1: Gradiente Conjugado:

Demuestre que si los vectores $p_1, \dots, p_\ell$ no nulos satisfacen

$$p_i^T A p_j = 0$$

con A s.p.d. $\Rightarrow \{p_k\}_k$ es L.I.

Dem

Sean α_k escalares t.q.

$$\sum_{i=1}^{\ell} \alpha_i p_i = 0$$

Sea p_j fijo

$$\Rightarrow \left(\sum_{i=1}^{\ell} \alpha_i p_i^T\right) A p_j = 0$$

$$\Rightarrow \sum_{i=1}^{\ell} \alpha_i p_i^T A p_j = 0$$

$$\Rightarrow \sum_{i \neq j}^{\ell} \alpha_i \cancel{p_i^T A p_i}^{\,0} + \alpha_j p_j^T A p_j = 0$$

$$\Rightarrow \alpha_j (p_j^T A p_j) = 0$$

Pero como A es s.p.d $\Rightarrow p_j^T A p_j > 0 \ \forall j = 1, \dots, \ell$

$$\Rightarrow \alpha_j = 0 \ \forall j = 1, \dots, \ell$$

$\therefore \{p_k\}$ es L.I.

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

Como $\{p_i\}$ es l.i

$$\Rightarrow \quad x^* - x_0 = \beta_1 p_1 + \cdots + \beta_n p_n$$

con β_i únicos

fijemos p_k y tomemos

$$p_k^T A (x^* - x_0) = \sum \beta_i p_k^T A p_i = \beta_k p_k^T A p_k$$

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

que conciden con los valores α_k del método de gradiente conjugado

$\Rightarrow$ En cada paso se calculan aproximaciones β_i

$\Rightarrow$ A lo más en n pasos converge.

2. Quasi-Newton

Muestre que la condición de Wolfe implica la condición de curvatura:

$$S_k^T y_k > 0$$

Tomemos $s_k = \alpha p_k$ y $y_k = \nabla f_{k+1} - \nabla f_k$

La condición de Wolfe.

$$|\nabla f(x_k + \alpha_k p_k)^T p_k| \leq c_2 |\nabla f(x_k)^T p_k|$$

$$\Rightarrow \nabla f(x_k + \alpha_k p_k)^T p_k \geq -c_2 |\nabla f(x_k)^T p_k|$$

Como p_k es dirección de descenso $\nabla f(x_k)^T p_k < 0$

Luego

$$\nabla f(x_k + \alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k \quad (y_k^T) p_k$$

$$> 0$$

$$\Rightarrow \quad \alpha_k y_k^T p_k \geq 0$$

$$\Rightarrow \quad y_k^T (\alpha_k p_k) = y_k^T s_k > 0$$

$$\Rightarrow \quad \underline{s_k^T y_k > 0}$$

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

$$\rho = \frac{1}{y_k^T s_k}$$

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

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

$$B_{k+1} H_{k+1} =$$

$$(I - \rho_k y_k s_k^T) B_k (I - \rho_k s_k y_k^T)(I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k)^T$$

$$+ (I - \rho_k y_k s_k^T) B_k (I - \rho_k s_k y_k^T) \rho_k s_k s_k^T$$

$$+ \rho_k y_k y_k^T (I - \rho_k s_k y_k)^T H_k (I - \rho_k y_k s_k^T)$$

$$+ \rho_k^2 y_k y_k^T s_k s_k^T$$

Notemos que:

$$(I - \rho_k y_k s_k^T) B_k (I - \rho_k s_k y_k^T)(I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) =$$

$$(B_k - \rho_k y_k s_k^T B_k - B_k \rho_k s_k y_k^T + \rho_k^2 y_k s_k^T B_k s_k y_k^T)$$

$$(H_k - \rho_k s_k y_k^T H_k - H_k \rho_k y_k s_k^T + \rho_k^2 s_k y_k^T H_k y_k s_k^T) =$$

$$I - \rho_k B_k s_k y_k^T H_k - \rho_k y_k s_k^T + \rho_k^2 B_k s_k y_k^T H_k y_k s_k^T$$

$$- \rho_k y_k s_k^T + \rho_k^2 y_k s_k^T B_k s_k y_k H_k + \rho_k^2 y_k s_k^T y_k s_k^T$$

$$- \rho_k^3 y_k s_k^T B_k s_k y_k^T H_k y_k s_k^T$$

$$- \rho_k B_k s_k y_k^T H_k + \rho_k^3 B_k s_k y_k^T H_k + \rho_k^2 B_k s_k y_k^T H_k y_k s_k^T$$

$$- \rho_k^4 B_k s_k y_k^T H_k y_k s_k^T$$

$$+ \rho_k^2 y_k s_k^T B_k s_k y_k^T H_k - \rho_k^4 y_k s_k B_k s_k y_k^T H_k$$

$$- \rho_k^3 y_k s_k^T B_k s_k y_k^T H_k y_k s_k^T$$

$$+ \rho_k^5 y_k s_k^T B_k s_k y_k^T H_k y_k s_k^T$$

Otro término: $(I - \rho_k y_k s_k^T) B_k (I - \rho_k s_k y_k^T) \rho_k s_k s_k^T =$

$$\rho_k B_k s_k s_k^T - \rho_k^2 y_k s_k^T B_k s_k s_k^T - \rho_k^3 B_k s_k s_k^T$$

$$+ \rho_k^4 y_k s_k^T B_k s_k s_k^T$$

Otro término: $\rho_k y_k y_k^T (I - \rho_k s_k y_k)^T H_k (I - \rho_k y_k s_k^T)$

$$\rho_k y_k y_k^T H_k - \rho_k^3 y_k y_k^T s_k y_k^T H_k - \rho_k^2 y_k y_k^T H_k y_k s_k^T$$

$$+ \rho_k^4 y_k y_k^T H_k y_k s_k^T$$

Todos los términos se cancelan y solo sobrevive la identidad por lo que la inversa de B_{k+1} es H_{k+1}.