

## Auxiliar #10 opti DIM

PI Tenemos:

- R dirección de descenso desde  $x_k$
- $\exists f(x_k + \alpha x_k \mid \alpha > 0) \geq M \in \mathbb{R}$
- $0 < C_1 < C_2 < \infty$

Vemos que en particular, si definimos:

$$\phi_k: \mathbb{R}_{++} \rightarrow \mathbb{R}$$
$$\alpha \mapsto f(x_k + \alpha p_k)$$

- Esta función es inferiormente acotada por  $M$ .
- Como  $p_k$  es de descenso ( $\nabla f_k^T p_k \leq 0$ )

$$\nabla f_k^T p_k < 0, \text{ luego como } 0 < C_1 < C_2$$

Tenemos que  $\ell(\alpha) = f(x_k) + \alpha C_1 \nabla f_k^T p_k$   
no es acotada inferiormente

$$\text{Luego } \exists \alpha \neq 0 \text{ tal que } \ell(\alpha) = \phi(\alpha)$$

- Así, sea  $\alpha'$  el mínimo de estos; claramente  $\forall \alpha < \alpha'$

$$(1) \quad f(x_k + \alpha' p_k) \leq f(x_k) + \alpha' C_1 \nabla f_k^T p_k$$

por  $C_1 \in (0, 1)$  "Aplicar" la propiedad.  
i.e.  $f$  decrece más rápido, al menos temporalmente

$$\text{Así } \exists \alpha'' \in (0, \alpha') \text{ tal que } (p_k^T \nabla f_k)$$

$$(2) \quad f(x_k + \alpha'' p_k) = f(x_k) + \alpha'' \nabla f(x_k + \alpha'' p_k)^T p_k$$

Luego como (1) es igualado en  $\alpha = \alpha'$ .

$$\nabla f(x_k + \alpha'' p_k)^T p_k = c_1 \nabla f(x_k)^T p_k > C_2 \nabla f(x_k)^T p_k$$

Pues  $c_1 < c_2$ . Con ello, se prueban las condiciones de Wolfe débiles estrictamente en  $x''$ .

Luego por las hipótesis de continuidad  $\exists \delta > 0$   
t.q. se satisfacen en  $(x'' - \delta, x'' + \delta)$



b) De la misma demostración, notemos que si  
 $c_2 < c_1$ :

- si llegamos a

$$\nabla f(x_k + \alpha'' p_k) p_k = \nabla f(x_k) p_k \cdot c_1$$

no podemos acotar por  $\nabla f(x_k) p_k \cdot c_2$ , por lo que no tenemos (2).

- si llegamos a

$$\nabla f(x_k + \alpha'' p_k) p_k > \nabla f(x_k) p_k \cdot c_2$$

No tenemos garantía de la existencia de  
Algun  $\delta > 0$  tal que

$$f(x_k + \alpha p_k) \leq f(x_k) + \alpha c_1 \nabla f(x_k) p_k$$

Por lo que no podemos llegar a (1)



P2 En efecto, basta ver que:

$$X = B^{-1}(BX)$$

Con lo que:

$$\|X\| = \|B^{-1}(B(x))\| \leq \|B^{-1}\| \|Bx\|$$

Esto por la def. de la norma de los operadores lineales continuos.

Luego, obtenemos que:

$$\frac{\|X\|}{\|B^{-1}\|} \leq \|Bx\|$$

Con lo que se prueba lo pedido.

P3 Recordemos que el algoritmo del gradiente consiste en la siguiente recursión:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

Consideremos entonces la función siguiente:

$$f(x) = \frac{x^T Q x}{2} - b^T x$$

Veremos del Hint que para este caso se considera:

$$\alpha_k = \frac{\nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_k)^T Q \nabla f(x_k)}$$

Con lo que:

$$x_1 = x_0 - \frac{\nabla f(x_0)^T \nabla f(x_0)}{\nabla f(x_0)^T Q \nabla f(x_0)} \cdot \nabla f(x_0)$$

Calculamos  $\nabla f(x_0)$ . Consideremos  $x^*$  óptimo.

En general sabemos:

$$\nabla f(x) = Qx - b$$

Luego

$$\nabla f(x_0) = Qx_0 - b = Qx_0 - Qx^* + Qx^* - b$$

$$= Q(x_0 - x^*) + Qx^* - b$$

$$\leftarrow \lambda(x_0 - x^*) \leftarrow$$

por linealidad con  
vector propio

$$\leftarrow \nabla f(x^*) \leftarrow$$

= 0 pues  
 $x^*$  es  
óptimo.

NOTA:  $\lambda$  no es el  $\nabla f$  asociado al  $\vec{v}$  paralelo a  $(x_0 - x^*)$  sino que:

Si el  $\vec{v}$  es  $\sigma$  tenemos

$$Q\sigma = \beta\sigma \quad \text{y} \quad \sigma = \gamma(x_0 - x^*)$$

obtenemos:

$$\lambda = \gamma\beta.$$

Con esto, vemos que:

- $\nabla f(x_0) = \lambda(x_0 - x^*)$
- $Q\nabla f(x_0) = Q \cdot (\lambda(x_0 - x^*)) = \lambda Q(x_0 - x^*)$   
 $= \lambda^2(x_0 - x^*)$
- $\nabla f(x_0)^T Q \nabla f(x_0) = \nabla f(x_0)^T [\lambda^2(x_0 - x^*)]$   
 $= \lambda^2 [\lambda(x_0 - x^*)^T] (x_0 - x^*) =$   
 $\lambda^3 \|x_0 - x^*\|^2$
- $\nabla f(x_0)^T \nabla f(x_0) = \lambda(x_0 - x^*)^T \lambda(x_0 - x^*)$   
 $= \lambda^2 (x_0 - x^*)^T (x_0 - x^*) = \lambda^2 \|x_0 - x^*\|^2$

Luego podemos computar  $x_1$ :

$$x_1 = x_0 - \frac{\lambda^2 \|x_0 - x^*\|^2}{\lambda^3 \|x_0 - x^*\|^2} \cdot \lambda(x_0 - x^*)$$
$$= x_0 - \frac{1}{\lambda} (\lambda(x_0 - x^*)) = x^*$$

Más aún, tenemos que podemos encontrar explícitamente el  $x^*$  pues, como estamos en caso irrestricto:

$$\nabla f(x^*) = 0 \quad \text{por (CUPD)}.$$

Luego:

$$[Qx^* = b] \quad \text{y como } Q \text{ es def. positiva} \\ \text{es invertible (teo. visto en} \\ \text{lineal.)}$$

Obtenemos

$$x^* = Q^{-1}b$$