1.1 Gradiente Conjugado

- Demuestre que los vectores no nulos $p_1, p_2, \dots, p_j$ t.q.

$$p_i^T A p_j = 0, \quad \forall i \neq j, \quad \text{donde}$$

$A \in \mathbb{R}^{n \times n}$ una matriz simétrica positiva definida, entonces los vectores son linealmente independientes

Demos →

Sin pérdida de generalidad supongamos que p_n es combinación lineal de los $n-1$ vectores anteriores i.e.

$$p_n = k_1 p_1 + k_2 p_2 + \dots + k_{n-1} p_{n-1}$$

$$= \sum_{i=1}^{n-1} k_i p_i$$

$$\Rightarrow p_1^T A p_n = p_1^T A (k_1 p_1 + k_2 p_2 + \dots + k_{n-1} p_{n-1})$$

$$= k_1 p_1^T A p_1 + k_2 p_1^T A p_2 + \dots + k_{n-1} p_1^T A p_{n-1} = 0$$

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

$$\Rightarrow k_1 p_1^T A p_1 = 0$$

y como A es positiva definida $\Rightarrow k_1 = 0$

Repitiendo este proceso para $\{p_2, p_3, \dots, p_{n-1}\}$ encontramos que $k_1 = k_2 = \dots = k_{n-1} = 0$

$\Rightarrow p_n = \vec{0}$! ya que p_n es un vector no nulo

$\therefore \{p_1, p_2, \dots, p_n\}$ son linealmente indep ■

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

Dado que $\{p_1, p_2, \ldots, p_n\}$ son l.i.

$$\Rightarrow \mathbb{R}^n = \operatorname{span}(\{p_1, p_2, \ldots, p_n\})$$

$\Rightarrow x_1 \in \mathbb{R}^n$ $\exists \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$ tales que

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

$$x^* - x_1 = \sum_{i=1}^{n} \sigma_k p_k$$

Dado que $x_{k+1} = x_k + a_k p_k$, si hacemos el análisis iterativo tenemos

$$x_{k+1} = x_1 + a_1 p_1 + a_2 p_2 + \ldots + a_k p_k$$

$$= x_1 + \sum_{i=1}^{k} a_i p_i$$

$$\Rightarrow A(x^* - x_1) = A\left(x^* - x_1 - \sum_{i=1}^{k} a_i p_i\right)$$

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

$$\Rightarrow \sigma_k = \frac{p_k^T (b - A x_0)}{p_k^T A p_k} = \frac{p_k^T r_k}{p_k^T A p_k} = a_k$$

y como depende directamente de los vectores l.i. que son n

$\Rightarrow$ converge a lo más en n pasos ■

1.2 BFGS

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

Demos $\rightarrow$

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

y

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

son inversas entre ellas, donde

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

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

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

y dado que $B_{k+1} s_k = y_k$

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

Por otro lado el último término de B_{k+1}

$$\frac{y_k y_k^T}{y_k^T s_k} = \rho_k y_k y_k^T$$ $\therefore$ sustituyendo el valor

de B_{k+1} en el primer paréntesis obtendríamos

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

Ahora recordemos que $H_k = B_k^{-1}$

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

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

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

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

$$= I$$

$\therefore$ B_{k+1} y H_{k+1} son inversas ■