

Tema 5. Optimització No Lineal amb Restriccions

Grau en Matemàtiques

Programació Matemàtica

Jordi Castro Javier Heredia Josep Homs

Departament d'Estadística i Investigació Operativa
Facultat de Matemàtiques i Estadística
Universitat Politècnica de Catalunya
Barcelona



Esquema del tema

- 1 Motivació de les condicions d'optimalitat
- 2 Condicions d'optimalitat
- 3 Exemples
- 4 Anàlisi de sensibilitat
- 5 Problemes convexos
- 6 Problemes lineals i quadràtics
- 7 Bibliografia

Problema d'optimització no lineal amb restriccions

- Formulació general amb restriccions d'igualtat i/o desigualtat:

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

- ▶ $f : \mathbb{R}^n \rightarrow \mathbb{R}$ és la funció objectiu (suposem "suau", $f \in \mathcal{C}^2$).
- ▶ $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (suposem també suaus, $h_i \in \mathcal{C}^2, i = 1, \dots, m$).
- ▶ $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, (suposem també suaus, $g_j \in \mathcal{C}^2, j = 1, \dots, p$).

Concepte de restriccions actives

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h_i(x) = 0 \quad i = 1, \dots, m \\ & g_j(x) \leq 0 \quad j = 1, \dots, p \end{array}$$

Definició

La restricció $g_j, j = 1, \dots, p$, és activa al punt factible x si $g_j(x) = 0$, i inactiva si $g_j(x) < 0$. Les restriccions $h_i, i = 1, \dots, m$, són sempre actives en un punt factible. El conjunt d'índexs de restriccions de desigualtat actives al punt x el denotarem per

$$\mathcal{A}(x) = \{j \in \{1, \dots, p\} : g_j(x) = 0\}.$$

Exemple 1: Una restricció d'igualtat I

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & x_1^2 + x_2^2 - 2 = 0 \end{array}$$

Tenim que:

$$f(x_1, x_2) = x_1 + x_2$$

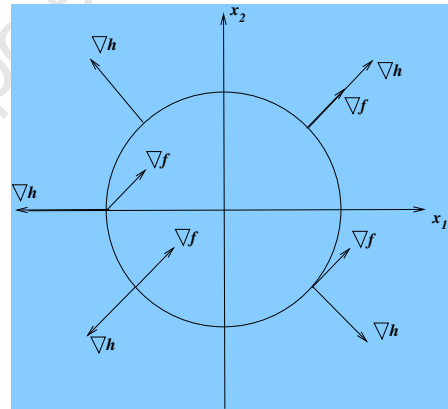
$$h(x_1, x_2) = x_1^2 + x_2^2 - 2$$

$$\nabla f(x_1, x_2) = (1, 1)^T$$

$$\nabla h(x_1, x_2) = (2x_1, 2x_2)^T$$

El mínim és $x^* = (-1, -1)^T$, i

$$\nabla h(x^*) = (-2, -2)^T.$$



- Veurem que en qualsevol punt que no sigui x^* podem moure'ns per la regió factible i disminuir f .

Exemple 1: Una restricció d'igualtat II

- El punt x^* verifica que $\nabla f(x^*)$ és paral·lel a $\nabla h(x^*)$:

$$\nabla f(x^*) = -\lambda^* \nabla h(x^*) \Leftrightarrow \nabla f(x^*) + \lambda^* \nabla h(x^*) = 0.$$

Al nostre cas

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda^* \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ on } \lambda^* = 1/2$$

- La condició anterior es deriva usant sèries de Taylor de primer ordre. Suposem estem a x factible ($h(x) = 0$), i ens movem al llarg de la direcció d :

Exemple 1: Una restricció d'igualtat III

- ▶ Mantenir factibilitat: $h(x + d) = 0$:

$$0 = h(x + d) \approx h(x) + \nabla h(x)^T d = \nabla h(x)^T d.$$

La direcció manté factibilitat, fins al primer ordre, si

$$\nabla h(x)^T d = 0.$$

- ▶ Millorar f : $f(x + d) < f(x)$:

$$f(x) > f(x + d) \approx f(x) + \nabla f(x)^T d.$$

La direcció millora f (és de descens), fins al primer ordre, si

$$\nabla f(x)^T d < 0.$$

Exemple 1: Una restricció d'igualtat IV

- ▶ Una condició necessària d'optimalitat d'un punt és que no hi hagi cap direcció en aquest punt que verifiqui les dues condicions anteriors. Equival a dir que si $\nabla h(x)^T d = 0$ llavors $\nabla f(x)^T d = 0$. Altrament tenim que o bé $\nabla f(x)^T d < 0$ (és de descens), o bé $\nabla f(x)^T d > 0$ (i llavors $-d$ és factible i de descens). Per tant

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0.$$

- Definint la funció Lagrangiana

$$L(x, \lambda) = f(x) + \lambda h(x)$$

la condició $\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$ és equivalent a dir que en la solució x^* hi ha un *multiplicador de Lagrange* λ^* tal que

$$\nabla_x L(x^*, \lambda^*) = 0$$

Exemple 1: Una restricció d'igualtat V

- La condició anterior és necessària, no suficient. P. ex., el punt $x = (1, 1)^T$ la satisfà per a $\lambda = -1/2$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1/2) \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

però no és un mínim (és un màxim).

- I (a diferència de restriccions de desigualtat) no podem imposar un signe a λ . P. ex., si definim $h(x) = 2 - x_1^2 - x_2^2$, tenim el mateix problema (mateixos f , Ω i x^*), però el multiplicador de Lagrange λ^* canvia de signe.

Exemple 2: Una restricció de desigualtat I

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & x_1^2 + x_2^2 - 2 \leq 0 \end{array}$$

Tenim que:

$$f(x_1, x_2) = x_1 + x_2$$

$$g(x_1, x_2) = x_1^2 + x_2^2 - 2$$

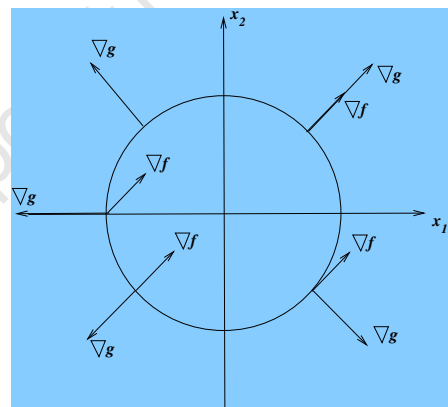
$$\nabla f(x_1, x_2) = (1, 1)^T$$

$$\nabla g(x_1, x_2) = (2x_1, 2x_2)^T$$

Com a l'exemple 1, el mínim és $x^* = (-1, -1)^T$ i es verifica $\nabla f(x^*) + \mu^* \nabla g(x^*) = 0$ per a x^* i $\mu^* = 1/2$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1/2 \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Però ara el signe de μ^* és rellevant:



Exemple 2: Una restricció de desigualtat II

- Si un punt x^* no és òptim podem tenir direcció factible i de descens:

- ▶ Mantenir factibilitat: $g(x + d) \leq 0$:

$$0 \geq g(x + d) \approx g(x) + \nabla g(x)^T d.$$

La direcció manté factibilitat, fins al primer ordre, si

$$g(x) + \nabla g(x)^T d \leq 0.$$

- ▶ Millorar f : $f(x + d) < f(x)$:

$$f(x) > f(x + d) \approx f(x) + \nabla f(x)^T d.$$

La direcció millora f (és de descens), fins al primer ordre, si

$$\nabla f(x)^T d < 0.$$

Exemple 2: Una restricció de desigualtat III

- Ara considerem dos casos, segons $g(x) \leq 0$ sigui o no activa:

- ▶ **Inactiva:** $g(x) < 0$ (el punt es troba a l'interior del cercle). Llavors qualsevol direcció (sempre que tingui una longitud suficientment petita) satisfà $g(x) + \nabla g(x)^T d \leq 0$. I si $\nabla f(x) \neq 0$ sempre podem trobar una direcció tal que $\nabla f(x)^T d < 0$, per exemple,

$$d = g(x) \frac{\nabla f(x)}{\|\nabla g(x)\| \|\nabla f(x)\|}.$$

(suposem $\|\nabla g(x)\| \neq 0$, altrament $d = -\nabla f(x)$ ja serveix).

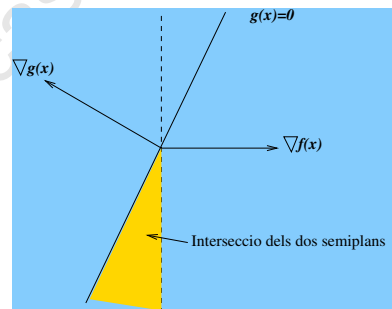
La única situació en que no existeix tal direcció és quan $\nabla f(x) = 0$.

Exemple 2: Una restricció de desigualtat IV

- **Activa:** $g(x) = 0$ (el punt es troba a la frontera de Ω , la circumferència). Les condicions de direcció factible i de descens, fins al primer ordre, són ara:

$$\begin{aligned}\nabla g(x)^T d &\leq 0 &\rightarrow &\text{defineix un semiplà tancat} \\ \nabla f(x)^T d &< 0 &\rightarrow &\text{defineix un semiplà obert}\end{aligned}$$

Si en el punt x hi ha una direcció d que pertany a la intersecció dels dos semiplans, llavors x no pot ser solució. Gràficament:

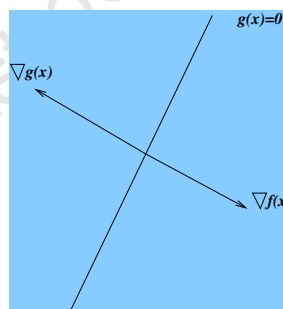


Exemple 2: Una restricció de desigualtat V

Per a que no hi hagi una d en la intersecció dels dos semiplans ha de passar que

$$\begin{aligned}\nabla f(x) + \mu \nabla g(x) &= 0 \\ \mu &\geq 0.\end{aligned}$$

Gràficament:

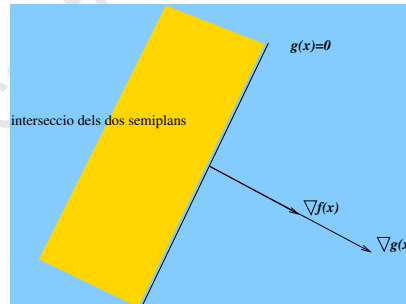


Exemple 2: Una restricció de desigualtat VI

El signe de μ és rellevant. Si es verifica:

$$\begin{aligned}\nabla f(x) + \mu \nabla g(x) &= 0 \\ \mu &< 0\end{aligned}$$

tenim moltes direccions factibles i de descens, fins al primer ordre, en el punt x i aquest no pot ser solució:



Exemple 2: Una restricció de desigualtat VII

- Els dos casos anteriors es resumeixen en les condicions següents:

$$\begin{aligned}\nabla_x L(x^*, \mu^*) &= \nabla f(x^*) + \mu^* \nabla g(x^*) = 0 \\ \mu^* &\geq 0 \\ \mu^* g(x^*) &= 0 \text{ (condició de complementarietat)}\end{aligned}$$

- ▶ Si $g(x^*) < 0$ (inactiva) llavors $\mu^* = 0$ i $\nabla f(x^*) + \mu^* \nabla g(x^*) = \nabla f(x^*) = 0$. Mateixa condició que en problemes sense restriccions.
- ▶ Si $g(x^*) = 0$ (activa) llavors $\mu^* \geq 0$, i junt amb $\nabla_x L(x^*, \mu^*) = 0$ garantim que no hi ha direcció factible de descens.

Exemple 3: Dues restriccions de desigualtat I

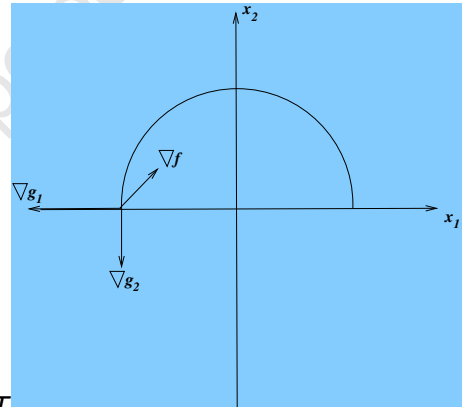
$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.a} \quad & x_1^2 + x_2^2 - 2 \leq 0 \\ & -x_2 \leq 0 \quad (\equiv x_2 \geq 0) \end{aligned}$$

Tenim que:

$$f(x) = x_1 + x_2 \quad \nabla f(x) = (1, 1)^T$$

$$g_1(x) = x_1^2 + x_2^2 - 2 \quad g_2(x) = -x_2$$

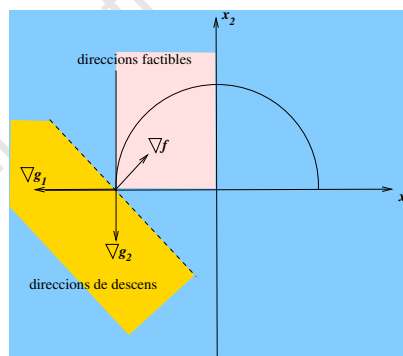
$$\nabla g_1(x) = (2x_1, 2x_2)^T \quad \nabla g_2(x) = (0, -1)^T$$



El mínim és $x^* = (-\sqrt{2}, 0)^T$, i les dues restriccions són actives en aquest punt.

Exemple 3: Dues restriccions de desigualtat II

- Un punt no és solució si hi ha una direcció factible i de descens, fins el primer ordre:
 - ▶ Mantenir factibilitat: $g_j(x) + \nabla g_j(x)^T d \leq 0 \quad j = 1, 2$
 - ▶ Descens: $\nabla f(x)^T d < 0$.
- Al punt $x^* = (-\sqrt{2}, 0)^T$ no hi ha cap direcció que verifiqui factibilitat i descens:



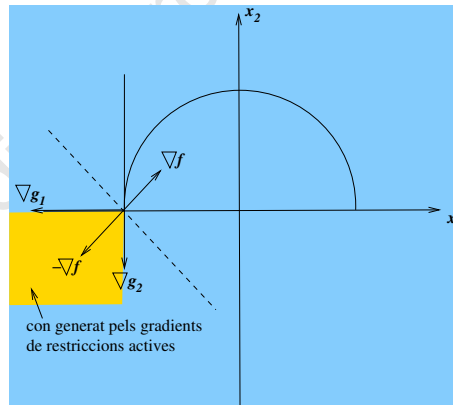
Exemple 3: Dues restriccions de desigualtat III

- Geomètricament, la condició anterior equival a dir que $-\nabla f(x^*) \in C$, on C és el con format pels gradients de les restriccions actives a x^*

$$C = \{v \in \mathbb{R}^n : v = \sum_{j \in \mathcal{A}(x^*)} \mu_j \nabla g_j(x^*), \mu_j \geq 0\}$$

Definició

Recordatori definició de con: $C \subseteq \mathbb{R}^n$ és un con si $\forall x \in C$ es té que $\alpha x \in C$, on $\alpha \geq 0, \alpha \in \mathbb{R}$.



Exemple 3: Dues restriccions de desigualtat IV

- Les condicions anteriors són:

$$-\nabla f(x^*) = \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*), \quad \mu_1^* \geq 0, \mu_2^* \geq 0$$

- Usant la Lagrangiana $L(x, \mu) = f(x) + \mu_1 g_1(x) + \mu_2 g_2(x)$, podem estendre les condicions amb una restricció de desigualtat:

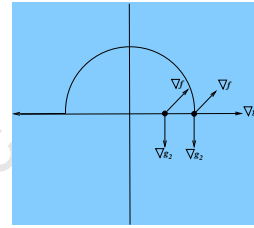
$$\begin{aligned} \nabla_x L(x^*, \mu^*) &= \nabla f(x^*) + \mu_1^* \nabla g_1(x^*) + \mu_2^* \nabla g_2(x^*) = 0 \\ \mu^* &= (\mu_1^*, \mu_2^*)^T \geq 0 \\ \mu_1^* g_1(x^*) &= 0, \mu_2^* g_2(x^*) = 0 \text{ (condició de complementarietat)} \end{aligned}$$

- Al punt $x^* = (-\sqrt{2}, 0)^T$ es té que $\nabla g_1(x) = (-2\sqrt{2}, 0)^T$ i $\nabla g_2(x) = (0, -1)^T$, i per complementarietat $\mu_1^* \geq 0$ i $\mu_2^* \geq 0$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_1^* \begin{pmatrix} -2\sqrt{2} \\ 0 \end{pmatrix} + \mu_2^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \mu^* = (1/(2\sqrt{2}), 1) \geq 0.$$

Exemple 3: Dues restriccions de desigualtat V

Examinem dos punts que no són mínim:



- $x = (\sqrt{2}, 0)^T$. Dues restriccions actives. $-\nabla f(x) \notin$ con generat pels gradients de les restriccions actives: hi ha direccions de descens i factibles, p.ex., $d = (0, -1)^T$. La condició $\nabla_x L(x, \mu) = 0, \mu \geq 0$ falla en aquest punt:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_1^* \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix} + \mu_2^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \mu^* = (-1/(2\sqrt{2}), 1) \not\geq 0.$$

- $x = (1, 0)^T$. $g_1(x) < 0$ (inactiva), $g_2(x) = 0$ (activa). $-\nabla f(x) \notin$ con generat per $\nabla g_1(x)$: hi ha direccions de descens i factibles, p.ex., $d = (-1/2, 1/4)^T$. La condició $\nabla_x L(x, \mu) = 0, \mu \geq 0$ falla en aquest punt:

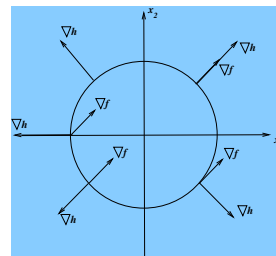
$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_2^* \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \nexists \mu_2^*.$$

Condicions de segon ordre I

- Si per les condicions de primer ordre hem usat ∇f , ∇h i ∇g , per les de segon ordre caldran $\nabla^2 f$, $\nabla^2 h$ i $\nabla^2 g$.
- Considerem l'exemple 1 anterior per derivar condicions de segon ordre:

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & x_1^2 + x_2^2 - 2 = 0 \end{array}$$

El mínim és $x^* = (-1, -1)^T$.



- La derivació és:

- 1 $L(x^*, \lambda^*) = f(x^*) + \lambda^* h(x^*)$ i $h(x^*) = 0$ llavors $L(x^*, \lambda^*) = f(x^*)$.

- 2 Condicions de primer ordre: $\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$.

Condicions de segon ordre II

- 3 Considerem direcció factible a x^* d : $h(x^* + d) = 0$ i $\nabla h(x^*)^T d = 0$.
 Llavors:

$$L(x^* + d, \lambda^*) = f(x^* + d) + \lambda^* h(x^* + d) = f(x^* + d).$$

- 4 Expansió de la sèrie de Taylor de $L(x^* + d, \lambda^*)$ en x^* :

$$\begin{aligned} f(x^* + d) &= L(x^* + d, \lambda^*) \\ &= L(x^*, \lambda^*) + \nabla_x L(x^*, \lambda^*)^T d + \frac{1}{2} d^T \nabla_{xx}^2 L(x^*, \lambda^*) d + o(\|d\|_2^2) \\ &= f(x^*) + \frac{1}{2} d^T \nabla_{xx}^2 L(x^*, \lambda^*) d + o(\|d\|_2^2) \end{aligned}$$

- 5 Usant l'expressió anterior i fent tendir $\|d\| \rightarrow 0$ llavors:

- ★ Si x^* és mínim, és necessari que

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d \geq 0 \quad \text{per a tot } d : \nabla h(x^*)^T d = 0.$$

- ★ És suficient per a que x^* sigui mínim que

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d > 0 \quad \text{per a tot } d : \nabla h(x^*)^T d = 0.$$

Condicions de segon ordre III

- A l'exemple 1 tenim:

$$\nabla f(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \nabla h(x^*) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \quad \lambda^* = \frac{1}{2}$$

$$\nabla^2 f(x^*) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \nabla^2 h(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\nabla_{xx}^2 L(x^*, \lambda^*) = \nabla^2 f(x^*) + \lambda^* \nabla^2 h(x^*) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

que és definida positiva per a tota d (no només pels $d : \nabla h(x^*)^T d = 0$).

- El punt estacionari $x = (1, 1)^T$ no satisfà la condició. En aquest punt $\lambda = -1/2$ i llavors

$$\nabla_{xx}^2 L(x, \lambda) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

que és definida negativa (això indica que el punt és màxim).

Caracterització de les direccions factibles I

- Considerem només restriccions actives. Si eliminem les inactives i considerem

$$\begin{aligned} h_i(x) &= 0 & i &= 1, \dots, m \\ g_j(x) &= 0 & j &\in \mathcal{A}(x^*) \end{aligned}$$

obtenim la mateixa solució x^* .

- Per tant podem considerar que la regió factible està definida per

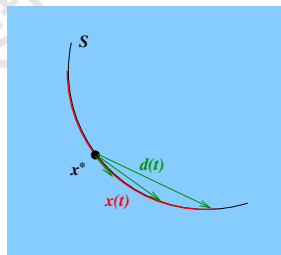
$$c(x) = 0 \rightarrow c_i(x) = 0 \quad i = 1, \dots, m' \quad (m' = m + |\mathcal{A}(x^*)|)$$

- $c(x) = 0$ defineix una superfície $S \subseteq \mathbb{R}^n$.
- Hem de caracteritzar les direccions factibles a S en x^* . Dues formes de fer-ho:

Caracterització de les direccions factibles II

1 Caracterització usant corbes factibles i diferenciables.

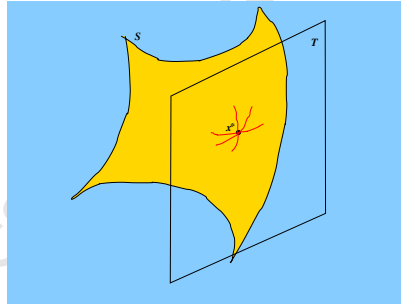
- Corba factible i diferenciable dos cops: família de punts $x(t) = (x_1(t), \dots, x_n(t))^T \in S$, $t \in \mathbb{R}$, $a \leq t \leq b$, tal que $\exists x'(t)$ i $x''(t)$.
- Si passa per x^* llavors $\exists a \leq t^* \leq b : x(t^*) = x^*$.
- $c(x(t)) = 0$ per a tot $a \leq t \leq b$.
- Amb $x(t)$ podem obtenir direccions factibles:



- En el límit la direcció factible d és la tangent a S en el punt x^* : $x'(t^*)$.

Caracterització de les direccions factibles III

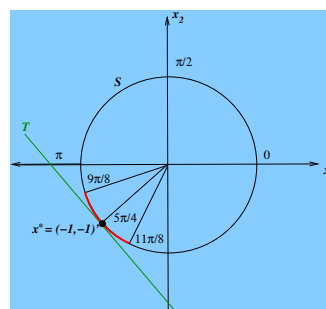
- El conjunt de direccions factibles equival al conjunt de derivades en x^* de totes les possibles corbes factibles i diferenciables que passen per x^* . Això s'anomena **pla tangent T en x^*** , i és un subespai de \mathbb{R}^n .



- Exemple.** La solució d'aquest problema era $x^* = (1, 1)^T$:

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & x_1^2 + x_2^2 - 2 = 0 \end{array}$$

Caracterització de les direccions factibles IV



La corba factible diferenciable (en vermell) és

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \sqrt{2} \cos t \\ \sqrt{2} \sin t \end{pmatrix} \quad 9\pi/8 \leq t \leq 11\pi/8$$

A $t^* = 5\pi/4$, $x(t^*) = x^*$, i $x'(t^*) = d$ pertany al pla tangent T :

$$d = x'(t^*) = \begin{pmatrix} x'_1(t) \\ x'_2(t) \end{pmatrix}_{t=5\pi/4} = \begin{pmatrix} -\sqrt{2} \sin t \\ \sqrt{2} \cos t \end{pmatrix}_{t=5\pi/4} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Caracterització de les direccions factibles V

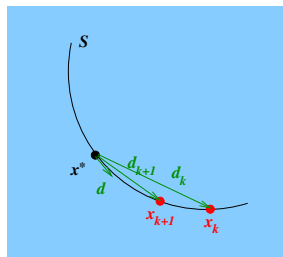
2 Caracterització usant seqüències factibles.

- Donat x^* , una seqüència de punts $\{x_k\}_{k=0}^{\infty}$, $x_k \in \mathbb{R}^n$, és factible si:

- 1 $x_k \neq x^* \quad \forall k$.
- 2 $\lim_{k \rightarrow \infty} x_k = x^*$
- 3 $\exists K : x_k$ és factible $\forall k \geq K$.

- Les direccions factibles són direccions límits de les direccions derivades de les seqüències factibles:

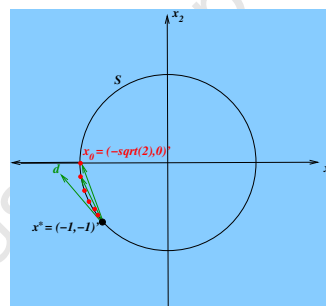
$$\lim_{k \rightarrow \infty} \frac{x_k - x^*}{\|x_k - x^*\|} = \lim_{k \rightarrow \infty} \frac{d_k}{\|d_k\|} = d$$



Caracterització de les direccions factibles VI

- **Exemple.** Usant de nou l'exemple de solució $x^* = (1, 1)^T$:

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & x_1^2 + x_2^2 - 2 = 0 \end{array}$$



La seqüència factible diferenciable (en vermell) és

$$x_k = \begin{pmatrix} -\sqrt{2 - (1 - \frac{1}{k+1})^2} \\ -(1 - \frac{1}{k+1}) \end{pmatrix}$$

Caracterització de les direccions factibles VII

❶ $x_k \neq x^* \quad \forall k.$

❷ $\lim_{k \rightarrow \infty} x_k = \begin{pmatrix} -1 \\ -1 \end{pmatrix} = x^*$

❸ x_k és factible $\forall k \geq 0$:

$$\left(-\sqrt{2 - \left(1 - \frac{1}{k+1}\right)^2}\right)^2 + \left(-\left(1 - \frac{1}{k+1}\right)\right)^2 = 2 - \left(1 - \frac{1}{k+1}\right)^2 + \left(1 - \frac{1}{k+1}\right)^2 = 2$$

El punt x_0 és:

$$x_0 = \begin{pmatrix} -\sqrt{2} \\ 0 \end{pmatrix}.$$

La direcció límit és

$$d = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

com quan hem usat la corba diferenciable factible.

Regularitat de x^* o LICQ

- Treballar amb el pla tangent T o direccions límits de les seqüències factibles no és còmode. Cal una representació en funció de h i g .
- Per poder fer això cal imposar que el punt x^* sigui regular:

Definició

Un punt factible x ($h(x) = 0$, $g(x) \leq 0$) es diu regular si els gradients $\nabla h_i(x)$ $i = 1, \dots, m$ i $\nabla g_j(x)$, $j \in \mathcal{A}(x)$ són linealment independents.

- La condició de que x^* sigui regular també s'anomena "LICQ" (*Linear Independence Constraint Qualification*).

Caracterització del pla tangent

Definició

Definim el subespai de direccions M com

$$\begin{aligned} M &= \{d : \nabla c(x^*)^T d = 0\} \\ &= \{d : \nabla h_i(x^*)^T d = 0 \ i = 1, \dots, m, \ \nabla g_j(x^*)^T d = 0 \ j \in \mathcal{A}(x^*)\} \end{aligned}$$

Teorema

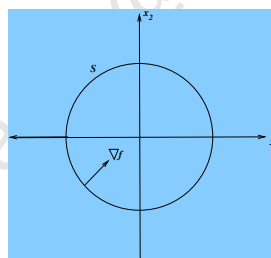
Si x^* és punt regular en la superfície S definida per $c(x) = 0$, llavors el pla tangent T de S en x^* és igual a M ($T = M$).

- El teorema anterior ens permet usar els vectors $d \in M$ com a direccions factibles.
- La regularitat no és una propietat de la superfície S , sino de com es representa amb $c(x) = 0$. El pla tangent T depèn de S , el conjunt M depèn de $c(x) = 0$.

Exemples de no-regularitat de x^* I

- 1 Mateix exemple anterior modificant $h(x) = 0$:

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & (x_1^2 + x_2^2 - 2)^2 = 0 \end{array} \quad \equiv \quad \begin{array}{ll} \min & x_1 + x_2 \\ \text{s.a} & x_1^2 + x_2^2 - 2 = 0 \end{array}$$



La superfície S és la mateixa en els dos problemes, però la seva representació és diferent.

Exemples de no-regularitat de x^* II

$$x^* = (-1, -1)^T \quad h(x) = (x_1^2 + x_2^2 - 2)^2$$

$$\nabla h(x) = \begin{pmatrix} 2(x_1^2 + x_2^2 - 2)2x_1 \\ 2((x_1^2 + x_2^2 - 2)2x_2) \end{pmatrix} \quad \nabla h(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

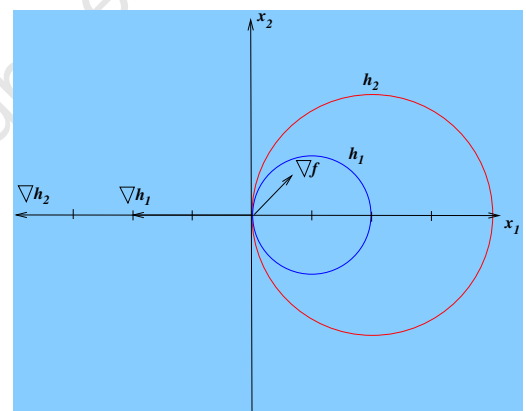
El punt x^* no és regular ($\nabla h(x^*)$ no és linealment independent), i no es verifica $\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda^* \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \nexists \lambda^*$$

Exemples de no-regularitat de x^* III

2 En aquest problema $x^* = (0, 0)^T$ és l'únic punt factible i per tant solució:

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.a} \quad & h_1(x) = (x_1 - 1)^2 + x_2^2 - 1 = 0 \\ & h_2(x) = (x_1 - 2)^2 + x_2^2 - 4 = 0 \end{aligned}$$



$$\nabla h_1(x) = \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix} \quad \nabla h_1(x^*) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\nabla h_2(x^*) = \begin{pmatrix} 2(x_1 - 2) \\ 2x_2 \end{pmatrix} \quad \nabla h_2(x^*) = \begin{pmatrix} -4 \\ 0 \end{pmatrix}$$

Exemples de no-regularitat de x^* IV

$\nabla h_1(x^*)$ i $\nabla h_2(x^*)$ no són linealment independents i no es verifica $\nabla f(x^*) + \sum_{i=1}^2 \lambda_i^* \nabla h_i(x^*) = 0$:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_1^* \begin{pmatrix} -2 \\ 0 \end{pmatrix} + \lambda_2^* \begin{pmatrix} -4 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \nexists \lambda^*$$

Condicions necessàries de primer ordre (KKT)

Teorema (**Condicions KKT**)

Donat

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

i la seva Lagrangiana

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x)$$

suposem que x^* és un mínim local. Si x^* és punt regular (es verifica la LICQ), llavors existeixen uns vector $\lambda^* \in \mathbb{R}^m$ i $\mu^* \in \mathbb{R}^p$ tals que:

- (i) $h(x^*) = 0, g(x^*) \leq 0$
- (ii) $\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*) \lambda^* + \nabla g(x^*) \mu^* = 0$
- (iii) $\mu^* \geq 0$ i $\mu^{*T} g(x^*) = 0$ (si $j \notin \mathcal{A}(x^*)$ llavors $\mu_j^* = 0$)

Condicions necessàries de segon ordre

Teorema (Condicions necessàries de segon ordre)

Donat

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

on $f, h, g \in \mathcal{C}^2$. Si x^* és mínim local, i és punt regular, llavors es satisfan les condicions KKT de primer ordre i a més

$$\nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) = \nabla^2 f(x) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j=1}^p \mu_j^* \nabla^2 g_j(x^*)$$

és semidefinida positiva per a tot

$d \in M = \{d : \nabla h_i(x^*)^T d = 0 \quad i = 1, \dots, m, \nabla g_j(x^*)^T d = 0 \quad j \in \mathcal{A}(x^*)\}$, és a dir

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d \geq 0 \quad \forall d \in M.$$

Condicions suficients de segon ordre

Teorema (Condicions suficients de segon ordre)

Donat

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

on $f, h, g \in \mathcal{C}^2$. Són condicions suficients per a que x^* sigui mínim local **estricta** que es satisfacin les condicions KKT de primer ordre en x^* **(no cal que x^* sigui regular)** i que

$$\nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) = \nabla^2 f(x) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j=1}^p \mu_j^* \nabla^2 g_j(x^*)$$

sigui **definida positiva** per a tot $d \in M' = \{d : \nabla h_i(x^*)^T d = 0 \quad i = 1, \dots, m, \nabla g_j(x^*)^T d = 0 \quad j \in \mathcal{A}(x^*) \cap \{j : \mu_j^* > 0\}\}$, és a dir

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d > 0 \quad \forall d \in M'.$$

Resum de condicions d'optimalitat necessàries

- Donat el problema

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

i la seva Lagrangiana

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$$

- Condicions necessàries** Si x^* punt regular és òptim local llavors han d'existir $\lambda^* \in \mathbb{R}^m$ i $\mu^* \in \mathbb{R}^p$ tals que:

Condicions de primer ordre (KKT)

- (i) $h(x^*) = 0, g(x^*) \leq 0$
- (ii) $\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* = 0$
- (iii) $\mu^* \geq 0$ i $\mu^{*T} g(x^*) = 0$ (si $g_j(x^*)$ és inactiva llavors $\mu_j^* = 0$)

Condicions de segon ordre

- (iv) $d^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d \geq 0$, per a tot $d \in M = \{d : \nabla h(x^*)^T d = 0, \nabla g_j(x^*)^T d = 0 \quad j \in \mathcal{A}(x^*)\}$.

Resum de condicions d'optimalitat suficients

- Condicions suficients d'optimalitat** El punt x^* és òptim local si satisfà:

Condicions de primer ordre (KKT)

- (i) $h(x^*) = 0, g(x^*) \leq 0$
- (ii) $\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \nabla h(x^*)\lambda^* + \nabla g(x^*)\mu^* = 0$
- (iii) $\mu^* \geq 0$ i $\mu^{*T} g(x^*) = 0$ (si $g_j(x^*)$ és inactiva llavors $\mu_j^* = 0$)

Condicions de segon ordre

- (iv) $d^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d > 0$, per a tot $d \in M' = \{d : \nabla h(x^*)^T d = 0, \nabla g_j(x^*)^T d = 0 \quad j \in \mathcal{A}(x^*) \cap \{j : \mu_j^* > 0\}\}$.

- La diferència entre condicions suficients i necessàries és:

- ▶ No cal imposar que x^* sigui punt regular
- ▶ la condició (iv):

$$\begin{array}{ll} d^T \nabla_{xx}^2 L(x^*, \lambda, \mu) d > 0 & d \in M' \quad [\text{suficients}] \\ d^T \nabla_{xx}^2 L(x^*, \lambda, \mu) d \geq 0 & d \in M \quad [\text{necessàries}] \end{array}$$

Condicions d'optimalitat: Exemple 1 I

Problema

Donada un segment de longitud a volem dividir-lo en dues parts, de forma que la suma de les àrees dels quadrats de costat cada una de les parts sigui mínima.

Solució:

- Variables: x_1 i x_2
- Formulació:

$$\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.a} & x_1 + x_2 = a \\ & x_1 \geq 0 \quad x_2 \geq 0 \end{array} \quad \equiv \quad \begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.a} & x_1 + x_2 = a \\ & -x_1 \leq 0 \quad -x_2 \leq 0 \end{array}$$

- Lagrangiana:

$$L(x_1, x_2, \lambda, \mu) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - a) - \mu_1 x_1 - \mu_2 x_2$$

Condicions d'optimalitat: Exemple 1 II

- Condicions necessàries de primer ordre:

$$\begin{array}{ll} (i) & x_1 + x_2 = a, \quad x_1 \geq 0, \quad x_2 \geq 0 \\ (ii.x_1) & \nabla_{x_1} L() = 2x_1 + \lambda - \mu_1 = 0 \\ (ii.x_2) & \nabla_{x_2} L() = 2x_2 + \lambda - \mu_2 = 0 \\ (iii) & \mu_i \geq 0, \mu_i = 0 \text{ si } x_i > 0, i = 1, 2 \end{array}$$

- S'han d'analitzar els 4 casos segons $x_1, x_2 \geq 0$ actives o no:

- ▶ $x_1 = 0, x_2 = 0$. No pot ser, infactibilitat $x_1 + x_2 = a$
- ▶ $x_1 > 0, x_2 = 0$. Llavors $x_1 = a$ i $\mu_1 = 0$. Si solucionem

$$\begin{array}{rcl} 2a + \lambda & = & 0 \\ \lambda - \mu_2 & = & 0 \end{array}$$

tenim que $\mu_2 = \lambda = -2a$, que viola $\mu_2 \geq 0$

- ▶ $x_1 = 0, x_2 > 0$. És simètric a l'anterior
- ▶ $x_1 > 0, x_2 > 0$. Llavors $\mu_1 = \mu_2 = 0$ i hem de solucionar:

$$\begin{array}{rcl} x_1 + x_2 & = & a \\ 2x_1 + \lambda & = & 0 \\ 2x_2 + \lambda & = & 0 \end{array}$$

obtenint $x_1 = x_2 = a/2, \lambda = -a$, candidat a òptim

Condicions d'optimalitat: Exemple 1 III

- Condicions suficients de segon ordre:

$$\nabla_{xx}^2 L() = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

que és definida positiva per a tot d , no només per als $d \in M$.

- Per tant el punt obtingut és mínim local (i com que el problema és convex, també global).

Condicions d'optimalitat: Exemple 2 I

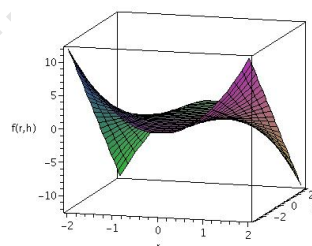
Problema

Volem construir un cilindre metàl·lic de volum màxim, de forma que la seva superfície sigui igual a 6π .

Solució:

- Variables: radi r i alçada h
- Formulació:

$$\begin{array}{ll} \max & \pi r^2 h \equiv -\min f(r, h) = -\pi r^2 h \\ \text{s.a} & c(r, h) = 2\pi rh + 2\pi r^2 - 6\pi = 0 \\ & r \geq 0 \quad h \geq 0 \end{array}$$



Condicions d'optimalitat: Exemple 2 II

- Lagrangiana:

$$L(r, h, \lambda, \mu) = -\pi r^2 h + \lambda(2\pi r h + 2\pi r^2 - 6\pi) - \mu_r r - \mu_h h$$

- Condicions necessàries de primer ordre:

$$\begin{aligned} (i) \quad & r \geq 0, \quad h \geq 0, \quad 2\pi r h + 2\pi r^2 - 6\pi = 0 \\ (ii.r) \quad & \nabla_r L(r, h, \lambda, \mu) = -2\pi r h + \lambda 2\pi h + \lambda 4\pi r - \mu_r = 0 \\ (ii.h) \quad & \nabla_h L(r, h, \lambda, \mu) = -\pi r^2 + \lambda 2\pi r - \mu_h = 0 \\ (iii) \quad & \mu_r r = 0, \quad \mu_h h = 0 \end{aligned}$$

- $r \geq 0$ $h \geq 0$ inactives, sino solució infactible i/o no òptima, i llavors $\mu_r = \mu_h = 0$
- Solucionant (i), (ii.r), (ii.h), i descartant les solucions infactibles ($r < 0, h < 0$) el punt $r = 1, h = 2, \lambda = 1/2$ satisfà les condicions de primer ordre, i és candidat a òptim.

Condicions d'optimalitat: Exemple 2 III

- Comprovem condicions de segon ordre:

$$d^T \nabla_{xx}^2 L(x, \lambda, \mu) d > 0 \quad \forall d : \nabla c(x)^T d = 0$$

(x en aquest cas és (r, h)).

- $\nabla_{xx}^2 L(x, \lambda)$ avaluada a $r = 1, h = 2, \lambda = 1/2$ és:

$$\nabla_{xx}^2 L() = \begin{bmatrix} \nabla_{rr}^2 L() & \nabla_{rh}^2 L() \\ \nabla_{hr}^2 L() & \nabla_{hh}^2 L() \end{bmatrix} = \begin{bmatrix} -2\pi h + 4\pi\lambda & -2\pi r + 2\lambda\pi \\ -2\pi r + 2\lambda\pi & 0 \end{bmatrix} = \begin{bmatrix} -2\pi & -\pi \\ -\pi & 0 \end{bmatrix}$$

- Vectors d tals que $\nabla c(x)^T d = 0$:

$$\nabla c(x) = \nabla_{(r,h)} (2\pi r h + 2\pi r^2 - 6\pi) = [8\pi \quad 2\pi]^T$$

$$[8\pi \quad 2\pi][d_1 \quad d_2]^T = 8\pi d_1 + 2\pi d_2 = 0 \Rightarrow d_2 = -4d_1$$

- $d^T \nabla_{xx}^2 L(x, \lambda, \mu) d > 0$:

$$[d_1 \quad -4d_1] \begin{bmatrix} -2\pi & -\pi \\ -\pi & 0 \end{bmatrix} [d_1 \quad -4d_1]^T = 6\pi d_1^2 > 0 \quad \forall d \neq 0$$

- Es garanteix la condició suficient de segon ordre: $r = 1, h = 2, \lambda = 1/2$ és mínim local.

Exemple 2 en AMPL+MINOS

Model AMPL

```
param pi := 4*atan(1);

var r >= 0, := 0.5; # 0.5 és punt inicial suggerit
var h >= 0, := 0.5; # 0.5 és punt inicial suggerit

minimize fx: -pi*r**2*h;
subject to cx: 2*pi*r*h + 2*pi*r**2 = 6*pi;
```

Sortida

```
ampl: solve;
MINOS 5.5: optimal solution found.
23 iterations, objective -6.283185307
ampl: display r,h;
r = 1
h = 2
ampl: display cx;
cx = -0.5
ampl: display r.rc,h.rc;
r.rc = 0
h.rc = -3.63141e-08
```

- `display $h_i(x)$` mostra $-\lambda_i$ o $-\mu_i$ amb el solver MINOS
- `display x.rc` mostra la $-\mu$ de $x \geq 0$ amb el solver MINOS
- Amb un altre solver els signes de λ i μ poden ser els contraris i coincidir amb els usats a les transparències anteriors

Exemple 2 en AMPL+MINOS: Problemes amb punts estacionaris (falsos òptims)

- Considerem de nou el problema del cilindre de volum màxim amb la restricció de desigualtat i punt inicial per defecte (és 0):

```
param pi := 4*atan(1);
var r >= 0; # per defecte el punt inicial és 0
var h >= 0; # per defecte el punt inicial és 0
maximize fx: pi*r**2*h;
subject to cx: 2*pi*r*h + 2*pi*r**2 <= 6*pi;
```

- No és un problema convex: funció objectiu convexa i maximitzem
- Si fem “solve”, però, obtenim:

```
MINOS 5.5: optimal solution found.
0 iterations, objective 0
```

i els punt òptims retornat són $r = h = 0$, el punt inicial.

- **El solver ha retornat la solució de volum mínim com a màxim!**: el punt inicial és un punt estacionari (verifica les condicions d'optimalitat de primer ordre), que és el que comprova MINOS.

Exemples de falsos òptims amb MAPLE

● Problema exemple de classe:

- ▶ El punt inicial satisfà KKT però no és mínim (és màxim):

```
> NLPSolve(x1 + x2, x1^2 + x2^2 - 2 = 0);
Warning, no iterations performed as initial point satisfies first-order conditions
[2., [x1 = 1.0, x2 = 1.0]]
```

- ▶ Indicant un punt inicial troba la solució:

```
> NLPSolve(x1 + x2, x1^2 + x2^2 - 2 = 0, initialpoint = x1 = 0, x2 = 0);
[-2., [x1 = -1.0, x2 = -1.0]]
```

● Problema del cilindre:

- ▶ El punt inicial satisfà KKT però no és mínim (és màxim):

```
> NLPSolve(-Pi * h * r^2, 2 * Pi * r^2 + 2 * Pi * r * h <= 6 * Pi, assume = nonnegative,
initialpoint = h = 0, r = 0);
Warning, no iterations performed as initial point satisfies first-order conditions
[-0., [h = 0.0, r = 0.0]]
```

- ▶ Usant un altre punt inicial troba la solució: > NLPSolve(-Pi * h * r^2,

```
2 * Pi * r^2 + 2 * Pi * r * h <= 6 * Pi, assume = nonnegative);
[-6.28318530717932244, h = 2.0000000000126468, r = 0.9999999999968173]]
```

Interpretació dels multiplicadors λ , μ

Teorema (Anàlisi de sensibilitat)

Donada la família de problemes

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = \epsilon \quad [h_i(x) = \epsilon_i \quad i = 1, \dots, m] \\ & g(x) \leq \delta \quad [g_j(x) \leq \delta_j \quad j = 1, \dots, p] \end{array}$$

on $f, h, g \in \mathcal{C}^2$. Suposem que per $a \epsilon = \delta = 0$ hi ha una solució x^* , punt regular, i λ^* , μ^* que satisfan les condicions d'optimalitat KKT de primer ordre i les suficients de segon ordre. Suposem també que si $j \in \mathcal{A}(x^*)$ llavors $\mu_j^* > 0$. Llavors per a tot (ϵ, δ) en un entorn de $(0, 0)$ hi ha una solució $x^*(\epsilon, \delta)$ tal que $x^*(0, 0) = x^*$. A més es verifica

$$\begin{aligned} \nabla_{\epsilon} f(x(\epsilon, \delta)) \Big|_{(0,0)} &= -\lambda^* \\ \nabla_{\delta} f(x(\epsilon, \delta)) \Big|_{(0,0)} &= -\mu^*. \end{aligned}$$

Exemple d'anàlisi de sensibilitat

- A l'exemple del cilindre

$$\begin{array}{ll} \min & f(r, h) = -\pi r^2 h \\ \text{s.a} & c(r, h) = 2\pi rh + 2\pi r^2 - 6\pi = 0 \\ & r \geq 0 \quad h \geq 0 \end{array}$$

la solució era $r = 1, h = 2, \lambda = 1/2, f(r, h) = -2\pi$.

- Considerem $c(r, h) = 2\pi rh + 2\pi r^2 - 6\pi = 0.01$
- El nou valor òptim és $f(r(\epsilon), h(\epsilon)) = -2\pi - 0.00500007$, i $\Delta f = f(r(\epsilon), h(\epsilon)) - f(r, h) = -0.00500007$.
- Observem que $\Delta f \approx \epsilon \cdot (-\lambda) = 0.01 \cdot (-1/2) = -0.005$.

Problema estrictament convex

- Donat

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

on f estrictament convexa, g_j convexes, i $h(x) = Ax - b$ funció afí.

- Llavors $\nabla^2 f(x)$ definida positiva, $\nabla^2 g_j(x)$ semidefinides positives, i $\nabla^2 h_i(x) = 0$.
- Per tant

$$\nabla_{xx} L(x^*, \lambda^*, \mu^*) = \nabla^2 f(x) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) + \sum_{j=1}^p \mu_j^* \nabla^2 g_j(x^*)$$

és definida positiva per a tot $d \in \mathbb{R}^n$, no només pels $d \in M$, i qualsevol punt que satisfaci les condicions KKT de primer ordre també satisfarà les de segon ordre i serà mínim global estricte.

Problema convex

Sense imposar convexitat estricta també es garanteix que en un problema convex amb les condicions de primer ordre n'hi ha prou per obtenir el mínim global:

Teorema

Donat

$$\begin{array}{ll} \min & f(x) \\ \text{s.a} & h(x) = 0 \quad [h_i(x) = 0 \quad i = 1, \dots, m] \\ & g(x) \leq 0 \quad [g_j(x) \leq 0 \quad j = 1, \dots, p] \end{array}$$

$f, h, g \in \mathcal{C}^1$, f i g_j convexes, i $h(x) = Ax - b$ funció afí. Si es verifiquen les condicions KKT de primer ordre a x^* , llavors x^* és mínim global.

Condicions d'optimalitat d'un PL

- PL

$$\begin{array}{ll} \min & c^T x \\ \text{s.a} & Ax = b \quad [\lambda] \\ & -x \leq 0 \quad [\mu] \end{array}$$

- Lagrangiana

$$L(x, \lambda, \mu) = c^T x + \lambda^T (b - Ax) - \mu^T x$$

- Condicions KKT

$$\begin{array}{ll} Ax = b, & x \geq 0 \\ c - A^T \lambda - \mu = 0 \\ \mu^T x = 0, & \mu \geq 0 \end{array}$$

- Reescrivim KKT:

$$\begin{array}{ll} Ax = b, & x \geq 0 \\ A^T \lambda + \mu = c, & \mu \geq 0 \\ \mu^T x = 0 \end{array}$$

Factibilitat primal

Factibilitat dual

Complementarietat: $c^T x = b^T \lambda$

Algorisme del simplex: una forma de solucionar KKT

- Condicions KKT:

$$\begin{aligned} Ax &= b \\ A^T \lambda + \mu &= c \\ \mu^T x &= 0 \\ x &\geq 0, \mu \geq 0 \end{aligned}$$

- Al símplex (primal) μ són els costos reduïts i λ les variables duals
- El símplex busca una partició entre bàsiques i no bàsiques, que verifiqui les condicions KKT:

$$\begin{aligned} x &= [x_B \ x_N]^T, \ x_N = 0, \ x_B > 0 \\ A &= [B \ N], \ Ax = Bx_B = b \\ \mu_B &= 0 \Rightarrow B^T \lambda = c_B \\ \mu_N &= c_N - N^T \lambda \text{ [definició cost reduït símplex]} \\ \mu_N &\geq 0 \Rightarrow \text{Condicció d'optimalitat símplex} \end{aligned}$$

- La complementarietat $\mu^T x = 0$ garantida a tota iteració del símplex: $\mu_B^T x_B = 0, \mu_N^T x_N = 0$
- Es viola (al símplex primal) la condició $\mu_N \geq 0$. Es va iterant fins aconseguir-la.

Condicions d'optimalitat d'un PQ

- PQ

$$\begin{aligned} \min \quad & c^T x + \frac{1}{2} x^T Q x \\ \text{s.a} \quad & Ax = b \quad [\lambda] \\ & -x \leq 0 \quad [\mu] \end{aligned}$$

- Lagrangiana

$$L(x, \lambda, \mu) = c^T x + \frac{1}{2} x^T Q x + \lambda^T (b - Ax) - \mu^T x$$




- Condicions KKT

$$\begin{aligned} Ax &= b, \quad x \geq 0 \\ c + Qx - A^T \lambda - \mu &= 0 \\ \mu^T x &= 0, \quad \mu \geq 0 \end{aligned}$$

- Reescriu KKT:

$$\begin{aligned} Ax &= b, \quad x \geq 0 && \text{Factibilitat primal} \\ A^T \lambda - Qx + \mu &= c, \quad \mu \geq 0 && \text{Factibilitat dual} \\ \mu^T x &= 0 && \text{Complementarietat} \end{aligned}$$

Bibliografia

-  D.P. Bertsekas, *Nonlinear Programming, 2nd Ed.*, 1999, Athena Scientific, Belmont, USA.
-  D.G. Luenberger, Y. Ye., *Linear and Nonlinear Programming, 3rd Ed.*, 2008, Springer, New York, USA.
-  J. Nocedal, S.J. Wright, *Numerical Optimization, 2nd Ed.*, 2006, Springer, New York, USA.