Final                Análisis Aplicado                David Isaac López Romero
CU: 173993

Teoría

1.1 Gradiente Conjugado

a) La demostración procede por contradicción

Supongamos que $p_1, \dots, p_l$ son linealmente dependientes. Entonces existe $p_k \in \mathbb{R}$ tal que

$$p_k = \sum_{i \neq k} \alpha_i p_i \quad (p_k \text{ es combinación lineal de los } p_i)$$

con $\alpha_i \in \mathbb{R}$ y $p_i \in \mathbb{R}^n$. Entonces, para alguna $j \in \{1, \dots, l\}$ se tiene que

$$\overset{(*)}{\Rightarrow} p_j^T A p_k = p_j^T A \left( \sum_{i \neq k} \alpha_i p_i \right) = \sum_{i \neq k} \left( p_j^T A \alpha_i p_i \right) = \sum_{i \neq k} \alpha_i \left( p_j^T A p_i \right)$$

$$= \alpha_j p_j^T A p_j$$

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

Así pues, por un lado $(*)$ $0 = p_j^T A p_k$ por hipótesis pero probamos por otro lado que $p_j^T A p_k = \alpha_j p_j^T A p_j \neq 0$ !

De este modo tenemos contradicción.

Así, $\{p_1, \dots, p_l\}$ es linealmente independiente

b) Sea $A \in \mathbb{R}^{n \times n}$ simétrica definida positiva y $x^*$ la solución del problema. Consideremos $p_0, \dots, p_{n-1}$ vectores tales que

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

Por el anterior inciso vemos que $\text{span}(\{p_0, \dots, p_{n-1}\}) = \mathbb{R}^n$

Consideremos a $x_0 \in \mathbb{R}^n$ y resulta que

$$x^* - x_0 = \sum_{i=0}^{n-1} \sigma_i p_i \quad \text{con } \sigma_i \in \mathbb{R}, \ p \in \mathbb{R}^n$$

donde

$$\sigma_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$$

Mini demo de $\sigma_k$

$$x^* = x_0 + \sigma_1 p_0 + \cdots + \sigma_{n-1} p_{n-1}$$

$$\Leftrightarrow x^* - x_0 = \sigma_1 p_0 + \cdots + \sigma_{n-1} p_{n-1}$$

$$p_k^T A (x^* - x_0) = p_k^T A \left( \sum_{i=0}^{n-1} \alpha_i p_i \right) \underset{|}{=} \sigma_k p_k^T A p_k$$

$$\underset{p_k^T A p_k \neq 0}{\Leftrightarrow} \quad \sigma_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$$

(análogo a inciso (1) $p_j^T A p_i = 0, \forall i \neq j$)

Ahora bien, $x^* - x_0 = \sum_{i=0}^{n-1} \sigma_i p_i$ ; $\sigma_k = \dfrac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$

Dado que el algoritmo genera $x_k$ de la forma

$$x_k = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{k-1} p_{k-1}$$

resulta que

$$x_k - x_0 = \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_{k-1} p_{k-1}$$

$$\Rightarrow p_k^T A (x_k - x_0) = p_k^T A \left( \sum_{i=0}^{k-1} \alpha_i p_i \right) \underset{|}{=} 0$$

$\left( p_j^T A p_i = 0 \ \forall i \neq j \right)$

Por lo tanto,

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

De este modo como $\sigma_k = \dfrac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$

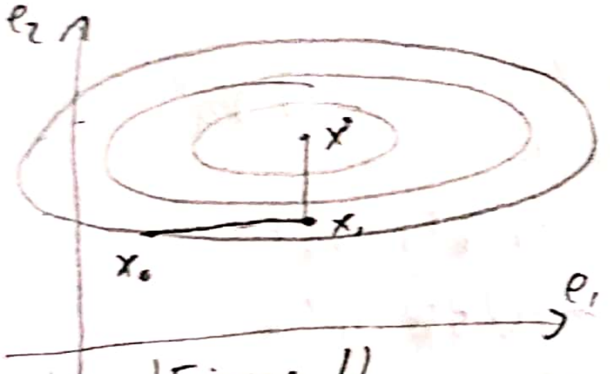
y como el paso es $x_{k+1} = x_k + \alpha_k p_k$

con $\alpha_k = \dfrac{-r_k^T p_k}{p_k^T A p_k}$

resulta que $\sigma_k = \alpha_k$

Así se haya el óptimo en $n$ pasos a lo más.

Este resultado es válido para cualquier problema. Primero si las coordenadas del problema tiene elipses paralelas (Figura 1) es fácil de ver pues se recorre junto a las elipses.




(Figura 1)

Si las elipses están rotadas como en la Figura siguiente (Figura 2) tarda más de $n$ pasos. Lo único que requerimos es cambiar de coordenadas con otra base canónica y se cumple como en la última Figura (3)

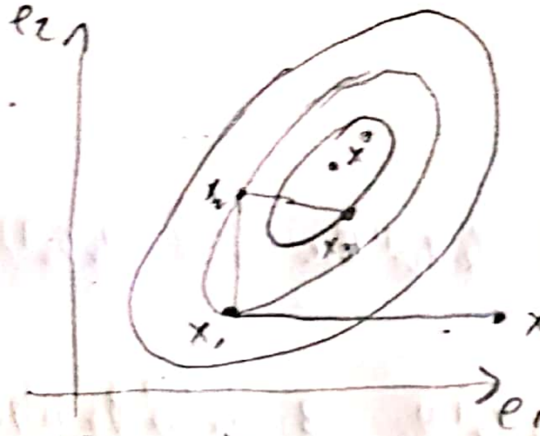



Figura (2)

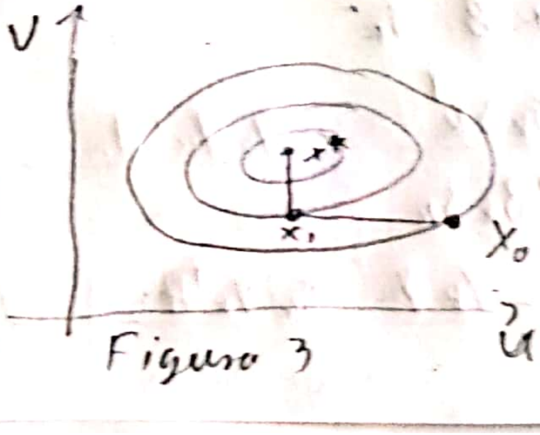



Figura 3

1.2) BFGS

a) P.D. $B_{k+1} H_{k+1} = I = H_{k+1} B_{k+1}$

Primero recordemos las definiciones y propiedades

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T$$

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

con $\rho_k = \frac{1}{y_k^T s_k}$, $B_{k+1} s_k = y_k$, $H_{k+1} y_k = s_k$, $H_k = B_k^{-1}$

$$B_{k+1} H_{k+1} = B_{k+1} \{ \underbrace{(I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T)}_{I} + \underbrace{\rho_k s_k s_k^T}_{II} \}$$

De II obtenemos que

$$B_{k+1} \rho_k s_k s_k^T = \left\{ B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \right\} \frac{1}{y_k^T s_k} s_k s_k^T$$

$$= \frac{B_k s_k s_k^T}{y_k^T s_k} - \frac{B_k s_k \cancel{(s_k^T B_k s_k)} s_k^T}{\cancel{(s_k^T B_k s_k)} y_k^T s_k} + \frac{y_k \cancel{(y_k^T s_k)} s_k^T}{\cancel{(y_k^T s_k)} y_k^T s_k}$$

$$= \cancel{\frac{B_k s_k s_k^T}{y_k s_k^T}} - \cancel{\frac{B_k s_k s_k^T}{y_k^T s_k}} + \frac{y_k s_k^T}{y_k^T s_k} = \frac{y_k s_k^T}{y_k^T s_k} \quad \ldots (*)$$

De I obtenemos que

$$B_{k+1} (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T)$$

$$= \left[ \left( B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \right) (H_k - \rho_k s_k y_k^T H_k) \right] (I - \rho_k y_k s_k^T)$$

$$= \left( I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \frac{y_k y_k^T H_k}{y_k^T s_k} - \frac{B_k s_k y_k^T H_k}{y_k^T s_k} + \frac{B_k s_k \cancel{(s_k^T B_k s_k)} y_k^T H_k}{\cancel{(s_k^T B_k s_k)} y_k^T s_k} - \frac{y_k y_k^T s_k y_k^T H_k}{y_k^T s_k \; y_k^T s_k} \right) \left( I - \frac{y_k s_k^T}{y_k^T s_k} \right)$$

$$= \left( I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \frac{y_k y_k^T H_k}{y_k^T s_k} - \cancel{\frac{B_k s_k y_k^T H_k}{y_k^T s_k}} + \cancel{\frac{B_k s_k y_k^T A_k}{y_k^T s_k}} - \frac{y_k y_k^T s_k s_k^T H_k}{y_k^T s_k \, y_k^T s_k} \right) \left( I - \frac{y_k s_k^T}{y_k^T s_k} \right)$$

$$= \left( I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \frac{s_k s_k^T H_k}{y_k^T s_k} - \frac{s_k \cancel{(y_k^T s_k)} s_k^T H_k}{y_k^T s_k \cancel{(y_k^T s_k)}} \right) \left( I - \frac{y_k s_k^T}{y_k^T s_k} \right)$$

$$= \left( I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \cancel{\frac{s_k s_k^T H_k}{y_k^T s_k}} - \cancel{\frac{s_k s_k^T H_k}{y_k^T s_k}} \right) \left( I - \frac{y_k s_k^T}{y_k^T s_k} \right)$$

$$= \left( I - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} \right) \left( I - \frac{y_k s_k^T}{y_k^T s_k} \right)$$

$$= I - \frac{s_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T}{s_k^T B_k s_k} + \frac{B_k s_k \cancel{(y_k^T s_k)} s_k^T}{s_k^T B_k s_k \cancel{(y_k^T s_k)}}$$

$$= I - \frac{s_k y_k^T}{y_k^T s_k} - \cancel{\frac{B_k s_k s_k^T}{s_k^T B_k s_k}} + \cancel{\frac{B_k s_k s_k^T}{s_k^T B_k s_k}}$$

$$= I - \frac{s_k y_k^T}{y_k^T s_k} \quad \ldots (**)$$

Finalmente sumando (*) y (**)

$$H_{k+1} B_{k+1} = I - \frac{s_k y_k^T}{y_k^T s_k} + \frac{s_k y_k^T}{y_k^T s_k} = I$$

Por lo tanto $\quad H_{k+1} = B_{k+1}^{-1}$