

Sistemas de Ecuaciones Lineales VI

- **Métodos de descenso:** Método del gradiente conjugado. Precondicionamiento.

Problema de minimización asociado a un sistema

- Algunos de los métodos iterativos más eficientes para resolver un sistema de ecuaciones lineales se basan en resolver un problema de minimización equivalente.
- Si $\mathbf{A} \in \mathbb{R}^{n \times n}$ es **simétrica y definida positiva**, $\mathbf{b} \in \mathbb{R}^n$ y $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^t \mathbf{y}$ denota el producto escalar usual en \mathbb{R}^n , entonces la función cuadrática

$$\begin{aligned} J : \quad \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{2} \langle \mathbf{A}x, x \rangle - \langle \mathbf{b}, x \rangle \end{aligned}$$

alcanza un único mínimo en el vector $\mathbf{x} \in \mathbb{R}^n$ solución del sistema $\mathbf{A}x = \mathbf{b}$.

Problema de minimización asociado a un sistema (cont.)

- En efecto,

$$\nabla J(\mathbf{x}) = \frac{1}{2} (\mathbf{A}\mathbf{x} + \mathbf{A}^t\mathbf{x}) - \mathbf{b} = \mathbf{A}\mathbf{x} - \mathbf{b}, \quad \text{si } \mathbf{A} \text{ es simétrica.}$$

Entonces J tiene un punto crítico en \mathbf{x} si y sólo si

$$\nabla J(\mathbf{x}) = \mathbf{0} \quad \Longleftrightarrow \quad \mathbf{A}\mathbf{x} = \mathbf{b}.$$

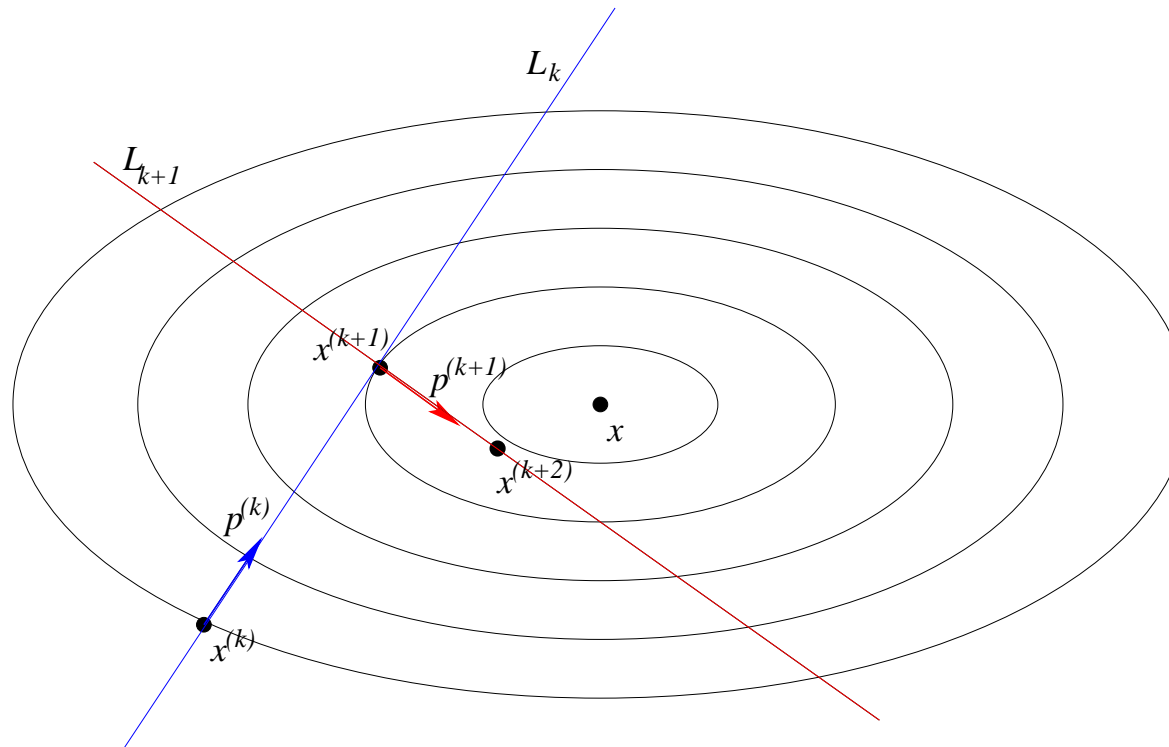
Además,

$$HJ(\mathbf{x}) := \left(\frac{\partial^2 J}{\partial x_i \partial x_j}(\mathbf{x}) \right) = (a_{ij}) = \mathbf{A}$$

y, por lo tanto, si \mathbf{A} es definida positiva, el punto crítico de J es un mínimo.

Métodos de descenso

- En los **métodos de descenso** se parte de un punto $\mathbf{x}^{(0)} \in \mathbb{R}^n$, y, en cada paso $k = 0, 1, 2, \dots$ se determina un nuevo punto $\mathbf{x}^{(k+1)} \in \mathbb{R}^n$ tal que $J(\mathbf{x}^{(k+1)}) < J(\mathbf{x}^{(k)})$ del siguiente modo:
 - se escoge una dirección $\mathbf{p}^{(k)}$ de descenso de J ,
 - se considera la recta L_k que pasa por el punto $\mathbf{x}^{(k)}$ con dirección $\mathbf{p}^{(k)}$,
 - se escoge el punto $\mathbf{x}^{(k+1)} \in L_k$ donde J alcanza su mínimo sobre L_k .



Métodos de descenso (cont.)

- Como $L_k = \{ \mathbf{x}^{(k)} + \alpha \mathbf{p}^{(k)} : \alpha \in \mathbb{R} \}$, entonces

$$\begin{aligned} J(\mathbf{x}^{(k)} + \alpha \mathbf{p}^{(k)}) \\ = \left[\frac{1}{2} \langle \mathbf{A} \mathbf{x}^{(k)}, \mathbf{x}^{(k)} \rangle - \langle \mathbf{b}, \mathbf{x}^{(k)} \rangle \right] + \langle \mathbf{A} \mathbf{x}^{(k)} - \mathbf{b}, \mathbf{p}^{(k)} \rangle \alpha + \frac{1}{2} \langle \mathbf{A} \mathbf{p}^{(k)}, \mathbf{p}^{(k)} \rangle \alpha^2 \end{aligned}$$

Por lo tanto,

$$\frac{dJ}{d\alpha} \left(\mathbf{x}^{(k)} + \alpha \mathbf{p}^{(k)} \right) = 0 \quad \Longleftrightarrow \quad \alpha = \alpha_k := \frac{\langle \mathbf{r}^{(k)}, \mathbf{p}^{(k)} \rangle}{\langle \mathbf{A} \mathbf{p}^{(k)}, \mathbf{p}^{(k)} \rangle},$$

donde $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A} \mathbf{x}^{(k)}$ es el residuo de $\mathbf{x}^{(k)}$.

Luego

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

Método del máximo descenso

- Los distintos métodos de descenso se distinguen por la manera de escoger la dirección de descenso $\mathbf{p}^{(k)}$.
- La elección más simple es escoger la dirección de máximo descenso de J :

$$\mathbf{p}^{(k)} = -\nabla J(\mathbf{x}^{(k)}) = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)} = \mathbf{r}^{(k)}.$$

- Esta elección conduce al **método del máximo descenso** o del **gradiente**.

- **Algoritmo:**

Dado el vector inicial $\mathbf{x}^{(0)}$,

$$\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)},$$

para $k = 0, 1, 2, \dots$

$$\alpha_k = \frac{\langle \mathbf{r}^{(k)}, \mathbf{r}^{(k)} \rangle}{\langle \mathbf{A}\mathbf{r}^{(k)}, \mathbf{r}^{(k)} \rangle},$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{r}^{(k)},$$

$$\mathbf{r}^{(k+1)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{A}\mathbf{r}^{(k)},$$

hasta que se satisfaga un criterio de detención.

Criterio de detención

- Un criterio usual en los métodos de descenso es detener el proceso cuando

$$\frac{\|\mathbf{r}^{(k)}\|}{\|\mathbf{b}\|} \leq \text{tol}.$$

- Sin embargo, eso no garantiza que el error satisfaga $\|\mathbf{e}^{(k)}\| = \|\mathbf{x} - \mathbf{x}^{(k)}\| \leq \text{tol} !!$

Observación. Si $\mathbf{A} \in \mathbb{R}^{n \times n}$ es simétrica y definida positiva, entonces la función

$$\begin{aligned} \langle \cdot, \cdot \rangle_{\mathbf{A}} : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) &\longmapsto \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} := \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

es el **producto interior definido por \mathbf{A}** y la correspondiente **norma inducida** es:

$$\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{A}}}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Convergencia del método del máximo descenso

- **Teorema.** (Convergencia) Sean $A \in \mathbb{R}^{n \times n}$ simétrica y definida positiva y $b \in \mathbb{R}^n$. Para cualquier $x^{(0)} \in \mathbb{R}^n$, la sucesión $\{x^{(k)}\}$ generada por el método del máximo descenso converge a la solución del sistema $Ax = b$.

Además, los errores $e^{(k)} = x - x^{(k)}$ satisfacen

$$\|e^{(k+1)}\|_A \leq \frac{\kappa - 1}{\kappa + 1} \|e^{(k)}\|_A,$$

donde $\kappa = \text{cond}_2(A)$.

- **Observación.** El factor de reducción del error en cada paso es

$$\frac{\kappa - 1}{\kappa + 1} = 1 - \frac{2}{\kappa + 1} < 1,$$

lo que asegura la convergencia.

Sin embargo, si la matriz es mal condicionada,

$$\kappa \gg 1 \implies 1 - \frac{2}{\kappa + 1} \approx 1$$

y el método converge muy lentamente.

Direcciones conjugadas

- Una elección de las direcciones de descenso $\mathbf{p}^{(k)}$, que genera un método más eficiente, se obtiene al usar vectores de dirección ortogonales respecto del producto interior definido por \mathbf{A} :

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} := \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

- Estas direcciones de descenso $\mathbf{p}^{(k)}$ satisfacen

$$\langle \mathbf{A}\mathbf{p}^{(k)}, \mathbf{p}^{(j)} \rangle = 0, \quad \forall j \neq k$$

y por ello se dice que son **direcciones conjugadas** o **\mathbf{A} -ortogonales**.

- El uso de estos vectores de descenso conduce al **método del gradiente conjugado**.

Partiendo de $\mathbf{x}^{(0)}$ arbitrario, y tomando $\mathbf{p}^{(0)} = \mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}$, los sucesivos vectores \mathbf{A} -ortogonales $\mathbf{p}^{(k)}$ se construyen como parte del método del gradiente conjugado y se presentan en el algoritmo que sigue.

Método del gradiente conjugado

- Algoritmo:

Dado el vector inicial $\mathbf{x}^{(0)}$,

$$\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)},$$

$$\mathbf{p}^{(0)} = \mathbf{r}^{(0)},$$

para $k = 0, 1, 2, \dots$

$$\alpha_k = \frac{\langle \mathbf{r}^{(k)}, \mathbf{p}^{(k)} \rangle}{\langle \mathbf{A}\mathbf{p}^{(k)}, \mathbf{p}^{(k)} \rangle},$$

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

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{A}\mathbf{p}^{(k)},$$

$$\beta_{k+1} = \frac{\langle \mathbf{r}^{(k+1)}, \mathbf{r}^{(k+1)} \rangle}{\langle \mathbf{r}^{(k)}, \mathbf{r}^{(k)} \rangle},$$

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

hasta que se satisfaga un criterio de detención.

Convergencia del método del gradiente conjugado

- **Teorema.** (Convergencia) Sean $\mathbf{A} \in \mathbb{R}^{n \times n}$ simétrica y definida positiva y $\mathbf{b} \in \mathbb{R}^n$.

Para cualquier $\mathbf{x}^{(0)} \in \mathbb{R}^n$, la sucesión $\{\mathbf{x}^{(k)}\}$ generada por el método del gradiente conjugado converge a la solución del sistema $\mathbf{Ax} = \mathbf{b}$ en a lo más n pasos.

Además, los errores $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$ satisfacen

$$\|\mathbf{e}^{(k)}\|_{\mathbf{A}} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \|\mathbf{e}^{(0)}\|_{\mathbf{A}}, \quad \text{donde } \kappa = \text{cond}_2(\mathbf{A}).$$

- **Observación.** Debido a los errores de redondeo, en las condiciones del teorema anterior, igual pueden ser necesarios más de n pasos del método.

Convergencia del método del gradiente conjugado (cont.)

- Ahora el **factor de reducción del error** en cada paso es

$$\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} = 1 - \frac{2}{\sqrt{\kappa} + 1} < 1.$$

- Si $\kappa > 1$, este método converge más velozmente que el del máximo descenso, pues

$$\frac{2}{\sqrt{\kappa} + 1} > \frac{2}{\kappa + 1}.$$

Por ejemplo, si $\kappa = 100$, entonces

- Máximo descenso: $\frac{\kappa - 1}{\kappa + 1} \approx 0.98$ (2 % de reducción del error por paso)
- Gradiente conjugado: $\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \approx 0.82$ (18 % de reducción del error por paso).

Sin embargo, **si la matriz es muy mal condicionada**, el método aún converge lentamente.

Precondicionamiento

- Si la matriz es mal condicionada, una estrategia usual para resolver un sistema $Ax = b$, se basa en encontrar una matriz P tal que $P^{-1}A$ sea bien condicionada, para luego resolver el sistema equivalente

$$(P^{-1}A)x = P^{-1}b.$$

- Sin embargo, si se quiere utilizar un método como el del gradiente conjugado, debe notarse que aunque A sea simétrica y definida positiva, en general $P^{-1}A$ no lo es!
- En el caso en que P es simétrica y definida positiva, $P = HH^t$, con H invertible (Cholesky). En tal caso, $P^{-1} = H^{-t}H^{-1}$ y

$$Ax = b \iff \hat{A}\hat{x} = \hat{b},$$

$$\text{con } \hat{A} = H^{-1}AH^{-t}, \quad \hat{x} = H^t x \quad \text{y} \quad \hat{b} = H^{-1}b.$$

- Si A es simétrica y definida positiva, la matriz $\hat{A} = H^{-1}AH^{-t}$ también lo es, por lo que pueden aplicarse métodos como el del gradiente conjugado a $\hat{A}\hat{x} = \hat{b}$. Además,

$$\sigma(P^{-1}A) = \sigma(\hat{A}) \implies \text{cond}_2(\hat{A}) = \frac{\max\{\sigma(P^{-1}A)\}}{\min\{\sigma(P^{-1}A)\}} = \frac{\lambda_{\max}(P^{-1}A)}{\lambda_{\min}(P^{-1}A)}.$$

Método del gradiente conjugado preconditionado

- **Algoritmo:**

Dado el vector inicial $\mathbf{x}^{(0)}$,

$$\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)},$$

resolver $\mathbf{P}\mathbf{z}^{(0)} = \mathbf{r}^{(0)}$,

$$\mathbf{p}^{(0)} = \mathbf{z}^{(0)},$$

para $k = 0, 1, 2, \dots$

$$\alpha_k = \frac{\langle \mathbf{z}^{(k)}, \mathbf{r}^{(k)} \rangle}{\langle \mathbf{A}\mathbf{p}^{(k)}, \mathbf{p}^{(k)} \rangle},$$

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

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{A}\mathbf{p}^{(k)},$$

resolver $\mathbf{P}\mathbf{z}^{(k+1)} = \mathbf{r}^{(k+1)}$,

$$\beta_{k+1} = \frac{\langle \mathbf{z}^{(k+1)}, \mathbf{r}^{(k+1)} \rangle}{\langle \mathbf{z}^{(k)}, \mathbf{r}^{(k)} \rangle},$$

$$\mathbf{p}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta_{k+1} \mathbf{p}^{(k)},$$

hasta que se satisfaga un criterio de detención.

Precondicionadores

- Para que el método del gradiente conjugado preconditionado sea eficiente, debe escogerse un preconditionador P tal que

1. P **simétrica y definida positiva**, de manera que $P = HH^t$ (Cholesky);
2. $P \approx A$, de manera que $P^{-1}A \approx I$, y así

$$\text{cond}_2(\hat{A}) = \frac{\lambda_{\max}(P^{-1}A)}{\lambda_{\min}(P^{-1}A)} \approx 1 \ll \text{cond}_2(A);$$

3. **resolver los sistemas $Pz = r$ sea poco costoso.**
- **Ejemplo. Precondicionador de Jacobi.** Consiste en tomar como P , la diagonal de A :

$$P = \text{diag}(A) = \text{diag}(a_{11}, \dots, a_{nn}).$$

Este preconditionador es efectivo cuando A es simétrica y definida positiva y con **elementos diagonales muy distintos**.

Además resolver sistemas $Pz = r$ con P diagonal es muy poco costoso.

Descomposición incompleta de Cholesky

- **Ejemplo. Precondicionador de Cholesky incompleto.** Consiste en utilizar $P = HH^t$, donde $H = (h_{ij})$ se obtiene aplicando el algoritmo del método de Cholesky, pero sólo para las entradas h_{ij} tales que $a_{ij} \neq 0$. Para las restantes se toma $h_{ij} = 0$.
- **Algoritmo de la descomposición incompleta de Cholesky.**

Para $j = 1, \dots, n$

$$h_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} h_{jk}^2}$$

para $i = j + 1, \dots, n$

si $a_{ij} = 0$, entonces $h_{ij} = 0$,

$$\text{si no, } h_{ij} = \frac{1}{h_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} h_{ik} h_{jk} \right)$$

Descomposición incompleta de Cholesky (cont.)

- Si $h_{jj} \neq 0, j = 1, \dots, n$, entonces \mathbf{H} resulta no singular y $\mathbf{P} = \mathbf{H}\mathbf{H}^t$ puede usarse como preconditionante.
- Si \mathbf{A} no tuviera entradas nulas, la descomposición de Cholesky sería completa y, en tal caso, $\mathbf{P} = \mathbf{H}\mathbf{H}^t = \mathbf{A}$.

Para matrices dispersas, esto no es cierto, pero usualmente $\mathbf{P} \approx \mathbf{A}$.

- Resolver los sistemas $\mathbf{P}\mathbf{z} = \mathbf{r}$ es poco costoso pues ya se dispone de la factorización de Cholesky $\mathbf{P} = \mathbf{H}\mathbf{H}^t$ y basta resolver dos sistemas triangulares.

Además \mathbf{H} es igualmente dispersa que \mathbf{A} .

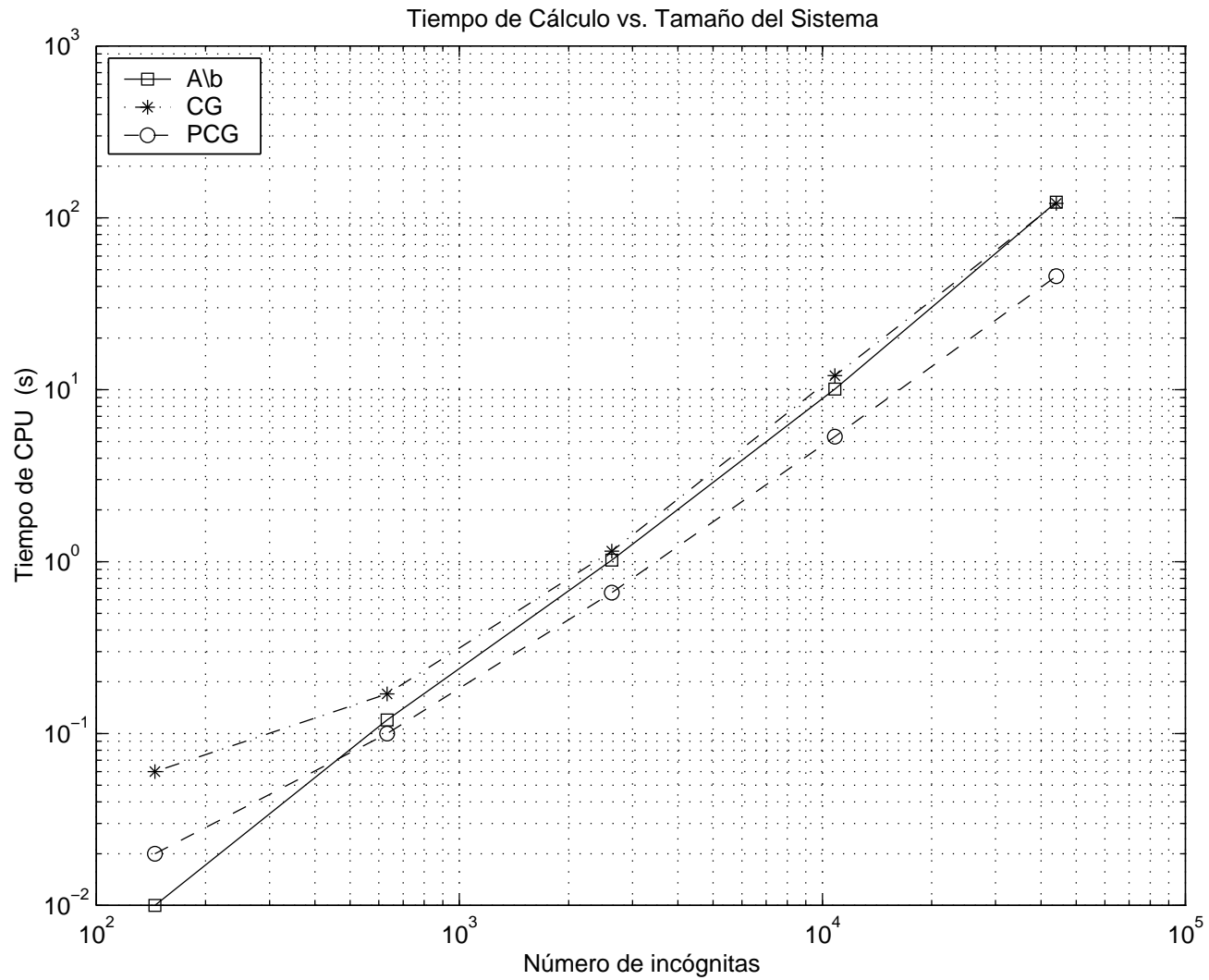
Comparación de costos computacionales

- Se considera como ejemplo un sistema con $\mathbf{A} \in \mathbb{R}^{n \times n}$ y $\mathbf{b} \in \mathbb{R}^n$ provenientes de un programa MATLAB para la resolución de un problema de EDPs que modela la deformación de una membrana bajo la acción de una fuerza (como se verá en el laboratorio correspondiente).
- Se utiliza una tolerancia $\text{tol} = 10^{-6}$.
- Se resuelve el sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ mediante el comando MATLAB **pcg** (método del gradiente conjugado preconditionado):
 1. sin preconditionador,
 2. con preconditionador \mathbf{P} obtenido mediante el comando MATLAB **cholinc** (Cholesky incompleto con relleno parcial), con el parámetro **droptol** $= 10^{-2}$.

Tabla I: Número de iteraciones para cada método

n	145	632	2629	10821	44071
MGC	32	66	125	261	500
MGCP	5	10	16	35	65

Comparación de costos operacionales



Comparación de costos de almacenamiento

