1 Teoría (40pts)

1.1 Gradiente Conjugado

Álvaro Acedo Trueba
174052

- Demuestre que si los vectores no nulos $p_1, p_2, ..., p_l$ satisfacen que :

$$p_i^T A p_j = 0, \forall i \neq j,$$

y A es simétrica y positiva definida, entonces los vectores son linealmente independientes.

① Suponemos que no son linealmente independientes

i.e. $\exists \; \alpha_K \neq 0$ t.q.

$$\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_K p_K + \dots + \alpha_l p_l = 0 \qquad (1 < K < l)$$

Premultiplicamos por $p_K A$

$$\Rightarrow \quad p_K A(\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_K p_K + \dots + \alpha_l p_l) = 0$$

$$\Rightarrow \quad \alpha_1 p_K A p_1 + \alpha_2 p_K A p_2 + \dots + \alpha_K p_K A p_K + \dots + \alpha_l p_K A p_l = 0$$

Pero sabemos que $p_K A p_j = 0 \quad \forall K \neq j \quad , j = 1, \dots, l$

$$\Rightarrow \alpha_1 \underbrace{\cancel{p_K A p_1}}_{0} + \alpha_2 \underbrace{\cancel{p_K A p_2}}_{0} + \dots + \alpha_K p_K A p_K + \dots + \alpha_l \underbrace{\cancel{p_K A p_l}}_{0} = 0$$

$$\Rightarrow \quad \alpha_K p_K A p_K = 0$$

$$\Rightarrow \quad \alpha_K = 0 \quad ó \quad p_K A p_K = 0 \qquad !$$

Es una contradicción ya que habíamos supuesto que $\alpha_K \neq 0$ y sabemos que A es positiva definida $\Rightarrow p_K A p_K > 0$

$\therefore \; \{p_1, p_2, \dots, p_l\}$ es un conjunto linealmente independiente //

- Dado este resultado, ¿Por qué el gradiente conjugado converge en a lo más n iteraciones?.

② Dado un punto inicial $x_0 \in \mathbb{R}^n$ y un conjunto de direcciones conjugadas $\{p_1, \dots, p_n\}$

Generamos la sucesión $\{x_K\}$ de la siguiente forma

$$x_{K+1} = x_K + \alpha_K p_K \qquad (*)$$

Continúa

donde $$\alpha_k = \frac{-r_k^T p_k}{p_k^T A p_k} \quad (**)$$

Notemos que en el inciso ① probamos que $\{p, \ldots, p_n\}$ es un conjunto linealmente independiente

$$\Rightarrow \quad \text{span}\{p_1, \ldots, p_n\} = \mathbb{R}^n$$

Entonces podemos escribir

$$x^* - x_0 = \gamma_1 p_1 + \gamma_2 p_2 + \ldots \gamma_k p_k + \ldots + \gamma_n p_n \quad , \text{ para } \gamma_1, \ldots, \gamma_n \text{ adecuados}$$

Multiplicando por $p_k^T A$

$$\Rightarrow p_k^T A (x^* - x_0) = \gamma_1 \cancelto{0}{p_k^T A p_1} + \gamma_2 \cancelto{0}{p_k^T A p_2} + \ldots + \gamma_k p_k^T A p_k + \ldots + \gamma_n \cancelto{0}{p_k^T A p_k}$$

Despejando γ_k

$$\Rightarrow \gamma_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k} \quad (\text{I}) \qquad (p_k^T A p_k \neq 0)$$

Veamos que estos coeficienter γ_k coinciden con los α_k generados por la fórmula $(**)$

Como x_k es generado por $(*)$ $\quad (x_{k+1} = x_k + \alpha_k p_k)$

$$\begin{aligned}
\Rightarrow \quad x_k &= x_{k-1} + \alpha_{k-1} p_{k-1} \\
&= x_{k-2} + \alpha_{k-2} p_{k-2} + \alpha_{k-1} p_{k-1} \\
&\vdots \\
&= x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \ldots + \alpha_{k-1} p_{k-1} \qquad (\text{II})
\end{aligned}$$

premultiplicando por $p_k^T A$:

$$\Rightarrow \quad p_k^T A x_k = p_k^T A x_0 + \alpha_0 \cancelto{0}{p_k^T A p_0} + \alpha_1 \cancelto{0}{p_k^T A p_1} + \ldots + \alpha_{k-1} \cancelto{0}{p_k^T A p_{k-1}}$$

continúa

$$\Rightarrow \quad p_k^T A x_k - p_k^T A x_0 = 0$$

$$\Rightarrow \quad p_k^T A (x_k - x_0) = 0$$

Entonces

$$p_k^T A(x^* - x_0) = p_k^T A(x^* - x_k)$$
$$= p_k^T A(b - Ax_k)$$
$$= -p_k^T r_k$$

Comparando $(**)$ con (I) resulta que $\underline{\underline{\alpha_k = \gamma_k}}$

$\therefore$ El algoritmo termina en n iteraciones, pues

(II) $x_n = x_0 + \alpha p_0 + \ldots + \alpha_n p_n = x_0 + \gamma_1 p_1 + \ldots + \gamma_n p_n = x^*$ (restando x_0)

$\therefore$ $x^* = x_n$ //

1.2 BFGS

- Verifique que B_{k+1} y H_{k+1} son inversas una de la otra.

Obs:
Para simplificar la notación, voy a omitir los subíndices k

$$B_{k+1} \cdot H_{k+1} = B_{k+1}\left((I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T\right)$$
$$= (B - B\rho s y^T) H (I - \rho y s^T) + B\rho s s^T$$

Por la ecuación secante, $Bs = y$

$$= (B - \rho y y^T) H (I - \rho y s^T) + \rho y s^T$$

$$= \left(B - \frac{Bss^TB}{s^TBs}\right) H (I - \rho y s^T) + \rho y s^T$$

$H = B^{-1}$

$$= \left(I - \frac{Bss^T}{s^TBs}\right)(I - \rho y s^T) + \rho y s^T$$

$$= (I - \cancel{\rho y s^T}) - \left(\frac{Bss^T}{s^TBs}\right)(I - \rho y s^T) + \cancel{\rho y s^T}$$

$$= I - \frac{Bss^T}{s^TBs}(I - \rho y s^T)$$

continúa

$$= I - \frac{Bss^T}{s^TBs} - \frac{Bss^T}{s^TBs}\frac{1}{s^Ty}ys^T$$

$\left\{ \text{Sustituyendo } \rho = \frac{1}{y^Ts} \right.$

$$= I - \frac{Bss^T}{s^TBs} - \frac{\cancel{(s^Ty)}\,Bss^T}{\cancel{(s^Ty)}\,s^TBs}$$

$$= I - \cancel{\frac{Bss^T}{s^TBs}} - \cancel{\frac{Bss^T}{s^TBs}}$$

$$= I$$

$\therefore$ B_{k+1} y H_{k+1} son inversas una de la otra //