Ejercicio 1:

Demostrar que si ~~A es~~ los vectores no nulos $p_1, p_2, \ldots, p_n$ satisfacen $p_i^T A p_j = 0 \ \forall i \neq j$ y A es simétrica y pos. def. entonces los vectores son lin. indep.

Supongamos que los vectores no son linealmente independientes entonces podemos escribir a uno en términos de los demás, es decir,

$$p_i = \alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_{i-1} p_{i-1} + \alpha_{i+1} p_{i+1} + \ldots + \alpha_n p_n \quad p_i \neq 0$$

luego pre multiplicamos la ecuación anterior por $p_i^T A$ y obtenemos

$$p_i^T A p_i = \alpha_1 p_i^T A p_1 + \ldots + \alpha_n p_i^T A p_n = 0 \leftarrow (p_i^T A p_j = 0 \ \forall i \neq j)$$

Sin embargo, como A es una matriz simétrica y positiva definida tenemos que $x^T A x > 0 \ \forall x \in \mathbb{R}^n$

∴ Hemos llegado a una contradicción que vino de suponer que los vectores no eran linealmente independientes.

∴ Los vectores son linealmente independientes

Ejercicio 2: De este resultado, ¿por qué el gradiente conjugado converge en a lo más n iteraciones.

Al ser linealmente independiente los vectores $\{p_i\}$ estos generan todo el espacio vectorial, en particular, $x^* - x_0$, con x^* óptimo y x_0 el inicial

Esto es $x^* - x_0 = \beta_1 p_1 + \ldots + \beta_n p_n$. Luego si multiplicamos por $p_i^T A$

obtenemos $p_i^T A (x^* - x_0) = \beta_i p_i^T A p_i$ (ya que $p_i^T A p_j = 0 \ \forall i \neq j$)

$$\therefore \beta_i = \frac{p_i^T A (x^* - x_0)}{p_i^T A p_i}$$

Luego, sea x_k el resultado del método del gradiente conjugado tenemos que este se puede escribir como:

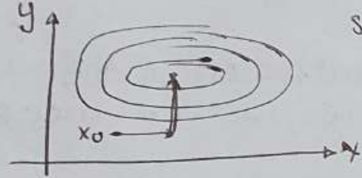
$$x_k = x_0 + \underbrace{\alpha_1 p_1 + \ldots + \alpha_{k-1} p_{k-1}}_{\text{se cancelan por ser conjugados}}$$ y pre multiplicando por $p_k^T A$ obtenemos

$$\underline{p_k^T A x_k = p_k^T A x_0}$$

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

$\therefore \beta_k = \dfrac{-p_k^T r_k}{p_k^T A p_k}$ que son el tamaño del paso generado por el método de gradiente → converge en n pasos.

En sí, lo que sucede es que se está minimizando en una dimensión siguiendo al eje y posteriormente a la otra.



$\therefore$ Como se minimiza por la dirección, esto es a lo más en n pasos.

Las elipses representan las curvas de nivel, cuando A es diagonal de $\emptyset$

Ejercicio 3:

Muestre que la segunda condición de wolfe fuerte implica la condición de curvatura: $S_k^T u_k > 0$

la segunda condición de wolfe nos dice que

$$|\nabla f(x_k + \alpha_k p_k)^T p_k| \leq C_2 |\nabla f(x_k)^T p_k|$$

luego, de esto tenemos que   al ser dirección de descenso

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq -C_2 |\nabla f(x_k)^T p_k| = C_2 \nabla f(x_k)^T p_k$$

restando $\nabla f(x_k)^T p_k$ de ambos lados obtenemos:

$$\nabla f(x_k + \alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k \geq (C_2 - 1) \nabla f(x_k)^T p_k$$

luego, como la condición de wolfe nos dice que $C_2 < 1$, tenemos que

$$(C_2 - 1) \nabla f(x_k)^T p_k > 0$$

Así,

$$\nabla f(x_k + \alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k > 0$$

si multiplicamos ambos lados por $\alpha_k (>0)$ obtenemos

$$0 < \alpha_k \nabla f(x_k + \alpha_k p_k)^T p_k - \alpha_k \nabla f(x_k)^T p_k = \nabla f(x_k + \alpha_k p_k)^T (\alpha_k p_k) - \nabla f(x_k)^T (\alpha_k p_k)$$

$$= (\nabla f(x_k + \alpha_k p_k)^T - \nabla f(x_k)^T) \cdot \alpha_k p_k.$$

$$= (\nabla f(x_k + \alpha_k p_k) - \nabla f(x_k))^T \cdot \alpha_k p_k > 0$$

luego, como $y_k = \nabla f(x_k + \underbrace{\alpha_k p_k}_{S_k}) - \nabla f(x_k)$ y $S_k = \alpha_k p_k$ tenemos:

$$\therefore\ y_k^T S_k > 0$$

EJERCICIO 4: verificar que B_{k+1} y H_{k+1} son inversas

Por construcción podemos suponer B_k, H_k inversas (B_0, H_0 lo son + inducción)
luego, para probar que son inversas basta ver que $B_{k+1}H_{k+1}=I$.

$$B_{k+1}H_{k+1} = B_{k+1}\left[(I-\rho_k s_k y_k^T)H_k(I-\rho_k y_k s_k^T)+\rho_k s_k s_k^T\right]$$

Notamos que $B_{k+1}s_k = y_k$ entonces:

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

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

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

como $B_kH_k = I$ por hipótesis tenemos:

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

Expandiendo obtenemos

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

luego, $\rho_k \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} y_k s_k^T = \frac{1}{s_k^T y_k}\cdot\frac{B_k s_k \; s_k^T \; y_k s_k^T}{s_k^T B_k s_k}$

luego como $B_k s_k$ es escalar, cancelamos $s_k^T y_k$

$$= \frac{B_k s_k s_k^T}{s_k^T B_k s_k} \Rightarrow \frac{B_k s_k s_k^T}{s_k^T B_k s_k} = \frac{B_k s_k s_k^T}{s_k^T B_k s_k}\rho_k y_k s_k$$

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

$\therefore B_{k+1}H_{k+1} = I$

con $\rho = y_k^T s_k$

$\ldots + \rho_k s_k s_k^T$

Ejercicio 1:

Demostrar que si ~~A es~~ los vectores no nulos $p_1, p_2, \ldots, p_n$ satisfacen $p_i^T A p_j = 0 \;\forall i \neq j$ y A es simétrica y pos. def. entonces los vectores son lin. indep.

Supongamos que los vectores no son linealmente independientes entonces podemos escribir a uno en términos de los demás, es decir,

$$p_i = \alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_{i-1} p_{i-1} + \alpha_{i+1} p_{i+1} + \ldots + \alpha_n p_n$$

Luego, premultiplicamos la ecuación por $p_j^T A$ y obtenemos

$\therefore \underbrace{p_j^T A p_i}_{0} = \alpha_j p_j^T A p_j$ luego, al ser A positiva definida $p_j^T A p_j > 0$

(por conjugados)

Así, $\alpha_j = 0$

Aplicando esto para $j = 1, 2, \ldots, n$; $j \neq i$ tenemos que

$\alpha_j = 0 \;\forall j = 1, 2, \ldots, n, \; j \neq i$

$\therefore \{p_1, \ldots, p_n\}$ es linealmente independiente.