1 Gradiente Conjugado

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

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

y A es simétrica y positiva definida $\Rightarrow$ los vectores son linealmente independientes.

Dem:

Supongamos $\alpha_0, \dots, \alpha_\ell \in \mathbb{R}$ t.

$$\sum_{i=1}^{\ell} \alpha_i p_i = \bar{0} \leftarrow \text{vector } 0 \text{ de } n \text{ entradas.}$$

vamos a demostrar que $\forall i \in \{1, \dots, \ell\}$ el escalar $\alpha_i = 0$.

multiplicando por $p_j^T A$, tenemos:

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

por hipótesis, tenemos que si $i \neq j$

$$\Rightarrow \quad p_j^T A \, p_i = 0$$

Por lo tanto nos queda únicamente el sumando para el cual $i = j$:

$$\Rightarrow \quad \alpha_j \, p_j^T A \, p_j = 0$$

pero sabemos que $p_j^T A \, p_j > 0 \Rightarrow \alpha_j = 0$

$\therefore$ son linealmente independientes $\blacksquare$

---

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

Dado que ya vimos que las direcciones p_i son linealmente independientes entre sí, entonces podemos asumir que $p_1, \ldots, p_n$ generan el espacio $\mathbb{R}^n$.
Ahora, combinando esto con el hecho de que (como vimos en clase)

$$\alpha_i = - \frac{r_i^T p_i}{p_i^T A p_i} \quad \text{con } r_i \text{ residuo}$$

es el minimizador de $\Phi(x) = \frac{1}{2} x^T A x - b^T x$

es el minimizador a lo largo de $x_i + \alpha_i p_i$, es decir, tenemos n 'minimizadores' cada uno para una de las p_i direcciones, pero solo 1 es el que minimiza toda la función Φ.

$\therefore$ tenemos que probar a lo más n iteraciones para llegar a la convergencia ✓.

2 Quasi-Newton

1. Muestre que la segunda condición fuerte de Wolfe implica la condición de curvatura

$$s_k^T y_k > 0.$$

Sabemos que $s_k = x_{k+1} - x_k = \alpha_k p_k$,

$$y_k = \nabla f_{k+1} - \nabla f_k$$

y además que: (2ª cond. Wolfe)

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 \nabla f_k^T p_k \quad , \quad 0 < c_1 < c_2 < 1$$

Multiplicando ambos lados por α_k tenemos:

$$\Rightarrow \nabla f_{k+1}^T \underbrace{\alpha_k p_k}_{s_k} \geq c_2 \nabla f_k^T \underbrace{\alpha_k p_k}_{s_k}$$

$y_k = \nabla f_{k+1} - \nabla f_k$

$$= \nabla f_{k+1}^T s_k \geq c_2 \nabla f_k^T s_k$$

- De $y_k = \nabla f_{k+1} - \nabla f_k \Rightarrow \nabla f_{k+1} = y_k + \nabla f_k$

$$\Rightarrow \nabla f_{k+1}^T = y_k^T + \nabla f_k^T$$

sustituyendo:

$$(y_k^T + \nabla f_k^T) s_k \geq c_2 \nabla f_k^T s_k$$

$$\Rightarrow y_k^T s_k + \nabla f_k^T s_k \geq c_2 \nabla f_k^T s_k$$

$$\Rightarrow y_k^T s_k \geq c_2 \nabla f_k^T s_k - \nabla f_k^T s_k$$

$$\Rightarrow y_k^T s_k \geq (c_2 - 1) \nabla f_k^T s_k$$

y dado que $c_2 < 1$ y que p_k es una dirección de descenso, $\Rightarrow (c_2 - 1) \nabla f_k^T s_k > 0$

$$\therefore y_k^T s_k = s_k^T y_k > 0 \checkmark \quad \blacksquare$$

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

Partiendo del hecho de que $H_k = B_k^{-1}$

Tenemos que

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

y

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

• Suponiendo de antemano que lo son, bastaría con multiplicarlas y verificar que llegamos a la identidad:

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

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

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

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

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

$$\overset{I}{\cancel{B_k H_k}} - \rho_k B_k s_k y_k^T H_k - \overset{I}{\cancel{B_k H_k}} \rho_k y_k s_k^T + \rho_k^2 B_k s_k y_k^T H_k y_k s_k^T + \rho_k B_k s_k s_k^T$$

$$- \rho_k y_k s_k^T \overset{I}{\cancel{B_k H_k}} + \rho_k^2 y_k s_k^T B_k s_k y_k^T H_k + \rho_k^2 y_k s_k^T \overset{I}{\cancel{B_k H_k}} y_k s_k^T$$

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

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

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

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

$$- \rho_k^3 y_k s_k^T B_k s_k y_k^T H_k y_k s_k^T + \rho_k^4 y_k s_k^T B_k s_k y_k^T s_k y_k^T H_k y_k s_k^T$$

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

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

me rindo