Repaso clase pasada

Dada A simétrica y definida positiva, vimos que:

x^* es la única solución de $Ax=b$ si y solo si x^* es el único mínimo de la forma cuadrática

$$g(x) = \frac{1}{2} x^t A x - b^t x.$$

Para encontrar ese x^* proponemos un método iterativo

$$x_{k+1} = x_k + t_{k+1} v_{k+1}$$

donde comenzando en x_0, si asumimos que tenemos x_k, en el paso siguiente determinamos una dirección v_{k+1} y vimos que el $t \in \mathbb{R}$ que hace mínimo $g(x_k + t v_{k+1})$ viene dado por

$$t_{k+1} = - \frac{v_{k+1}^t (A x_k - b)}{v_{k+1}^t A v_{k+1}}$$

$$= - \frac{\langle v_{k+1}, A x_k - b \rangle}{\langle v_{k+1}, A v_{k+1} \rangle}$$

Método del descenso más rápido

(o método del gradiente)

Consiste en tomar las direcciones

$$v_{k+1} = - \nabla g(x_k)$$

(dirección de + rápido decrecimiento).

En este caso, vimos que $\nabla g(x) = Ax - b$

$\Rightarrow \quad r_{k+1} = -(Ax_k - b) = \underbrace{b - Ax_k}_{r_k \text{ (residuo)}}$

En este caso

$$t_{k+1} = -\frac{\langle r_k, -r_k\rangle}{\langle r_k, Ar_k\rangle} = \frac{\langle r_k, r_k\rangle}{\langle r_k, Ar_k\rangle}$$

y $\quad x_{k+1} = x_k + \dfrac{\langle r_k, r_k\rangle}{\langle r_k, Ar_k\rangle} \cdot r_k$

Vimos que este método converge en 1 paso si las curvas de nivel son circunferencias $(A = \lambda I)$ pero puede hacer muchas iteraciones (zig-zag) si ese no es el caso.

Método de direcciones conjugadas

Definimos $\quad \langle x, y\rangle_A = x^t A y$

que es un p.i en $\mathbb{R}^n$ si A es simétrica y definida positiva.

Decimos que x e y son A- conjugados o A- ortogonales si $\langle x, y\rangle_A = 0$.

<u>Teorema</u>: Dada $A \in \mathbb{R}^{n\times n}$ simétrica y definida positiva, se consideran $\{q_1, \dots, q_n\}$ n direcciones A- conjugadas y no nulas. Dado $b \in \mathbb{R}^n$, para cualquier $x_0 \in \mathbb{R}^n$, la iteración

$$x_{k+1} = x_k + t_{k+1}\, q_{k+1}$$

con $$t_{k+1} = -\frac{\langle q_{k+1}, Ax_k - b\rangle}{\langle q_{k+1}, Aq_{k+1}\rangle}$$

con $k = 0, \dots, n-1$ converge a la solución x^* de $Ax = b$ en a lo sumo n pasos, es decir $Ax_n = b$.

Dem: para $n=2$.

Veamos que $\{q_1, q_2\}$ es un conjunto

li: $\alpha q_1 + \beta q_2 = 0 \Rightarrow A(\alpha q_1 + \beta q_2) = 0$

$\Rightarrow \alpha A q_1 + \beta A q_2 = 0$

$\Rightarrow 0 = \langle q_1, \alpha A q_1 + \beta A q_2 \rangle$

$= \alpha \langle q_1, A q_1 \rangle + \beta \underbrace{\langle q_1, A q_2 \rangle}_{=0}$

xq' q_1 y q_2 son A-conj

$= \alpha \underbrace{\langle q_1, A q_1 \rangle}_{>0}$ porque $q_1 \neq 0$ y A es sim y def +

$\Rightarrow \alpha = 0 \Rightarrow \beta = 0$

Tomamos $x_0 \in \mathbb{R}^n \Rightarrow$

$x_1 = x_0 + t_1 q_1$ con $t_1 = -\dfrac{\langle q_1, A x_0 - b \rangle}{\langle q_1, A q_1 \rangle}$

$$x_2 = x_1 + t_2 q_2 \quad \text{con } t_2 = -\frac{\langle q_2, Ax_1 - b\rangle}{\langle q_2, Aq_2\rangle}$$

$$\Rightarrow \quad Ax_2 = Ax_1 + t_2 Aq_2$$

$$= Ax_0 + t_1 Aq_1 + t_2 Aq_2$$

$$\Rightarrow \quad Ax_2 - b = Ax_0 - b + t_1 Aq_1 + t_2 Aq_2$$

$$\Rightarrow \langle q_1, Ax_2 - b\rangle$$

$$= \langle q_1, Ax_0 - b\rangle + \underbrace{t_1}_{= -\frac{\langle q_1, Ax_0 - b\rangle}{\langle q_1, Aq_1\rangle}} \langle q_1, Aq_1\rangle + t_2 \underbrace{\langle q_1, Aq_2\rangle}_{=0}$$

$$\Rightarrow \quad \langle q_1, Ax_2 - b\rangle = 0$$

Además

$$\langle q_2, Ax_2 - b\rangle$$

$$= \langle q_2, Ax_0 - b\rangle + t_1 \underbrace{\langle q_2, Aq_1\rangle}_{=0} + \overbrace{t_2}^{-\frac{\langle q_2, Ax_0 - b\rangle}{\langle q_2, Aq_2\rangle}} \langle q_2, Aq_2\rangle$$

$\Rightarrow \langle q_2, Ax_2 - b\rangle = 0$

Como $\{q_1, q_2\}$ son base de $\mathbb{R}^2$ concluimos que $Ax_2 = b$.

Recién probamos que

- $\langle q_i, r_2\rangle = 0$ para $i = 1, 2$

Además

- $\langle q_1, r_1\rangle = \langle q_1, b - Ax_1\rangle$

Recordemos $x_1 = x_0 + t_1 q_1$

$\Rightarrow Ax_1 = Ax_0 + t_1 A q_1$

$\Rightarrow Ax_1 - b = Ax_0 - b + t_1 A q_1$

$\Rightarrow \langle q_1, Ax_1 - b\rangle$

$$= \langle q_1, Ax_0 - b\rangle + \underbrace{t_1}_{= -\frac{\langle q_1, Ax_0 - b\rangle}{\langle q_1, Aq_1\rangle}} \langle q_1, Aq_1\rangle = 0$$

$\Rightarrow \langle q_1, r_1\rangle = 0.$

En general (en $\mathbb{R}^n$) se prueba que:

$$\langle q_j, r_k \rangle = 0 \quad \forall \ j = 1, \dots, k.$$

Método de gradiente conjugado:

En este método no se eligen las direcciones A-conjugadas previamente como en el método anterior sino que se van eligiendo de manera que los residuos r_k sean mutuamente ortogonales.

Dado x_0, comenzamos con

$$r_1 = -\nabla g(x_0) = \underbrace{b - Ax_0}_{r_0}$$

$x_1 = x_0 + t_1 v_1$ con

$$t_1 = -\frac{\langle v_1, Ax_0 - b\rangle}{\langle v_1, Av_1\rangle} = \frac{\langle r_0, r_0\rangle}{\langle r_0, Ar_0\rangle}$$

Si $Ax_0 = b$ listo, si no $v_0 \neq 0$

Definimos

$$v_2 = r_1 + s_1 v_1 \Bigg\}$$ Es una comb. lineal de $r_1 = -\nabla g(x_1)$, $v_1 = -\nabla g(x_0)$

con $r_1 = \underbrace{b - Ax_1}_{-\nabla g(x_1)} = \underbrace{b - Ax_0}_{r_0} - t_1 A v_1$

y elegimos s_1 de manera que

$$\langle v_1, v_2\rangle_A = 0 :$$

$$\langle v_1, v_2\rangle_A = \langle v_1, Av_2\rangle$$

$$= \langle v_1, Ar_1 + s_1 Av_1\rangle$$

$$= \langle v_1, A r_1 \rangle + s_1 \langle v_1, A v_1 \rangle = 0$$

$$\Longleftrightarrow \quad s_1 = - \frac{\langle v_1, A r_1 \rangle}{\langle v_1, A v_1 \rangle}$$

Además

$$\langle r_0, r_1 \rangle = \langle r_0, r_0 - t_1 A v_1 \rangle$$

$$= \langle r_0, r_0 \rangle - t_1 \langle r_0, A r_0 \rangle = 0$$

por def de t_1.

Ahora tendremos

$$x_2 = x_1 + t_2 v_2 \quad \text{con } t_2 = - \frac{\langle v_2, A x_1 - b \rangle}{\langle v_2, A v_2 \rangle}$$

$$r_2 = b - A x_2 = - \nabla g(x_2)$$

$$= b - A x_1 - t_2 A v_2 = r_1 - t_2 A v_2$$

$$\text{y} \quad v_3 = r_2 + s_2 v_2$$

con δ_2 de manera que

$$\langle v_2, v_3 \rangle_A = 0 \ldots\ldots$$

En general: Dado $x_0 \in \mathbb{R}^n$

$$v_1 = r_0 = b - Ax_0 \neq 0$$

$$t_1 = -\langle v_1, Ax_0 - b \rangle / \langle v_1, Av_1 \rangle$$

$$= \langle r_0, r_0 \rangle / \langle r_0, Ar_0 \rangle$$

Supongamos que tenemos

$\{v_1, \ldots, v_{k-1}\}$ y $\{x_1, \ldots, x_{k-1}\}$

con $\langle v_i, v_j \rangle_A = 0 \quad i \neq j$

$i, j \leq k-1$

$$\langle r_i, r_j \rangle = 0 \quad i \neq j$$

$i, j \leq k-1$

$$x_{k-1} = x_{k-2} + t_{k-1} v_{k-1}$$

Definimos $v_k = r_{k-1} + s_{k-1} v_{k-1}$

con s_{k-1} de forma que

$$\langle v_{k-1}, v_k \rangle_A = 0$$

$$\shortparallel$$

$$\langle v_{k-1}, A v_k \rangle = \langle v_{k-1}, A r_{k-1} \rangle + s_{k-1} \langle v_{k-1}, A v_{k-1} \rangle$$

$$\Rightarrow s_{k-1} = - \frac{\langle v_{k-1}, A r_{k-1} \rangle}{\langle v_{k-1}, A v_{k-1} \rangle}$$

Con este s_{k-1} se puede ver que

$$\langle v_j, v_k \rangle_A = 0 \quad \forall j = 1.., k-1$$

ya que si $j \leq k-2$

$$\langle v_j, v_k \rangle_A = \langle v_j, r_{k-1} + s_{k-1} v_{k-1} \rangle$$

$$= \underbrace{\langle v_j, r_{k-1} \rangle_A}_{=0} + s_{k-1} \underbrace{\langle v_j, v_{k-1} \rangle_A}_{=0} = 0$$

porque

$$r_j = b - A x_j = b - [A x_{j-1} + t_j A v_j] = r_{j-1} - t_j A v_j$$

$$\Rightarrow A v_j = \frac{1}{t_j}\left(r_{j-1} - r_j\right) \Rightarrow \langle v_j, r_{k-1}\rangle_A =$$

$$= \langle A v_j, r_{k-1}\rangle = \frac{1}{t_j}\left(\langle r_{j-1}, r_{k-1}\rangle - \langle r_j, r_{k-1}\rangle\right) = 0$$

con $j \leq k-2$.

Ahora calculamos

$$t_k = -\frac{\langle v_k, \overbrace{A x_{k-1} - b}^{-r_{k-1}}\rangle}{\langle v_k, A v_k\rangle}$$

$$= \frac{\langle r_{k-1} + s_{k-1} v_{k-1}, r_{k-1}\rangle}{\langle v_k, A v_k\rangle}$$

$$= \frac{\langle r_{k-1}, r_{k-1}\rangle}{\langle v_k, A v_k\rangle} + s_{k-1} \underbrace{\frac{\langle v_{k-1}, r_{k-1}\rangle}{\langle v_k, A v_k\rangle}}_{=0}$$

$$\Rightarrow \boxed{t_k = \frac{\langle r_{k-1}, r_{k-1}\rangle}{\langle v_k, A v_k\rangle}}$$

Además
$$x_k = x_{k-1} + t_k v_k$$

$$r_k = b - A x_k$$

$$= \underbrace{b - A x_{k-1}}_{r_{k-1}} - t_k A v_k$$

Usando que $v_k = r_{k-1} + s_{k-1} v_{k-1}$

tenemos que

$$\langle v_k, A v_k \rangle = \langle r_{k-1} + s_{k-1} v_{k-1}, A v_k \rangle$$

$$= \langle r_{k-1}, A v_k \rangle + s_{k-1} \langle \underbrace{v_{k-1}, A v_k}_{=0} \rangle$$

Juntando todo, tenemos:

$$\langle r_{k-1}, r_k \rangle = \langle r_{k-1}, r_{k-1} \rangle$$

$$+ \; t_k \langle r_{k-1}, A v_k \rangle$$

$$= \langle r_{k-1}, r_{k-1} \rangle + t_k \langle v_k, A v_k \rangle = 0$$

por la definición de t_k.

y análogamente se puede probar que:

$$\boxed{\langle r_j, r_k \rangle = 0 \quad \forall\, j = 1, \dots, k-1.}$$

Además

$$S_k = - \frac{\langle v_k, A r_k \rangle}{\langle v_k, A v_k \rangle}$$

$$\underset{\substack{\downarrow \\ A \\ \text{simétrica}}}{=} - \frac{\langle A v_k, r_k \rangle}{\langle v_k, A v_k \rangle} = \frac{-t_k \langle A v_k, r_k \rangle}{\underbrace{t_k \langle v_k, A v_k \rangle}_{t_k \neq 0 \text{ porque } r_k \neq 0}}$$

$$t_k A v_k = r_k - r_{k-1}$$

$$\langle t_k A v_k, r_k \rangle = \langle r_k, r_k \rangle - \underbrace{\langle r_{k-1}, r_k \rangle}_{\shortparallel\atop 0}$$

$$= \langle r_k, r_k \rangle$$

y

$$t_k \langle v_k, A v_k \rangle = - \langle r_{k-1}, r_{k-1} \rangle$$

$$\Rightarrow \boxed{s_k = \frac{\langle r_k, r_k \rangle}{\langle r_{k-1}, r_{k-1} \rangle}}$$

En resumen: Dada A sim y def positiva y x^0 cualquiera

Definimos

$$r_0 = b - Ax_0, \quad v_1 = r_0$$

y para $k=1,\dots,n-1$, si tenemos

$\{v_1,\dots,v_{k-1}\}$ y $\{x_1,\dots,x_{k-1}\}$

con $r_{k-1}=b-Ax_k \neq 0$

y $\langle v_i, v_j\rangle_A = 0 \quad \forall i\neq j$
$1\leq i,j\leq k-1$

$$\langle r_i, r_j\rangle = 0 \quad \forall i\neq j$$
$$1\leq i,j\leq k-1$$

$\Rightarrow$ definimos

$$s_{k-1}=\frac{\langle r_{k-1}, r_{k-1}\rangle}{\langle r_{k-2}, r_{k-2}\rangle}$$

$$v_k = r_{k-1}+s_{k-1}v_{k-1}$$

$$t_k = -\frac{\langle r_{k-1}, r_{k-1}\rangle}{\langle v_k, Av_k\rangle}$$

$$x_k = x_{k-1} + t_k v_k$$

$$r_k = r_{k-1} - t_k A v_k$$

Ver Burden Sección 7.6

pág 354 a pág 360.