

Métodos iterativos

de resolución de Sistemas Ecuaciones Algebraicas Lineales

Sistemas de ecuaciones lineales

- La resolución de un Sistemas Ecuaciones Algebraicas Lineales (SEAL):

$$\mathbf{Ax} = \mathbf{b}$$

donde

- \mathbf{x} es un vector conteniendo n incógnitas;
- \mathbf{A} una matriz cuadrada de $n \times n$ con los coeficientes del sistema;
- \mathbf{b} un vector con los términos independientes,

puede realizarse por distintos métodos. Estos suelen clasificarse en: Métodos Directos y Métodos Iterativos.

- Los Métodos *Iterativos* están basados en construir una secuencia de soluciones aproximadas $\{\mathbf{x}^{(k)}\}$ ($k = 1, 2, \dots$), que -a medida que el contador de iteraciones k aumente- se aproxime a la solución exacta.

- Esta secuencia se construye usando *fórmulas de recurrencia*. Hay varias fórmulas de recurrencia posibles y varias maneras de obtenerlas.
- Por ejemplo, una fórmulas de recurrencia puede ser:

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

- Hace falta una estimación inicial $\mathbf{x}^{(0)}$ para comenzar el proceso. Luego dando valores $k = 1, 2, \dots$ al contador de iteraciones se obtiene la secuencia buscada.
- ¿Cómo puede obtenerse una fórmula de recurrencia? Se puede hacer de distintas maneras. Por ejemplo, supóngase una matriz \mathbf{M} para la cual sea fácil resolver un sistema $\mathbf{M}\mathbf{x} = \mathbf{d}$. Entonces el sistema original $\mathbf{A}\mathbf{x} = \mathbf{b}$ puede escribirse:

$$(\mathbf{M} - \mathbf{M} + \mathbf{A})\mathbf{x} = \mathbf{b}$$

- Que puede escribirse:

$$\mathbf{M}\mathbf{x} = \mathbf{N}\mathbf{x} + \mathbf{b}$$

con $\mathbf{N} = \mathbf{M} - \mathbf{A}$.

- Esta fórmula puede usarse para recurrencia

$$\mathbf{M}\mathbf{x}^{(k)} = \mathbf{N}\mathbf{x}^{(k-1)} + \mathbf{b}$$

- En cada iteración debe efectuarse un producto ($\mathbf{N}\mathbf{x}$) y una resolución de sistema ($\mathbf{M}\mathbf{x} = \mathbf{d}$)

- Esta fórmula de recurrencia puede ser escrita:

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

- donde se ha definido

$$\mathbf{T} = (\mathbf{I} - \mathbf{M}^{-1}\mathbf{A})$$

$$\mathbf{c} = \mathbf{M}^{-1}\mathbf{b}$$

La matriz \mathbf{T} se denomina *Matriz de iteración*.

- La solución exacta \mathbf{x} verifica esa ecuación:

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

$$\mathbf{x} = \mathbf{T}\mathbf{x} + \mathbf{c}$$

- Restando, puede obtenerse el vector error en la iteración k :

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x} = \mathbf{T}(\mathbf{x}^{(k-1)} - \mathbf{x})$$

- La norma del error:

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| \leq \|\mathbf{T}\| \|\mathbf{x}^{(k-1)} - \mathbf{x}\|$$

- Si $\|\mathbf{T}\| < 1$ la secuencia $\{\mathbf{x}^{(k)}\}$ converge a \mathbf{x} .

- Hay muchos métodos iterativos para resolver sistemas de ecuaciones.
- Hay métodos *clásicos* como:
 - Método de Richardson (no lo veremos en este curso)
 - Método de Jacobi
 - Método de Gauss-Seidel
 - Método SOR

y técnicas de reducción del error para ellos.

- Por otro lado hay otros basados en técnicas de optimización, como los métodos de iteración por subespacios, entre los cuales tenemos:
 - Método del descenso más rápido
 - Método del gradiente conjugado
 - Método GMRES
 - Método del gradiente biconjugado
- A continuación veremos algunos de ellos.

- La matriz \mathbf{A} puede descomponerse de manera aditiva:

$$\mathbf{A} = \mathbf{L}_s + \mathbf{D}_s + \mathbf{U}_s$$

donde

- \mathbf{D}_s matriz diagonal con la diagonal de \mathbf{A} ($d_{ii} = a_{ii}$);
- \mathbf{L}_s matriz triangular inferior con los términos de \mathbf{A} excluyendo la diagonal ($l_{ij} = a_{ij}$ para $i > j$);
- \mathbf{U}_s matriz triangular superior con los términos de \mathbf{A} excluyendo la diagonal ($u_{ij} = a_{ij}$ para $i < j$).

- La iteración de Jacobi puede escribirse:

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

con

$$\mathbf{T}_j = -\mathbf{D}_s^{-1}(\mathbf{L}_s + \mathbf{U}_s)$$

- En la iteración k , la componente i de \mathbf{x} se calcula:

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

Método de Jacobi

Algoritmo del Método de Jacobi

Dados: $n, \mathbf{A}, \mathbf{b}, \mathbf{x}^0, Tol, Kmax$

```
for  $k = 1, \dots, Kmax$  do comienzo iteraciones
  for  $i = 1, \dots, n$  do
    
$$y_i \leftarrow \left[ b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^n a_{ij}x_j \right] / a_{ii}$$

  end for i
   $\epsilon \leftarrow ||y - x||$ 
  for  $i = 1, n$  do
     $x_i \leftarrow y_i$ 
  end for i
  if  $(\epsilon < Tol)$  stop
end for k
```

Convergencia del Método de Jacobi

- Un método iterativo converge si $\|\mathbf{T}\| < 1$
- En la iteración de Jacobi: $\mathbf{T}_j = -\mathbf{D}_s^{-1}(\mathbf{L}_s + \mathbf{U}_s) = (\mathbf{I} - \mathbf{D}_s^{-1}\mathbf{A})$
- Los elementos diagonales de $(\mathbf{D}_s^{-1}\mathbf{A})$ valen 1
- Una norma para \mathbf{T}_j :

$$\|\mathbf{T}_j\| = \max_{1 \leq i \leq n} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}/a_{ii}|$$

- Si una matriz es diagonalmente dominante, entonces el método de Jacobi converge para cualquier vector inicial.

Definición: Una matriz se dice diagonalmente dominante si $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$

Ejemplo Método de Jacobi

Ejemplo:

- Considérese el siguiente problema:

$$\begin{aligned}2x_1 - x_2 &= 1 \\ -x_1 + 2x_2 &= 1\end{aligned}$$

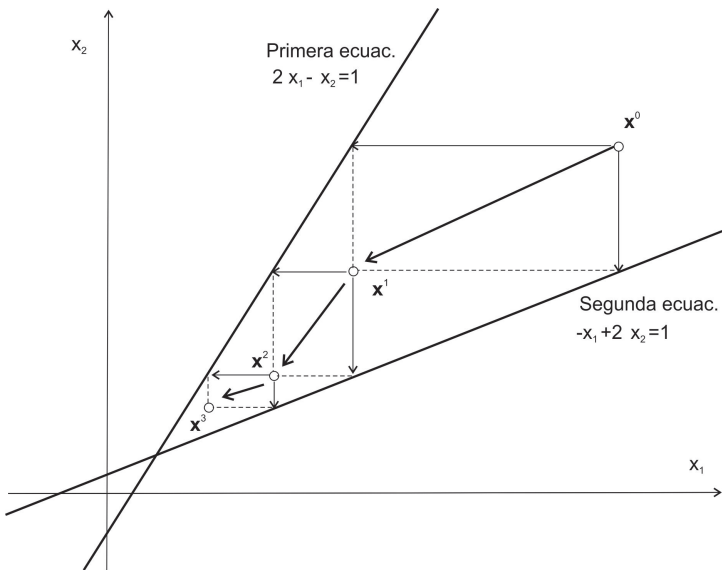
La solución es $x_1 = 1$, $x_2 = 1$ y el proceso iterativo está ilustrado en la figura siguiente. De la primera ecuación se obtiene la abscisa y de la segunda la ordenada del nuevo punto.

- Si, en cambio, se intercambian las filas:

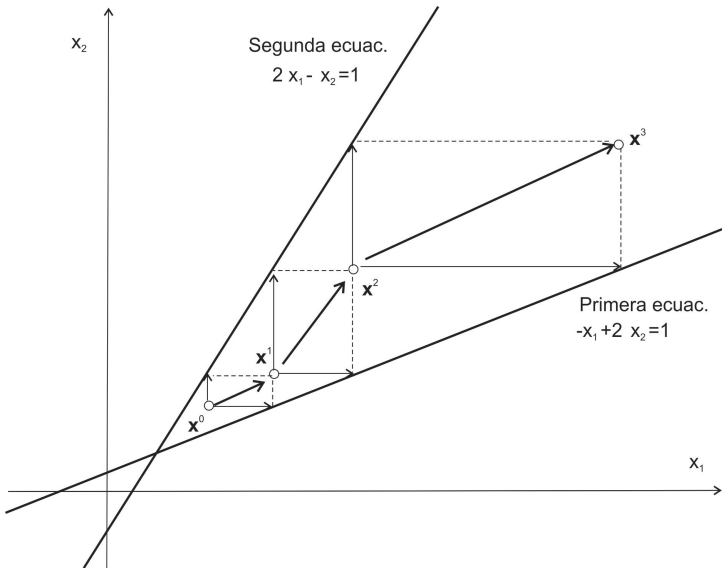
$$\begin{aligned}-x_1 + 2x_2 &= 1 \\ 2x_1 - x_2 &= 1\end{aligned}$$

el método de Jacobi no converge, como se ilustra en la figura.

Ejemplo Método de Jacobi



Ejemplo Método de Jacobi



Método de Gauss-Seidel

- Dada \mathbf{A} descompuesta en forma aditiva:

$$\mathbf{A} = \mathbf{L}_s + \mathbf{D}_s + \mathbf{U}_s$$

- La fórmula de iteración Gauss-Seidel:

$$\mathbf{x}^{(k)} = \mathbf{T}_{gs}\mathbf{x}^{(k-1)} + \mathbf{c}$$

con

$$\mathbf{T}_{gs} = -(\mathbf{D}_s + \mathbf{L}_s)^{-1}\mathbf{U}_s$$

- En la iteración k , la componente i de \mathbf{x} se calcula:

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

Método de Gauss-Seidel

Algoritmo del Método de Gauss-Seidel

Dados: $n, \mathbf{A}, \mathbf{b}, \mathbf{x}^0, Tol, Kmax$

```
for  $k = 1, \dots, Kmax$  do comienzo iteraciones
  for  $i = 1, \dots, n$  do
    
$$y_i \leftarrow \left[ b_i - \sum_{j=1}^{i-1} a_{ij}y_j - \sum_{j=i+1}^n a_{ij}x_j \right] / a_{ii}$$

  end for i
   $\epsilon \leftarrow ||y - x||$ 
  for  $i = 1, n$  do
     $x_i \leftarrow y_i$ 
  end for i
  if  $(\epsilon < Tol)$  stop
end for k
```


Convergencia del Método de Gauss-Seidel

- Si una matriz es *diagonalmente dominante*, entonces el método de Gauss-Seidel converge.
- Si una matriz es *simetrica, definida positiva*, entonces el método de Gauss-Seidel converge.

Obs:

Una matriz A se dice definida positiva si

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$$

Tambien una matriz es definida positiva si y solo si todos sus autovalores son positivos.

- Se define como *radio espectral* de una matriz \mathbf{A} :

$$\rho(\mathbf{A}) = \max\{|\lambda| \mid \det(\mathbf{A} - \lambda\mathbf{I}) = 0\}$$

- λ son los autovalores de \mathbf{A} , soluciones de la ecuación:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

Una matriz de $n \times n$ tiene n autovalores (algunos pueden ser repetidos; los autovalores pueden ser reales o complejos)

- $\rho(\mathbf{A})$ es el módulo del mayor autovalor; o sea el radio del menor círculo en el plano complejo, que contiene a todos los autovalores.

- Se puede demostrar que:
 - $[\rho(\mathbf{A}^T \mathbf{A})]^{1/2} = \|\mathbf{A}\|_2$
 - $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$ para toda norma de \mathbf{A}
- En particular, si \mathbf{A} es simétrica:

$$\rho(\mathbf{A}) = \|\mathbf{A}\|_2$$

- Para que la sucesión

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

converja a \mathbf{x} es necesario y suficiente que

$$\rho(\mathbf{T}) < 1$$

Ejemplo Método de Gauss-Seidel

- Considérese el mismo problema resuelto por el Método de Jacobi:

$$\begin{aligned}2x_1 - x_2 &= 1 \\ -x_1 + 2x_2 &= 1\end{aligned}$$

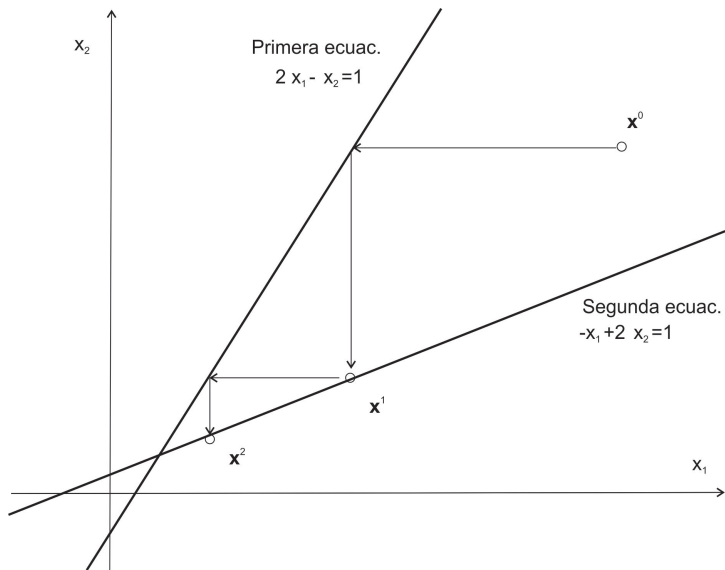
cuya solución es $x_1 = 1$, $x_2 = 1$. El proceso iterativo de Gauss Seidel está ilustrado en la figura siguiente.

- Si, en cambio, se intercambian las filas:

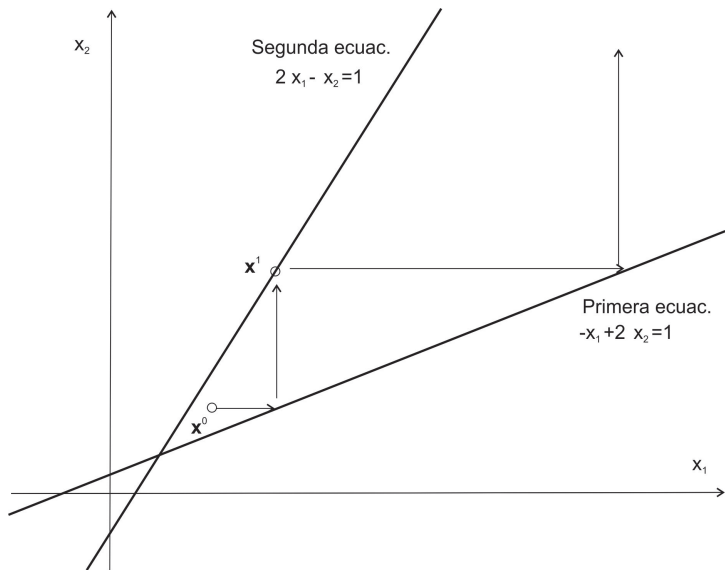
$$\begin{aligned}-x_1 + 2x_2 &= 1 \\ 2x_1 - x_2 &= 1\end{aligned}$$

el método de Gauss-Seidel tampoco converge, como se ilustra en la figura.

Ejemplo Método de Gauss-Seidel



Ejemplo Método de Gauss-Seidel



Método de Sobre-relajaciones Sucesivas (Successive Over Relaxation)

- Dada A descompuesta en forma aditiva:

$$\mathbf{A} = \mathbf{L}_s + \mathbf{D}_s + \mathbf{U}_s$$

- La fórmula de iteración SOR:

$$\mathbf{M}_\omega \mathbf{x}^{(k)} = \mathbf{N}_\omega \mathbf{x}^{(k-1)} + \omega \mathbf{b}$$

con

$$\mathbf{M}_\omega = \mathbf{D}_s + \omega \mathbf{L}_s$$

$$\mathbf{N}_\omega = (1 - \omega) \mathbf{D}_s - \omega \mathbf{U}_s$$

y donde ω es un parámetro propio del método.

- La matriz de iteración \mathbf{T}_ω es, entonces:

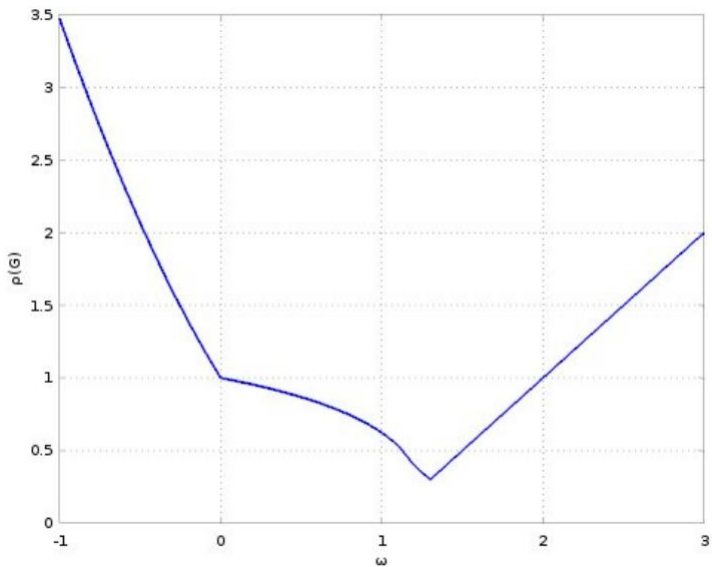
$$\mathbf{T}_\omega = (\mathbf{D}_s + \omega \mathbf{L}_s)^{-1}((1 - \omega) \mathbf{D}_s - \omega \mathbf{U}_s)$$

- En la iteración k , la componente i de \mathbf{x} se calcula:

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

- Si $\omega = 1$ SOR \rightarrow Gauss-Seidel.
- Si $\rho(\mathbf{T}_\omega) < 1$ SOR converge.
- Si $a_{ii} \neq 0, \forall i$, entonces $\rho(\mathbf{T}_\omega) \geq |\omega - 1|$, luego debería ser $0 < \omega < 2$
- Si $0 < \omega < 1 \rightarrow$ método sub-relajación
- Si $1 < \omega < 2 \rightarrow$ método sobre-relajación
- Para una matriz se puede hallar un ω óptimo, que minimice el radio espectral

Método SOR



Método SOR

Algoritmo del Método SOR

Dados: $n, \mathbf{A}, \mathbf{b}, \mathbf{x}^0, \omega, Tol, Kmax$

```
for  $k = 1, \dots, Kmax$  do comienzo iteraciones
  for  $i = 1, \dots, n$  do
    
$$y_i \leftarrow (1 - \omega)x_i + \omega \left[ b_i - \sum_{j=1}^{i-1} a_{ij}y_j - \sum_{j=i+1}^n a_{ij}x_j \right] / a_{ii}$$

  end for i
   $\epsilon \leftarrow ||y - x||$ 
  for  $i = 1, n$  do
     $x_i \leftarrow y_i$ 
  end for i
  if  $(\epsilon < Tol)$  stop
end for k
```

Métodos de iteración por subespacios

- Los métodos iterativos que hemos visto (Jacobi, Gauss-Seidel, SOR) son muy sencillos y fáciles de aplicar. Sin embargo su convergencia en general es lenta y no son muy utilizados en la práctica.
- Los métodos directos, resultan caros para resolver grandes sistemas de ecuaciones.
- En estos casos se utilizan otros métodos iterativos que podemos designar como métodos de iteración por subespacios, o métodos de Krylov.
- Uno de ellos, que describiremos a continuación, es el Método del Gradiente Conjugado. Este método, que puede ser mirado como un método directo, ya que en un número finito de pasos conduce a la solución exacta, en la práctica se utiliza como un método iterativo ya que se realiza una cantidad -relativamente- pequeña de iteraciones hasta lograr la convergencia deseada.

Métodos de iteración por subespacios

- Considérese la solución del sistema $\mathbf{Ax} = \mathbf{b}$ con \mathbf{A} simétrica, positiva definida.

El Metodo del Gradiente Conjugado permite resolver sistemas con matrices simétricas y definida positivas. Para el caso de matrices no simétricas existen otros metodos similares (Generalized Minimal Residual Method -GMRES-, Biconjugate gradient method -BiCG-) que no seran tratados aquí.

- Se define una *forma cuadrática* asociada a este sistema:

$$\phi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Ax} - \mathbf{b}^T\mathbf{x}$$

Puede observarse que ϕ es un escalar.

- Se puede demostrar que el vector \mathbf{x} que minimiza ϕ es -a su vez- solución de $\mathbf{Ax} = \mathbf{b}$. → ver Apéndice A
- En esto se basan estos métodos de iteración por subespacios. En vez que buscar el vector solución de ese sistema, se busca el que minimiza la forma cuadrática.

Métodos de iteración por subespacios

- Hay métodos que buscan un mínimo de ϕ con iteraciones

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}$$

- Dado $\mathbf{x}^{(k)}$ en la iteración k , los algoritmos buscan una dirección de descenso $\mathbf{p}^{(k)}$ y un paso $\alpha^{(k)}$ a lo largo de esa dirección, para hallar la nueva iteración.
- En la iteración k se define el *residuo* como :

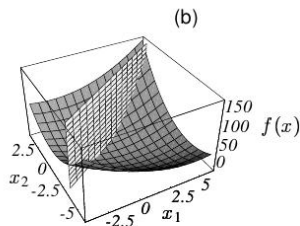
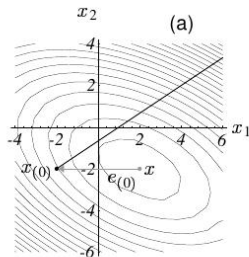
$$\mathbf{r}^{(k)} = (\mathbf{b} - \mathbf{A}\mathbf{x}^{(k)})$$

Está claro que se busca que el residuo tienda a cero.

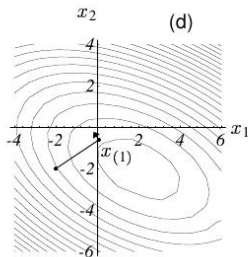
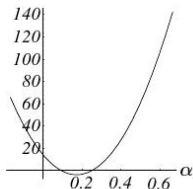
- Veremos entonces como calcular: 1) La dirección de búsqueda (o de descenso) \mathbf{p} ; y 2) El paso en esa dirección (α)

Métodos de iteración por subespacios

Por ejemplo, para vectores en \mathbb{R}^2 :



$f(x_{(i)} + \alpha r_{(i)})$ (c)



Métodos de iteración por subespacios

Métodos para calcular la dirección de búsqueda \mathbf{p} :

- Mencionaremos aquí dos métodos: el método del descenso más rápido y el método del gradiente conjugado.
- *Método del descenso más rápido (steepest descent):*

$$\mathbf{p}^{(k)} = -\left. \frac{\partial \phi}{\partial \mathbf{x}} \right|_{\mathbf{x}^{(k)}} = \mathbf{r}^{(k)}$$

El gradiente de la función ϕ en un punto, marca la dirección en que ésta crece con mayor pendiente. Por tanto el signo negativo indica la dirección en que desciende más rápidamente.

Puede verse que esa dirección resulta ser igual al vector residuo. Este método no se usa mucho pues su convergencia puede ser lenta.

Métodos de iteración por subespacios

- *Método del gradiente conjugado* :

En la iteración k se calcula la nueva iteración:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(0)} + \alpha^{(0)} \mathbf{p}^{(0)} + \alpha^{(1)} \mathbf{p}^{(1)} + \cdots + \alpha^{(k)} \mathbf{p}^{(k)}$$

La dirección de descenso $\mathbf{p}^{(k)}$ se busca tal que sea *conjugada* a todas las anteriores.

Se dice que dos direcciones son *conjugadas* o *A-ortogonales* si se verifica:

$$\mathbf{p}^{(i)T} \mathbf{A} \mathbf{p}^{(j)} = \begin{cases} 0 & \text{if } i \neq j \\ > 0 & \text{if } i = j \end{cases} \quad (3)$$

Es decir, ese doble producto para dos direcciones diferentes es cero. (Si se trata de la misma dirección es un número positivo, ya que la matriz es definida positiva).

Método del Gradiente Conjugado

- La nueva dirección se calcula con la fórmula:

$$\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \gamma^{(k)} \mathbf{p}^{(k)}$$

- Si se introduce ésta en la condición de conjugadas:

$$(\mathbf{p}^{(k+1)})^T \mathbf{A} \mathbf{p}^{(k)} = 0$$

de allí puede despejarse $\gamma^{(k)}$ que resulta:

$$\gamma^{(k)} = -\frac{\mathbf{r}^{(k+1)T} \mathbf{A} \mathbf{p}^{(k)}}{\mathbf{p}^{(k)T} \mathbf{A} \mathbf{p}^{(k)}}$$

Método del Gradiente Conjugado

Cálculo del paso de avance (α)

- ϕ puede escribirse como función de α :

$$\phi(\mathbf{x}^{(k+1)}) = \phi(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}) = \tilde{\phi}(\alpha^{(k)})$$

- El paso $\alpha^{(k)}$ se calcula tal que minimice $\tilde{\phi}(\alpha^{(k)})$:

$$\frac{d\tilde{\phi}(\alpha^{(k)})}{d\alpha^{(k)}} = \mathbf{p}^{(k)T} \frac{\partial \phi}{\partial \mathbf{x}^{(k)}} = \mathbf{p}^{(k)T} [\mathbf{A}(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}) - \mathbf{b}] = 0$$

$$(\mathbf{p}^{(k)T} \mathbf{A} \mathbf{p}^{(k)}) \alpha^{(k)} = \mathbf{p}^{(k)T} (\mathbf{b} - \mathbf{A} \mathbf{x}^{(k)}) = \mathbf{p}^{(k)T} \mathbf{r}^{(k)}$$

siendo $\mathbf{r}^{(k)}$ el residuo en la iteración k . → ver Apéndice A

- De aquí el valor del paso se calcula:

$$\alpha^{(k)} = \frac{\mathbf{p}^{(k)T} \mathbf{r}^{(k)}}{\mathbf{p}^{(k)T} \mathbf{A} \mathbf{p}^{(k)}}$$

Método del Gradiente Conjugado

Resumiendo, entonces, en cada iteración del método del gradiente conjugado se calcula:

- $\alpha^{(k)} = \frac{\mathbf{r}^{(k)T} \mathbf{p}^{(k)}}{\mathbf{p}^{(k)T} \mathbf{A} \mathbf{p}^{(k)}}$
- $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}$
- $\mathbf{r}^{(k+1)} = \mathbf{b} - \mathbf{A} \mathbf{x}^{(k+1)}$ (en realidad se usa $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha^{(k)} \mathbf{A} \mathbf{p}^{(k)}$)
- $\gamma^{(k)} = -\frac{\mathbf{r}^{(k+1)T} \mathbf{A} \mathbf{p}^{(k)}}{\mathbf{p}^{(k)T} \mathbf{A} \mathbf{p}^{(k)}}$ (en realidad se usa $\gamma^{(k)} = \frac{(\mathbf{r}^{(k)}, \mathbf{r}^{(k)})}{(\mathbf{r}^{(k-1)}, \mathbf{r}^{(k-1)})}$)
→ ver Apéndice A
- $\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} + \gamma^{(k)} \mathbf{p}^{(k)}$

Método del Gradiente Conjugado

- Se puede demostrar:

- $\text{span}\{\mathbf{p}^{(0)}, \mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \dots, \mathbf{p}^{(k)}\} = \text{span}\{\mathbf{r}^{(0)}, \mathbf{r}^{(1)}, \mathbf{r}^{(2)}, \dots, \mathbf{r}^{(k)}\} = \text{span}\{\mathbf{r}^{(0)}, \mathbf{A}\mathbf{r}^{(0)}, \mathbf{A}^2\mathbf{r}^{(0)}, \dots, \mathbf{A}^k\mathbf{r}^{(0)}\}$

- $(\mathbf{r}^{(i)}, \mathbf{r}^{(j)}) = 0 \quad \text{for } i \neq j$

- El espacio definido por

$$K_{(k+1)}(\mathbf{r}^{(0)}, \mathbf{A}) = \text{span}\{\mathbf{r}^{(0)}, \mathbf{A}\mathbf{r}^{(0)}, \mathbf{A}^2\mathbf{r}^{(0)}, \dots, \mathbf{A}^k\mathbf{r}^{(0)}\}$$

se llama *espacio de Krylov*

Método del Gradiente Conjugado

- El Método del Gradiente Conjugado busca un $\mathbf{x}^{(k)}$ tal que minimice el error (en norma \mathbf{A}) en el espacio:

$$\mathbf{x}^{(0)} + K_k(\mathbf{r}^{(0)}, \mathbf{A})$$

o sea

$$((\mathbf{x}^{(k)} - \mathbf{x}), \mathbf{A}(\mathbf{x}^{(k)} - \mathbf{x})) = (\mathbf{x}^{(k)} - \mathbf{x})^T \mathbf{A}(\mathbf{x}^{(k)} - \mathbf{x})$$

es mínimo para todo $\mathbf{x}^{(k)} \in \mathbf{x}^{(0)} + K_k(\mathbf{r}^{(0)}, \mathbf{A})$

- Puede demostrarse que si la matriz \mathbf{A} tiene rango r , el metodo del gradiente conjugado converge en a lo sumo r iteraciones.

Método del Gradiente Conjugado

A. Inicialización

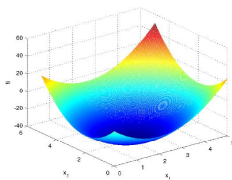
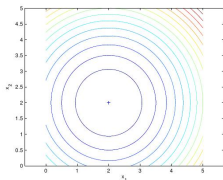
- A.1 \mathbf{x} estimación inicial
- A.2 $\mathbf{r} = \mathbf{b} - \mathbf{Ax}$ prod.matriz-vector+suma vectores
- A.3 $\mathbf{p} = \mathbf{r}$
- A.4 $\rho = \mathbf{r}^T \mathbf{p}$ producto interno
- A.5 $\rho_0 = \rho$
- A.6 $k = 1$

B. Iteraciones: while $k < Kmax$ do

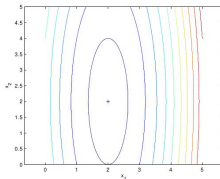
- B.1 Test convergencia: si $\rho < Tol$ ρ_0 fin
- B.2 $\mathbf{a} = \mathbf{Ap}$ prod.matriz-vector
- B.3 $m = \mathbf{p}^T \mathbf{a}$ producto interno
- B.4 $\alpha = \frac{\rho}{m}$
- B.5 $\mathbf{x} = \mathbf{x} + \alpha \mathbf{p}$ SAXPY
- B.6 $\mathbf{r} = \mathbf{r} - \alpha \mathbf{a}$ SAXPY
- B.7 $\rho_{old} = \rho$
- B.8 $\rho = \mathbf{r}^T \mathbf{r}$ producto interno
- B.9 $\gamma = \frac{\rho}{\rho_{old}}$
- B.10 $\mathbf{p} = \mathbf{r} + \gamma \mathbf{p}$ SAXPY
- B.11 $k = k + 1$, go to B.1

Método del Gradiente Conjugado

- Una matriz esférica (matriz identidad multiplicada por una constante) en \mathbb{R}^2 tiene una forma cuadrática como la de la figura. Su número de condición es 1. En una iteración se debería hallar el valor exacto.



- Una matriz con $\kappa(\mathbf{A}) = 10$ tiene una forma cuadrática como la de la figura:



Método del Gradiente Conjugado Precondicionado

- El método del gradiente conjugado es muy usado en cálculos científicos. Por ejemplo en mecánica computacional (elementos finitos).
- En problemas estructurales la matriz suele ser simétrica, positiva definida y rala (banda).
- En problemas de fluidos suele ser no simétrica. Pueden usarse métodos análogos (GMRES).
- El método de gradiente conjugado es eficiente si el número de condición es bajo (o la matriz está *bien condicionada*).
- El número de condición:

$$\kappa(\mathbf{A}) = \frac{|\lambda_{max}|}{|\lambda_{min}|}$$

donde λ_{max} y λ_{min} son los autovalores extremos de la matriz, puede ser muy alto.

Método del Gradiente Conjugado Precondicionado

- Hay varios factores que contribuyen a elevar el número de condición de la matriz. Por ejemplo, en elementos finitos:
 - granularidad de la malla:
cuanto mayor es la cantidad de elementos en la malla, mayor es el número de condición.
 - distorsión de la malla:
al aumentar la relación de aspecto, aumenta el número de condición.
*(por ejemplo para elementos finitos de placa,
con relaciones de aspecto 1:1 $\kappa(\mathbf{A}) = 14$;
con relac. de aspecto 1:10 $\kappa(\mathbf{A}) = 31000$ [Cook, 1974])*
 - grandes variaciones en los términos de la matriz:
*(por ej. una matriz en elem. finitos de:
estado plano de tensiones, puede tener $\kappa(\mathbf{A}) = 10^3 \sim 10^5$;
placa, puede tener $\kappa(\mathbf{A}) = 10^{12}$)*

Método del Gradiente Conjugado Precondicionado

- Para alto número de condición, la convergencia del método del gradiente conjugado es lenta.
- Para remediarlo se propone un *precondicionamiento*. O sea hallar una matriz \mathbf{P} tal que

$$\kappa(\mathbf{P}^{-1}\mathbf{A}) \simeq 1$$

- Si $\mathbf{P} \simeq \mathbf{A}$ entonces $\kappa(\mathbf{P}^{-1}\mathbf{A}) \simeq 1$
- El problema a resolver es

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{x} = \mathbf{P}^{-1}\mathbf{b}$$

- Si \mathbf{P} se escribe

$$\mathbf{P} = (\mathbf{S}\mathbf{S}^T)^{-1}$$

Método del Gradiente Conjugado Precondicionado

- siendo S matriz cuadrada, regular, simétrica, positiva definida

$$SS^T \mathbf{Ax} = SS^T \mathbf{b}$$

- El problema se puede escribir

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

con

$$\hat{\mathbf{A}} = S^T \mathbf{A} S$$

$$\hat{\mathbf{x}} = S^{-1} \mathbf{x}$$

$$\hat{\mathbf{b}} = S^T \mathbf{b}$$

- $\mathbf{P}^{-1} \mathbf{A}$ y $\hat{\mathbf{A}}$ tienen los mismo autovalores. Luego

$$\kappa(\mathbf{P}^{-1} \mathbf{A}) = \kappa(\hat{\mathbf{A}})$$

Algunos preconditionadores:

- Jacobi:

El preconditionador más simple es

$$\mathbf{P} = \mathbf{D}_s$$

con \mathbf{D}_s diagonal de \mathbf{A} . Es efectivo si la matriz es diagonal dominante.

- Block-Jacobi:

Una matriz con bloques de submatrices sobre la diagonal de \mathbf{A}

- SOR:

$$\mathbf{P} = (\mathbf{D}_s + \omega \mathbf{L}_s) \mathbf{D}_s^{-1} (\mathbf{D}_s + \omega \mathbf{U}_s)$$

- Cholesky incompleta (o LU incompleta):

Si A es simétrica, positiva definida, puede descomponerse en :

$$A = C^T C$$

Se puede escribir:

$$A = C^{*T} C^* + R$$

con C^* una descomposición de Cholesky restringida a la estructura rala de A .

Tomando

$$P = C^{*T} C^*$$

se espera que el número de condición sea pequeño.

Método del Gradiente Conjugado Precondicionado

A. Inicialización

A.1	\mathbf{x}	estimac.inicial
A.2	$\mathbf{r} = \mathbf{b} - \mathbf{Ax}$	prod.matriz-vector+suma vect.
A.3	resuelve $\mathbf{Pz} = \mathbf{r}$	solucion SEAL
A.4	$\rho = \mathbf{r}^T \mathbf{z}$	prod.interno
A.5	$\rho_0 = \rho$	
A.6	$\mathbf{p} = \mathbf{z}$	
A.7	$k = 1$	

B. Iteraciones: while $k < Kmax$ do

B.1	Test converg.: si $\rho < Tol$ ρ_0 fin	
B.2	$\mathbf{a} = \mathbf{Ap}$	prod.matriz-vector
B.3	$\mathbf{m} = \mathbf{p}^T \mathbf{a}$	prod.interno
B.4	$\alpha = \frac{\rho}{m}$	
B.5	$\mathbf{x} = \mathbf{x} + \alpha \mathbf{p}$	SAXPY
B.6	$\mathbf{r} = \mathbf{r} - \alpha \mathbf{a}$	SAXPY
B.7	solve $\mathbf{Pz} = \mathbf{r}$	soluc.SEAL
B.8	$\rho_{old} = \rho$	
B.9	$\rho = \mathbf{r}^T \mathbf{z}$	prod.interno
B.10	$\gamma = \frac{\rho}{\rho_{old}}$	
B.11	$\mathbf{p} = \mathbf{z} + \gamma \mathbf{p}$	SAXPY
B.12	$k = k + 1$, go to B.1	

Ejemplos

- El SEAL con

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 24 \\ 30 \\ -24 \end{bmatrix}$$

(ref. Burden y Faires), tiene solución: $\mathbf{x}^T = [3 \quad 4 \quad -5]$

$$\|\mathbf{A}\|_{\infty} = 8, \kappa(\mathbf{A}) = 8.55 \quad (\lambda_{\max} = 7.1623 \text{ y } \lambda_{\min} = 0.8377)$$

- Con una tolerancia de 10^{-7} sobre el residuo, se requirieron:
 - Con Gauss- Seidel: 34 iteraciones
 - Con SOR ($\omega = 1.25$): 14 iteraciones
 - Con Gradiente Conjugado: 3 iteraciones

- El SEAL con

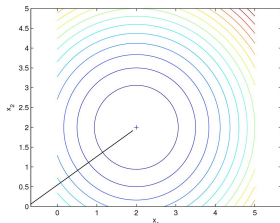
$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.1 & 1 & 1 & 0 \\ 0.1 & 4 & -1 & 1 & -1 \\ 1 & -1 & 60 & 0 & -2 \\ 1 & 1 & 0 & 8 & 4 \\ 0 & -1 & -2 & 4 & 700 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

$$\kappa(\mathbf{A}) = 12264$$

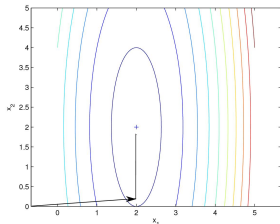
- Con una tolerancia de 10^{-2} sobre el residuo, la cantidad de iteraciones y el error absoluto en norma infinito:
 - Con Jacobi: 49 iteraciones, $\|\mathbf{e}\|_{\infty} = 0,00305834$
 - Con Gauss- Seidel: 15 iteraciones, $\|\mathbf{e}\|_{\infty} = 0,02445559$
 - Con SOR ($\omega = 1.25$): 7 iteraciones, $\|\mathbf{e}\|_{\infty} = 0,00818607$
 - Con Gradiente Conjugado: 5 iteraciones, $\|\mathbf{e}\|_{\infty} = 0,00629785$
 - Con Grad. Conj. Precond. (Jacobi): 4 iterac., $\|\mathbf{e}\|_{\infty} = 0,00009312$

Ejemplos

- Para una matriz de 2×2 esférica el método de GC converge en una iteración



- Para una matriz de 2×2 con $\kappa(\mathbf{A}) = 10$ el método de GC converge en dos iteraciones.



Ejemplos

- Sistema de 10 ecuaciones $\kappa(\mathbf{A}) = 9.3025$, , $\|\mathbf{A}\|_{\infty} = 70$,
 $\|\mathbf{A}\|_2 = 60.4085$
- Con una tolerancia de 10^{-3} sobre $\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$, la cantidad de iteraciones y el error absoluto en norma infinito:
 - Con Jacobi: 7 iteraciones, $\|\mathbf{e}\|_{\infty} = 0,0271$
 - Con Gauss- Seidel: 4 iteraciones, $\|\mathbf{e}\|_{\infty} = 0.1690$
 - Con SOR ($\omega = 1.25$): 6 iteraciones, $\|\mathbf{e}\|_{\infty} = 0.0672$
 - Con Gradiente Conjugado: 4 iteraciones, $\|\mathbf{e}\|_{\infty} = 0,4540$
 - Con Grad. Conj. Precond. (Jacobi): 2 iterac., $\|\mathbf{e}\|_{\infty} = 0.3851$

Criterios para detener las iteraciones

- En los métodos iterativos vistos, se construye una secuencia $\{\mathbf{x}^{(k)}\}$ que en el límite tiende a la solución exacta \mathbf{x} .
- Para la iteración k se define:
 - El error: $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$
 - El residuo de la ecuación: $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$
 - El incremento en esa iteración: $\delta^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$,
- Se ha visto que el error $\mathbf{e}^{(k)}$ y el residuo $\mathbf{r}^{(k)}$, en la iteración k , están vinculados por la ecuación:

$$\mathbf{A}\mathbf{e}^{(k)} = \mathbf{r}^{(k)}$$

Por tanto, cuando el error tiende a cero, el residuo también lo hace.

Criterios para detener las iteraciones

- El error absoluto puede escribirse:

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$$

$$\mathbf{e}^{(k+1)} = \mathbf{x}^{(k+1)} - \mathbf{x} = \mathbf{x}^{(k)} - \mathbf{x} + \boldsymbol{\delta}^{(k)} = \mathbf{e}^{(k)} + \boldsymbol{\delta}^{(k)}$$

- De allí:

$$\mathbf{e}^{(k)} = \mathbf{e}^{(k+1)} - \boldsymbol{\delta}^{(k)}$$

$$\|\mathbf{e}^{(k)}\| = \|\mathbf{e}^{(k+1)} - \boldsymbol{\delta}^{(k)}\| \leq \|\mathbf{e}^{(k+1)}\| + \|\boldsymbol{\delta}^{(k)}\| \leq \|\mathbf{T}\| \|\mathbf{e}^{(k)}\| + \|\boldsymbol{\delta}^{(k)}\|$$

La primera desigualdad surge por la Desigualdad Triangular, que debe cumplir toda norma. La segunda por el método iterativo convergente.

Criterios para detener las iteraciones

- Para matrices simétricas y positiva definidas:

$$\|\mathbf{e}^{(k)}\| \leq \rho(\mathbf{T}) \|\mathbf{e}^k\| + \|\delta^{(k)}\|$$

$$\|\mathbf{e}^{(k)}\| \leq \frac{1}{1 - \rho(\mathbf{T})} \|\delta^{(k)}\|$$

- La diferencia entre dos iteraciones $\delta^{(k)}$ puede servir de cota del error $\mathbf{e}^{(k)}$, siempre que el radio espectral sea pequeño ($\ll 1$).
- Si el radio espectral es cercano a 1 la diferencia entre dos iteraciones $\delta^{(k)}$ no es una buena medida del error.

Criterios para detener las iteraciones

- En vista de esto hay varias formas de decidir cuando detener el proceso iterativo

① $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq Tol_1$

② $\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|}{\|\mathbf{x}^{(k)}\|} \leq Tol_2$

③ $\|\mathbf{r}^{(k)}\| \leq Tol_3 \|\mathbf{b}\|$

- La primera usa la diferencia entre dos iteraciones. La tolerancia Tol_1 depende del significado físico de la variable \mathbf{x} (de sus unidades).
- La segunda usa un valor relativo del incremento. Aquí Tol_2 es independiente de las unidades (se puede especificar 0,01, 10^{-4} , etc.).
- En el tercer criterio la tolerancia Tol_3 también es independiente de las unidades ya que se refiere a la norma del vector de términos independientes (que tiene las mismas unidades que el residuo).

En este capítulo hemos visto:

- Qué son los métodos iterativos de resolución de SEAL
- Métodos clásicos: Jacobi, Gauss-Seidel, SOR

$$\mathbf{Ax} = \mathbf{b} \underset{\text{sist.equiv.}}{\longrightarrow} \mathbf{x} = \mathbf{Tx} + \mathbf{c} \underset{\text{f.iterativa}}{\longrightarrow} \mathbf{x}^{(k+1)} = \mathbf{Tx}^{(k)} + \mathbf{c}$$

- Condiciones de convergencia de cada uno de ellos
- Métodos de iteración por subespacios: Gradiente conjugado

$$\mathbf{Ax} = \mathbf{b} \underset{\text{prob.asoc.}}{\longrightarrow} \min_{\mathbf{x}} \phi(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{Ax} - \mathbf{b}^T \mathbf{x} \underset{\text{f.iterativa}}{\longrightarrow} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}$$

- El método del gradiente conjugado preconditionado
- Algunos ejemplos comparativos
- Criterios para detener las interacciones