círculo verde es el círculo unidad en el plano
- La elipse azul es la imagen de este círculo tras la transformación por A
- v_1 y v_2 son las columnas de V (base ortogonal) y están rotadas 45° respecto a los ejes
- u_1 y u_2 son las columans de U (base ortogonal) y están rotadas 30° respecto a los ejes
- A transforma

$$v_1 \to \sigma_1 u_1$$
$$v_2 \to \sigma_2 u_2$$

Interpretación geométrica



Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...

Aplicaciones SVD

SVD provee de valores significativos para una matriz, como los autovalores pero computacionalmente más interesantes. Por ello es útil casi en cualquier disciplina que use matrices como

- Análisis de Imágenes
- Estadística
- Problemas inversos
- Predicción del tiempo meteorológico
- Reconocimiento de formas (Pattern recognition)
- Indexación Semántica Latente (Latent Semantic Indexing)
  Wikipedia: es un método de indexación y recuperación que utiliza un método numérico llamado descomposición en valores singulares (SVD por sus siglas en inglés) para identificar patrones en las relaciones entre los términos contenidos en una colección de textos no estructurados.
- Análisis de la expresión de los genes
- Identificación de proteinas
- ...