1. Gradiente Conjugado

1. **P.D** si $p_1, p_2, \dots, p_\ell \neq 0$ satisfacen q

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

y A es simétrica positiva definida entonces los vectores son l.I

Dem

para $\alpha_i \in \mathbb{R}$ debemos probar que:

$$\sum \alpha_i p_i = 0 \Rightarrow \alpha_i = 0 \quad \forall\ i$$

vamos a suponer que NO es l.I $\Rightarrow \exists\ \alpha_j \in \mathbb{R}$ tal que $\alpha_j \neq 0$
donde tenemos que $\sum_{i=1}^{\ell} \alpha_i p_i = 0$

si multiplicamos por la izquierda

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

$$0 = A[\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_\ell p_\ell] = \alpha_1 A p_1 + \alpha_2 A p_2 + \dots + \alpha_\ell A p_\ell$$
$$= \sum_{i=1}^{\ell} \alpha_i A p_i \quad \dots (*)$$

si multiplicamos $(*)$ p_j^T por la izquierda

$$\Rightarrow p_j^T 0 = p_j^T \sum_{i=1}^{\ell} \alpha_i A p_i = p_j^T [\alpha_1 A p_1 + \alpha_2 A p_2 + \dots + \alpha_\ell A p_\ell]$$
$$= [p_j^T \alpha_1 A p_1 + p_j^T \alpha_2 A p_2 + \dots + p_j^T \alpha_\ell A p_\ell]$$
$$= \alpha_1 p_j^T A p_1 + \alpha_2 p_j^T A p_2 + \dots + \alpha_\ell p_j^T A p_\ell$$
$$= \sum \alpha_i p_j^T A p_i$$

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

tenemos que A es positiva definida por lo que $x^T A x > 0 \quad \forall\ x \in \mathbb{R}^\ell$

y como $\exists\ \alpha_j \in \mathbb{R}$ tal que $\alpha_j \neq 0$

$$\Rightarrow \sum \alpha_i p_j^T A p_i = \alpha_1 p_j^T A p_1 + \alpha_2 p_j^T A p_2 + \dots + \alpha_\ell p_j^T A p_\ell$$
$$= \alpha_j \underbrace{p_j^T A p_j}_{>0} \ \overset{?}{=}\ 0 \quad \text{dado que } \sum \alpha_i p_j^T A p_i = 0$$

Por lo tanto $\alpha_j = 0$ con lo cual demostramos que $\{p_1, \dots, p_\ell\}$ son vectores L.I

2. Dado este resultado, ¿Por que el gradiente conjugado converge en a lo más n iteraciones? Por teorema tenemos que

Para cualquier $x_0 \in \mathbb{R}^n$ la sucesion $\{x_n\}$ generada por el algoritmo de gradiente conjugado converge a la solución x^* del sistema lineal $Ax = b$ en a lo más n pasos

nuestro caso tomamos el sistema lineal definido con $\sum_{i=1}^{\ell} \alpha_i p_i$ donde $x_{k+1} = x_k + \alpha_k p_k$

Dado que $\{p_i\}$ son direcciones L.I $\Rightarrow$ podemos escribir la diferencia entre x_0 y x^*

como $x^* - x_0 = \sigma_0 p_0 + \sigma_1 p_1 + \dots + \sigma_{n-1} p_{n-1}$ ... $(*)$

p.a escalar σ_k

si premultiplicamos $(*)$ por $p_k^T A$ y tomando en cuenta que $p_i^T A p_j = 0 \quad \forall\ i \neq j$

obtenemos $p_k^T A (x^* - x_0) = p_k^T A (\sigma_0 p_0 + \sigma_1 p_1 + \dots + \sigma_{n-1} p_{n-1})$

$$= \sigma_0 \cancel{p_k^T A p_0}^{0} + \dots + \sigma_k p_k^T A p_k + \dots + \sigma_{n-1} \cancel{p_k^T A p_{n-1}}^{0} = \sigma_k p_k^T A p_k \Rightarrow \sigma_k = \frac{p_k^T A (x^* - x_0)}{p_k^T A p_k}$$

vamos a probar que σ_k coincide con el tamaño de paso α_k dado por

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

si x_k es generado por el algoritmo de gradiente conjugado dado $x_{k+1} = x_k + \alpha_k p_k$ y $\alpha_k = \frac{-r_k^T p_k}{p_k^T A p_k}$

entonces tenemos que dado que $\{p_i\}$ son l.I $\Rightarrow$ podemos escribir la diferencia entre x_k y x_0 como $x_k - x_0 = \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{k-1} p_{k-1}$

premultiplicando por $p_k^T A$ obtenemos:

$$p_k^T A (x_k - x_0) = \overbrace{p_k^T A(\alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{k-1} p_{k-1})}^{0} = 0$$

para x^* solución del sistema lineal $Ax=b$

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

$$= p_k^T (b - A x_k) \quad \dots (*)$$

como $r_k = A x_k - b \Rightarrow (*) = -p_k^T r_k$

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

y como $\sigma_k = \dfrac{p_k^T A(x^* - x_0)}{p_k^T A p_k}$ $\Rightarrow$ concluimos que σ_k coincide con α_k

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

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

tenemos que $B_{k+1} S_k = y_k$

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

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

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

(*) $= \left(I - \dfrac{B_k S_k S_k^T}{S_k^T B_k S_k} \right) (I - \rho_k y_k S_k^T) + \rho_k y_k S_k^T$

$\underset{(*)}{=} I - \cancel{\dfrac{B_k S_k S_k^T}{S_k^T B_k S_k}} - \cancel{\rho_k y_k S_k^T} + \cancel{\dfrac{B_k S_k S_k^T}{S_k^T B_k S_k}} + \cancel{\rho_k y_k S_k^T}$

$= \underline{\underline{I}}$ //

(*) $H_k = B_k^{-1}$

(*) $\left(\dfrac{B_k S_k S_k^T}{S_k^T B_k S_k} \right) (\rho_k y_k S_k^T) = \rho_k \dfrac{B_k S_k S_k^T}{S_k^T B_k S_k} y_k S_k^T = \dfrac{1}{S_k^T y_k} \dfrac{B_k S_k (S_k^T y_k) S_k^T}{S_k^T B_k S_k}$

$= \dfrac{B_k S_k S_k^T}{S_k^T B_k S_k}$