Gradiente Conjugado y Espacios de Krylov

M. Sepúlveda

1 de Junio, 2021

Motivación: métodos de descenso

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

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

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

Motivación: métodos de descenso

Problema de minimización asociado a un sistema (cont.)

- En efecto,

$$\nabla J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{A}\mathbf{x} + \mathbf{A}^T\mathbf{x}\right) - \mathbf{b} = \mathbf{A}\mathbf{x} - \mathbf{b}, \qquad \text{si } \mathbf{A} \text{ es } \textbf{simétrica}.$$

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,

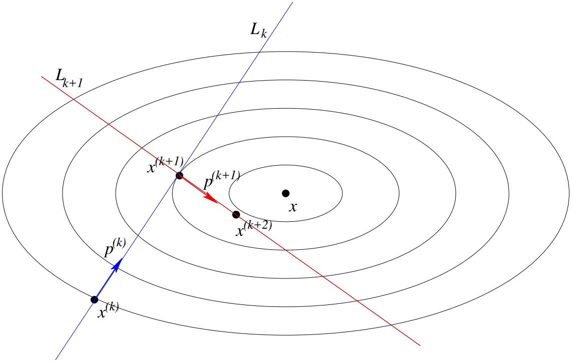
$$HJ(\mathbf{x}) := \left(\frac{\partial^2 J}{\partial x_i \partial x_j}(\mathbf{x})\right) = (a_{ij}) = \mathbf{A}$$

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

Motivación: métodos de descenso

Métodos de descenso

- En los **métodos de descenso** se parte de un punto $\mathbf{x}_0 \in \mathbb{R}^n$, y, en cada paso $k = 0, 1, 2, \ldots$ 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:
  1. se escoge una dirección $\boldsymbol{p}_k$ de descenso de J,
  2. se considera la recta L_k que pasa por el punto $\mathbf{x}_k$ con dirección $\boldsymbol{p}_k$,
  3. se escoge el punto $\mathbf{x}_{k+1} \in L_k$ donde J alcanza su mínimo sobre L_k.



Métodos de descenso (cont.)

- Como $L_k = \{\mathbf{x}_k + \alpha \boldsymbol{p}_k : \ \alpha \in \mathbb{R}\}$, entonces

$$J(\mathbf{x}_k + \alpha \boldsymbol{p}_k) = \left[\frac{1}{2}\mathbf{x}_k^T \mathbf{A}\mathbf{x}_k - \mathbf{b}^T \mathbf{x}_k\right] + (\mathbf{A}\mathbf{x}_k - \mathbf{b})^T \boldsymbol{p}_k \alpha + \frac{1}{2}\boldsymbol{p}_k^T \mathbf{A} \boldsymbol{p}_k \alpha^2$$

Por lo tanto,

$$\frac{dJ}{d\alpha}(\mathbf{x}_k + \alpha \boldsymbol{p}_k) = 0 \quad \Longleftrightarrow \quad \alpha = \alpha_k := \frac{\boldsymbol{r}_k^T \boldsymbol{p}_k}{\boldsymbol{p}_k^T \mathbf{A} \boldsymbol{p}_k},$$

donde $\boldsymbol{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k$ es el residuo de $\mathbf{x}_k$.
Luego

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \boldsymbol{p}_k.$$

Método del máximo descenso

- Los distintos métodos de descenso se distinguen por la manera de escoger la dirección de descenso $\boldsymbol{p}_k$.
- La elección más simple es escoger la dirección de máximo descenso de J:

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

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

> Dado el vector inicial $\mathbf{x}_0$,
> $\boldsymbol{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}_0$,
> para $k = 0, 1, 2, \ldots$
> $$\alpha_k = \frac{\boldsymbol{r}_k^T \boldsymbol{r}_k}{\boldsymbol{r}_k^T \mathbf{A} \boldsymbol{r}_k},$$
> $$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \boldsymbol{r}_k,$$
> $$\boldsymbol{r}_{k+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_{k+1} = \boldsymbol{r}_k - \alpha_k \mathbf{A} \boldsymbol{r}_k,$$
> hasta que se satisfaga un criterio de detención.

Criterio de detención

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

$$\frac{\|\boldsymbol{r}_k\|}{\|\mathbf{b}\|} \leq \mathtt{tol}.$$

- Sin embargo, eso no garantiza que el error satisfaga $\|\mathbf{e}_k\| = \|\mathbf{x} - \mathbf{x}_k\| \leq \mathtt{tol}!!$

Convergencia del método del máximo descenso

- **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 máximo descenso converge a la solución del sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$.
  Además, los errores $\mathbf{e}_k = \mathbf{x} - \mathbf{x}_k$ satisfacen

$$\|\mathbf{e}_{k+1}\|_2 \leqslant \frac{\kappa - 1}{\kappa + 1}\|\mathbf{e}_k\|_2,$$

  donde $\kappa = \operatorname{cond}_2(\mathbf{A})$ y $\|\mathbf{e}\|_{\mathbf{A}} = \sqrt{\mathbf{e}^T\mathbf{A}\mathbf{e}}$.
- **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.

Métodos basados en Espacios de Krylov

Sea $\mathbf{A}$ una matriz real nosingular de $n \times n$. Para la solución de un sistema de ecuaciones lineales

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

empezando con un vector $\mathbf{x}_0$, los Métodos basados en espacios de Krylov producen una cadena de vectores

$$\mathbf{x}_0 \to \mathbf{x}_1 \to \cdots \to \mathbf{x}_m,$$

la cual termina de manera aritméticamente exacta, con una solución exacta $\mathbf{x}_m = \bar{x} := \mathbf{A}^{-1}\mathbf{b}$ en a lo más n pasos, $m \leqslant n$. La característica especial de los métodos de espacio de Krylov es que la iteración $\mathbf{x}_k$ satisface para todo $k \geqslant 0$

$$\mathbf{x}_k \in \mathbf{x}_0 + K_k(\boldsymbol{r}_0, \mathbf{A})$$

donde

$$K_k(\boldsymbol{r}_0, \mathbf{A}) := \operatorname{span}\{\boldsymbol{r}_0, \mathbf{A}\boldsymbol{r}_0, \ldots, \mathbf{A}^{k-1}\boldsymbol{r}_0\} \qquad , k = 1, 2, \ldots$$

es el espacio de Krylov perteneciente a la matriz $\mathbf{A}$ y empezando por el vector $\boldsymbol{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$, dado por el residual de $\mathbf{x}_0$.

Método de Gradiente Conjugado (Idea General)

Un primero método de este tipo es el *método de Gradiente Conjugado* (gc), fue propuesto por Hestenes y Stiefel (1952) para sistemas con una matriz simétrica definida positiva. Estas matrices definen una norma $\|\mathbf{x}\|_{\mathbf{A}} := \sqrt{\mathbf{x}^T\mathbf{A}\mathbf{x}}$, y el método ($gc$) genera una sucesión $\mathbf{x}_k \in \mathbf{x}_0 + K_k(\boldsymbol{r}_0, \mathbf{A})$ con la propiedad de minimización:

$$\|\mathbf{x}_k - \bar{\mathbf{x}}\|_A = \min_{\mathbf{z}\in\mathbf{x}_0+K_k(\boldsymbol{r}_0,\mathbf{A})} \|\mathbf{z} - \bar{\mathbf{x}}\|_{\mathbf{A}}.$$

En este método, un rol importante lo juegan unos vectores $\mathbf{A}$-conjugados, $\boldsymbol{p}_k \in \mathbb{R}^n$, $k = 0, 1, \ldots$

$$\boldsymbol{p}_i^T\mathbf{A}\boldsymbol{p}_k = 0, \qquad \text{para } i \neq k$$

generadores del espacio de Krylov

$$K_k(\boldsymbol{r}_0, \mathbf{A}) = \operatorname{span}\{\boldsymbol{p}_0, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{k-1}\} \qquad k = 1, 2 \ldots$$

Método GMRES (Idea General)

El método del Residuo Mínimo Generalizado [Saad y Schulz (1986), Saad(1996)] es más caro, pero está definido para sistemas lineales más generales con una matriz nosingular no-simétrica. Genera vectores $\mathbf{x}_k \in \mathbf{x}_0 + K_k(\boldsymbol{r}_0, \mathbf{A})$ con un residuo mínimo $\mathbf{b} - \mathbf{A}\mathbf{x}_k$,

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}_k\|_2 = \min_{\mathbf{z} \in \mathbf{x}_0 + K_k(\boldsymbol{r}_0, \mathbf{A})} \|\mathbf{b} - \mathbf{A}\mathbf{z}\|_2$$

y usa, como herramienta principal, vectores ortonormales $\mathbf{v}_1, \mathbf{v}_2, \ldots$ que entregan una base ortonormal de los espacios de Krylov $K_k(\boldsymbol{r}_0, \mathbf{A})$ de dimensión k

$$K_k(\boldsymbol{r}_0, \mathbf{A}) = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \qquad k = 1, 2 \ldots$$

Gradiente Conjugado y Funcional Cuadrático

Sea $\mathbf{A}$ una matriz real nosingular de $n \times n$ simétrica definida positiva. Consideremos sistema de ecuaciones lineales

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

y la solución exacta $\bar{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{b}$. Consideremos la norma $\|\mathbf{x}\|_{\mathbf{A}} := \sqrt{\mathbf{x}^T\mathbf{A}\mathbf{x}}$, fuertemente relacionada con el funcional cuadrático

$$\begin{aligned} F(\mathbf{z}) &= \frac{1}{2}\left(\mathbf{b} - \mathbf{A}\mathbf{z}\right)^T \mathbf{A}^{-1}\left(\mathbf{b} - \mathbf{A}\mathbf{z}\right) \\ &= \frac{1}{2}\mathbf{z}^T\mathbf{A}\mathbf{z} - \mathbf{b}^T\mathbf{z} + \frac{1}{2}\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} \\ &= \frac{1}{2}(\mathbf{z} - \bar{\mathbf{x}})^T\mathbf{A}(\mathbf{z} - \bar{\mathbf{x}}) \\ &= \frac{1}{2}\|\mathbf{z} - \bar{\mathbf{x}}\|_{\mathbf{A}}^2, \end{aligned}$$

minimizado por $\bar{\mathbf{x}}$,

$$0 = F(\bar{\mathbf{x}}) = \min_{\mathbf{z}\in\mathbb{R}^n} F(\mathbf{z})$$

Método de Gradiente Conjugado v/s Método de Gradiente

Al considerar el método de máximo descenso, la sucesión $\mathbf{x}_0 \to \mathbf{x}_1 \to \cdots$ se obtiene a través de una minimización uni-dimensional de F en la dirección del gradiente:

$$F(\mathbf{x}_{k+1}) = \min_{d} F(\mathbf{x}_k + d\boldsymbol{r}_k), \qquad \text{con } \boldsymbol{r}_k = -\nabla F = \mathbf{b} - \mathbf{A}\boldsymbol{r}_k.$$

En cambio, en el caso del método de Gradiente Conjugado, en cada paso $\mathbf{x}_k \to \mathbf{x}_{k+1}$ se resuelve un problema de minimización $(k+1)$-dimensional:

$$F(\mathbf{x}_{k+1}) = \min_{d_0,\ldots,d_k} F(\mathbf{x}_k + d_0\boldsymbol{r}_0 + \cdots + d_k\boldsymbol{r}_k),$$

con $\boldsymbol{r}_i = \mathbf{b} - \mathbf{A}\mathbf{x}_i$ para $i = 0, \ldots, k$.

Algoritmo Gradiente Conjugado

- **Algoritmo:**

Dado el vector inicial $\mathbf{x}_0$,
$\boldsymbol{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}_0$,
$\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$,
para $k = 0, 1, 2, \dots$

$$\alpha_k = \frac{\boldsymbol{r}_k^T \boldsymbol{r}_k}{\boldsymbol{p}_k \mathbf{A} \boldsymbol{p}_k},$$
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \boldsymbol{p}_k,$$
$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k \mathbf{A} \boldsymbol{p}_k,$$
$$\beta_{k+1} = \frac{\boldsymbol{r}_{k+1}^T \boldsymbol{r}_{k+1}}{\boldsymbol{r}_k^T \boldsymbol{r}_k},$$
$$\boldsymbol{p}_{k+1} = \boldsymbol{r}_{k+1} + \beta_{k+1} \boldsymbol{p}_k,$$

hasta que se satisfaga un criterio de detención,
o bien hasta que $\boldsymbol{p}_{k+1} = 0$.

Teorema

Teorema

Sea $\mathbf{A}$ *una matriz real simétrica definida positiva, y* $\mathbf{b} \in \mathbb{R}^n$*. Entonces para cada vector inicial* $\mathbf{x}_0 \in \mathbb{R}^n$ *hay un entero nonegativo más pequeño* $m \leqslant n$*, tal que* $\boldsymbol{p}_m = 0$*. Los vectores* $\mathbf{x}_k$*,* $\boldsymbol{p}_k$*,* $\boldsymbol{r}_k$*,* $k = 0, \ldots, m$*, generados por el algoritmo de Gradiente Conjugado, satisfacen:*

1. $\mathbf{A}\mathbf{x}_m = \mathbf{b}$*. El método entrega la solución exacta en a lo más* n *pasos.*
2. $\boldsymbol{r}_j^T \boldsymbol{p}_i = 0,$ *para* $0 \leqslant i < j \leqslant m$*.*
3. $\boldsymbol{r}_i^T \boldsymbol{p}_i = \boldsymbol{r}_i^T \boldsymbol{r}_i,$ *para* $i \leqslant m$*.*
4. $\boldsymbol{p}_i^T \mathbf{A} \boldsymbol{p}_j = 0,$ *para* $0 \leqslant i < j \leqslant m$*,* $\quad \boldsymbol{p}_j^T \mathbf{A} \boldsymbol{p}_j > 0,$ *para* $j < m$*.*
5. $\boldsymbol{r}_i^T \boldsymbol{r}_j = 0,$ *para* $0 \leqslant i < j \leqslant m$*,* $\quad \boldsymbol{r}_j^T \boldsymbol{r}_j > 0,$ *para* $j < m$*.*
6. $\boldsymbol{r}_i = \mathbf{b} - \mathbf{A}\mathbf{x}_i,$ *para* $i \leqslant m$*.*

De este Teorema se deduce en particular que el método está bien definido mientras $\boldsymbol{r}_k^T \boldsymbol{r}_k > 0$, $\boldsymbol{p}_k^T \mathbf{A} \boldsymbol{p}_k > 0$, para $\boldsymbol{p}_k \neq 0$. Además los vectores $\boldsymbol{p}_k$ son $\mathbf{A}$-conjugados, lo cual justifica el nombre del método.

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 $\mathbf{A}\mathbf{x} = \mathbf{b}$ en **a lo más n pasos**.
  Además, los errores $\mathbf{e}_k = \mathbf{x} - \mathbf{x}_k$ satisfacen

$$\|\mathbf{e}_k\|_{\mathbf{A}} \leqslant 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|\mathbf{e}_0\|_{\mathbf{A}}, \qquad \text{donde } \kappa = \mathrm{cond}_2(\mathbf{A}).$$

- **Observación.**
  - Debido a los errores de redondeo, en las condiciones del teorema anterior, igual pueden ser necesarios más de n pasos del método.
  - Se puede escribir el error en téminos de la norma 2 quedando la desigualdad

$$\|\mathbf{e}_k\|_2 \leqslant 2\kappa\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|\mathbf{e}_0\|_2,$$