Problemes d'Optimització No Lineal

PROGRAMACIÓ MATEMÀTICA – GRAU EN MATEMÀTIQUES – UPC

---

1. Problemes de conceptes bàsics

1.1. Raoneu si els següents conjunts són convexos o no:

i. $C = \{x \in \mathbb{R}^n \mid Ax = b,\ A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m\}$

ii. $C = \{x \in \mathbb{R}^n \mid Ax \leq b,\ A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m\}$

iii. $C = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$

1.2. Determineu la convexitat de les següents funcions sobre punts $x \in \mathbb{R}^2_+$:

i. $f(x_1, x_2) = 4x_1^2 - 5x_1x_2 + x_2^2$

ii. $f(x_1, x_2) = 2e^{x_1} + x_2^2 x_1 + 1$

iii. $f(x_1, x_2) = 2\ln x_1x_2 - 3x_1^2 + 2x_1x_2 - x_2^2$

1.3. Demostreu que la intersecció de conjunts convexos és un conjunt convex, però no la unió.

1.4. Demostreu que:

i. Si f_1 i f_2 són funcions convexes, $f_1 + f_2$ és també una funció convexa.

ii. Si f és una funció convexa, af és convexa $\forall a \geq 0$.

1.5. Donats el vector $a \in \mathbb{R}^n$ i la matriu simètrica $A \in \mathbb{R}^{n\times n}$, calculeu el gradient i la Hessiana de les funcions $f_1(x) = a^T x$ i $f_2(x) = x^T Ax$.

1.6. Demostreu que si $f : D \in \mathbb{R}^n \to \mathbb{R}$ és funció convexa en D, llavors el conjunt $\Omega_c = \{x \in D : f(x) \leq c\}$ és convex per a tot $c \in \mathbb{R}$.

1.7. Demostreu que en un problema d'optimització convex el conjunt de solucions és convex.

1.8. Demostreu que tot mínim local aïllat és mínim local estricte.

1.9. Sabem que f és convexa si i només si $\nabla^2 f$ és semidefinida positiva. Però per funcions estrictaments convexes aquesta doble implicació no la tenim: si $\nabla^2 f$ és definida positiva llavors f és estrictament convexa; però si f és estrictament convexa, $\nabla^2 f$ no ha de ser definida positiva. Comproveu que això succeeix amb la funció $f(x) = x^4$: és estrictament convexa a tot $x \in \mathbb{R}$, però no és definida positiva a tot $x \in \mathbb{R}$.

1.10. Sabem que si $g(x)$ és una funció convexa, el conjunt $\{x \mid g(x) \leq c\}$ és convex $\forall c \in \mathbb{R}$. Creieu que això també és cert si la restricció és $g(x) = c$? Demostreu-ho en cas afirmatiu o doneu un contraexemple en cas contrari.

1.11. Si apliquéssim un determinat algorisme d'optimització al problema:

$$\begin{array}{ll} \min & x^4 + 3y^2 - 2xy - 4x + 2y \\ \text{s. a} & \ln(x+y) \geq 0 \\ & x^2 + y^2 \leq 20 \\ & x \geq 1/2 \quad y \geq 1/2 \end{array}$$

podríem garantir que el punt solució obtingut correspon a un mínim global?

1.12. Siguin $f : \mathbb{R}^m \to \mathbb{R}$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Definim la funció $g : \mathbb{R}^n \to \mathbb{R}$ com $g(x) = f(Ax + b)$ (composició de la funció f amb la transformació lineal $Ax + b$). Proveu que si f és funció convexa, llavors g també és funció convexa.

Aquest resultat és útil per veure immediatament que funcions com $f(x_1, x_2) = (x_1 - 3x_2)^4 + 3e^{2x_1+5x_2+1}$ són convexes, per ser suma de dues funcions, cada una d'elles convexes per ser la composició d'una funció convexa amb una transformació lineal.

1.13. Siguin $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ un conjunt de m funcions convexes. Proveu que la funció "màxim punt a punt" definida com $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ és convexa.

Seria convexa la funció "mínim punt a punt" $g(x) = \min\{f_1(x), \ldots, f_m(x)\}\}$? Proveu que no gràficament.

1.14. Sigui $f : \mathbb{R}^n \to \mathbb{R}$ funció convexa, i $x, d \in \mathbb{R}^n$. Proveu que la funció $g : \mathbb{R} \to \mathbb{R}$ definida com $g(\alpha) = f(x + \alpha d)$ (restricció de f a una línia que passa per x amb direcció d) és convexa.

1.15. Si $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ són convexes, digueu si les següents operacions donarien o no una funció convexa (demostreu que sí, o en cas contrari trobeu un contraexemple):

i. $f_1 - f_2$.
ii. $f_1 \cdot f_2$.
iii. $\frac{f_1}{f_2}$.

1.16. Sigui $f : \mathbb{R}^n \to \mathbb{R}$ funció convexa, i $g : \mathbb{R} \to \mathbb{R}$ funció convexa i no-decreixent. Proveu que la composició $(g \circ f)$ és una funció convexa.

2. Problemes d'optimització sense restriccions

2.1. Considereu el problema

$$\min f(x_1, x_2) = 5x_1^2 + 5x_2^2 - x_1x_2 - 11x_1 + 11x_2 + 11$$

i. Trobeu un punt que satisfaci les condicions necessàries de primer ordre.
ii. Proveu que aquest punt és un mínim global.

2.2. Considereu el següent problema de minimització:

$$\min x_1^2 - x_1x_2 + 2x_2^2 - 2x_1 + e^{x_1 - x_2}$$

i. Escriviu les condicions necessàries de primer ordre. Són aquestes condicions suficients? Per què?
ii. És $x_a = (0, 0)^T$ una solució òptima? Si no ho és, identifiqueu una direcció d al llarg de la qual la funció decreix.
iii. Minimitzeu la funció començant des del punt x_a al llarg de la direcció d obtinguda a (b).
iv. Menyspreant el darrer terme de la funció objectiu, utilitzeu les condicions necessàries i suficients per resoldre el problema.

2.3. Considereu el problema de minimització de la funció $f(x + \alpha d)$, on $\alpha \in \mathbb{R}_{\geq 0}$. Proveu que una condició necessària per a α^* és $d^T \nabla f(y) = 0$, on $y = x + \alpha^* d$. Sota quines hipòtesis és aquesta condició suficient?

2.4. Considereu el problema de minimització de la funció $f(x + \alpha d)$ sotmés a $x + \alpha d \in S$ i $\alpha \geq 0$, on S és un conjunt convex i compacte i f és una funció convexa. Suposem, a més, que d és una direcció de descens. Proveu que la solució òptima ve donada per $\alpha^* = \min(\alpha_1, \alpha_2)$, on α_1 satisfà $d^T \nabla f(x + \alpha_1 d) = 0$ i $\alpha_2 = \max\{\alpha \mid x + \alpha d \in S\}$.

2.5. Trobeu el màxim de la funció $f(x_1, x_2) = \frac{x_1 x_2 - x_1^2}{3} - x_2^2$ mitjançant el mètode de Newton i prenent com a punt inicial $x_0 = (1, 1)^T$.

2.6. Un dels incovenients del mètode de Newton és que no té convergència global. Comproveu aquest fet a partir de la funció $f(x) = 2/3|x|^{3/2}$ i prenent com a punt inicial $x_0 = 1$ (aquesta funció té un mínim global en $x^* = 0$).

2.7. Considereu la funció $f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + x_2^2$ i el punt $x_a = (1, 1)^T$.

i. Quina és la direcció que proporciona el mètode del gradient en x_a?
ii. És de descens la direcció $(1, -1)^T$?

2.8. Considereu el problema de minimització de la funció $f(x_1, x_2) = 2x_1^4 - 4x_1x_2 + x_2^2 + 5x_2$. Calculeu els dos primers termes de la seqüència

i. del mètode del gradient
ii. del mètode de Newton

prenent com a punt inicial $x_0 = (0, 0)^T$.

2.9. Torneu a repetir l'exercici 2.8 però considerant en aquest cas la funció $f(x_1, x_2) = 2x_1^2 - x_1x_2 + x_2^2$ i el punt inicial $x_0 = (1, 4)^T$.

2.10. Donada la funció quadràtica $f(x_1, x_2) = x_1^2 + 2x_2^2 - 2x_1x_2 + 2x_2$:

i. Determineu el seu mínim, si existeix.
ii. Efectueu un pas del mètode del gradient a partir del punt $x = (0, 0)^T$.

2.11. Trobeu el mínim, si existeix, de la funció

$$f(x_1, x_2, x_3) = 50x_1^2 + 3x_1x_3 + 4x_2^2 + 15x_3^2 - 6x_1 + 8x_2 - 60x_3$$

i. de forma analítica.
ii. a partir del punt $x = (0, 0, 2)^T$, de forma aproximada, mitjançant dues iteracions del mètode del gradient. Determineu el factor teòric de convergència.

2.12. Proveu que les seqüències $x^k = 2^{-k}$, $x^k = k^{-k}$ i $x^k = 2^{-2^k}$ tenen, respectivament, convergència lineal, superlineal i quadràtica.

2.13. Donada la funció quadràtica convexa $f(x) = \frac{1}{2}x^TQx - b^Tx$, proveu que aplicant el mètode del gradient la longitud de pas òptima ve donada per $\alpha^k = \dfrac{\nabla f(x^k)^T\nabla f(x^k)}{\nabla f(x^k)^TQ\nabla f(x^k)}$.

2.14. Per al cas particular del problema 2.13 (funció quadràtica convexa i longitud pas calculada de forma exacta) comproveu que dues direccions successives del mètode del gradient són perpendiculars (és a dir, comproveu que $\nabla f(x^k)$ i $\nabla f(x^{k+1})$ són perpendiculars).

2.15. Considereu les funcions següents:

$$f_1(x) = x_1^2 + x_2^2 + x_1x_2 - x_1 - 2x_2$$

$$f_2(x) = 5x_1^2 + x_2^2 - 2x_2$$

i. Comproveu que les dues tenen el mateix punt mínim.
ii. Raoneu quina funció tindrà un factor de convergència més lent si apliquem el mètode del gradient.
iii. Realitzeu dues iteracions del mètode del gradient per a cada funció, prenent en ambdos casos com a punt inicial $x_0 = (2, 2)^T$.
iv. Comproveu que les direccions obtingudes (per a cada problema) són perpendiculars.

2.16. Una de les propietats del mètode de Newton és que, quan s'aplica a funcions quadràtiques, troba el punt òptim en una iteració. Això és degut a que es basa en fer una aproximació quadràtica de la funció a minimitzar a cada iteració i trobar aleshores el punt òptim de l'aproximació. Quan la funció és quadràtica, l'aproximació coincideix amb la funció original. Per tant amb una iteració en tenim prou per trobar l'òptim. Comproveu experimentalment aquest fet aplicant el mètode de Newton a les funcions del problema 2.15.

2.17. Demostreu que si $\nabla^2 f(x)$ és definida positiva llavors $(\nabla^2 f(x))^{-1}$ també és definida positiva. (Això ens garanteix que la direcció de Newton $-(\nabla^2 f(x))^{-1}\nabla f(x)$ és de descens.)

2.18. El Teorema de Zoutendijk s'usa per demostrar la convergència global del mètode de Newton si $\cos\theta_k \geq \delta > 0$, per a tot $k \geq 0$. Demostreu que si $\nabla^2 f(x^k)$ és definida positiva per a tot $k \geq 0$, i $\nabla^2 f(x^k)$ té un nombre de condició uniformement afitat, és a dir,

$$\exists M : \quad \|\nabla^2 f(x^k)\|\|(\nabla^2 f(x^k))^{-1}\| \leq M \quad \forall k \geq 0,$$

llavors

$$\cos\theta_k \geq \frac{1}{M}.$$

Haureu d'usar, i per tant prèviament demostrar, la següent propietat: $\|Bx\| \geq \|x\|/\|B^{-1}\|$, per a quasevol vector x i matriu no singular B.

2.19. Considereu la funció $f(x_1, x_2) = x_1^3 - 12x_1x_2 + 8x_2^3$ i el punt $x_a = (-1, 0)^T$.

i. Obteniu la solució òptima del problema aplicant les condicions de primer i segon ordre.

ii. Caracteritzeu les direccions de descens estricte en x_a. És de descens la direcció proporcionada pel mètode de Newton en el punt x_a? Podrieu justificar-ho?

iii. Realitzeu una iteració del mètode de Newton unidimensional en el punt x_a amb $\alpha_0 = 0$ per obtenir un nou punt al llarg de la direcció $d = (1, -1)^T$.

iv. Comproveu si el nou punt obtingut a l'apartat (c) verifica o no les condicions d'Armijo-Wolfe per a $c_1 = 0.1$ i $c_2 = 0.5$.

2.20. Considereu la funció

$$f(x) = a(x_1^2 - x_2^2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + b(x_3^2 - x_4)^2 + c[(x_2 - 1)^2 + (x_4 - 1)^2] + d(x_2 - 1)(x_4 - 1),$$

on $a = 99$, $b = 90$, $c = 10.1$ i $d = 20.1$. Donats el punt $x_a = (0, 0, 0, 0)^T$ i la direcció $d = (1, 1, 1, 1)^T$:

i. comproveu que d és una direcció de descens per a f en x_a.

ii. satisfà el punt obtingut a l'apartat (b) les condicions d'Armijo-Wolfe?

2.21. Donada la funció $f(x_1, x_2) = x_1^4 + x_2^2$ i el punt $x_a = (1, 1)^T$, considereu com a direcció de moviment l'obtinguda amb el mètode de Newton.

i. Satisfà les condicions d'Armijo-Wolfe el punt obtingut amb aquesta direcció?

ii. Quina longitud de pas hi hauria que prendre (exploració lineal) perquè es satisfacin?

iii. Quin és l'òptim analític en aquesta direcció?

2.22. [Problema examen juliol 2013]

i. Donada la funció quadràtica $f(x) = \frac{1}{2}x^T Q x - b^T x$, Q definida positiva, considereu el problema

$$\min_{x \in \mathbb{R}^n} f(x),$$

que té com a solució el punt x^*. Considereu un punt $x^0 \neq x^*$. Demostreu que si $e^0 = x^0 - x^*$ és un vector propi de Q, llavors el mètode de màxim descens (o del gradient) amb exploració lineal exacta arribarà a x^* en una iteració.

ii. A les hipòtesis de l'apartat (a) hagués estat equivalent sustituir la primera frase per "Donada la funció quadràtica estrictament convexa $f(x) = \frac{1}{2}x^T Q x - b^T x$, considereu el problema..."? Justifiqueu-ho breument (sense cap demostració).

iii. Si a les hipòtesis de l'apartat (a) haguéssim considerat "Q **semidefinida** positiva", què implicaria que e^0 fos vector propi associat al valor propi $\lambda = 0$? Seria encara cert el resultat de l'apartat (a)? Suposeu que existeix un punt x^* solució.

iv. Considereu

$$\min_{x \in \mathbb{R}^2} \frac{1}{2}x^T Q x - b^T x \qquad Q = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}.$$

A. Per a quins valors de b té solució aquest problema i calculeu el mínim.

B. Quin inconvenient tindria el mètode de Newton si l'apliquem (numèricament) a aquest problema?

2.23. [Problema examen gener 2013]

i. Considereu el problema sense restriccions

$$\min_{x \in \mathbb{R}^2} f(x) = x_1^4 + 2x_1^2 x_2^2 + x_2^4.$$

A. Comproveu que la funció objectiu és convexa.

B. Suposeu que $x^k = (a,a)^T$, $a \in \mathbb{R}, a \neq 0$ (és a dir, té les dues components iguals). Deriveu l'expressió de x^{k+1} usant el mètode de Newton amb $\alpha^k = 1$. Considerant llavors que $x^0 = (1,1)^T$, proveu que la seqüència generada pel mètode de Newton (usant $\alpha^k = 1$ a tota iteració k) és

$$x^k = \begin{bmatrix} \left(\frac{2}{3}\right)^k \\ \left(\frac{2}{3}\right)^k \end{bmatrix}.$$

C. Us porta al mínim del problema aquesta seqüència?

ii. A. La seqüència de punts x^k generada pel mètode de Newton a l'apartat (a) anterior, quina convergència local té (quadràtica, lineal, superlineal)? Justifiqueu-ho.

B. Contradiu el trobat al punt i. anterior el teorema de convergència local del mètode de Newton vist a classe, que afirma que sota certes hipòtesis el mètode de Newton convergeix quadràticament quan estem suficientment a prop de la solució? Justifiqueu-ho (màxim tres/quatre línies de justificació).

iii. A classe vam veure que quan s'aplica el mètode de màxim descens a una funció quadràtica convexa $f(x) = 1/2x^TQx - b^Tx$, dues direccions consecutives són ortogonals si es fa exploració lineal exacta. Demostreu que aquesta propietat és generalitzable a funcions qualsevols $f \in \mathcal{C}^1$: si es fa exploració lineal exacta es verifica que $(d^k)^Td^{k+1} = 0$, on d^i és la direcció de moviment del mètode de màxim descens aplicat a f al punt x^i.

2.24. [Problema examen gener 2014] El mètode de Newton no garanteix direccions de descens si la Hessiana no és definida positiva. Per evitar això hi ha modificacions al mètode de Newton, i també un tipus de mètodes anomenats quasi-Newton. El mètodes quasi-Newton funcionen com el mètode de Newton, però usen en comptes de la Hessiana $\nabla^2 f(x^k)$ una matriu B_k simètrica i definida positiva (aquí no indiquem quina expressió té, no ens cal). La seqüència de punts dels mètodes quasi-Newton (com al de Newton) és $x^{k+1} = x^k + \alpha_k d^k$, on $d^k = B_k^{-1}(-\nabla f(x^k))$. Les matrius B_k garanteixen:

$$B_{k+1}s^k = y^k \qquad \text{on } s^k = x^{k+1} - x^k \quad y^k = \nabla f(x^{k+1}) - \nabla f(x^k). \tag{1}$$

i. Usant (1), proveu que $(s^k)^Ty^k > 0$ és condició necessària per a que B_{k+1} sigui definida positiva.

ii. Proveu que si α_k satisfà la segona condició d'Armijo-Wolfe llavors sempre es garanteix que $(s^k)^Ty^k > 0$.

2.25. [Problema examen gener 2015] Una de les aplicacions actuals més rellevants de l'optimització és la solució del problema anomenat "compressed sensing" i d'altres relacionats. Aquest problema s'utilitza en reconstrucció o decodificació d'imatges i senyals. Es tracta de recuperar un vector $x \in \mathbb{R}^n$ *espars* (això és, amb pocs elements diferents de zero) a partir d'unes observacions $b \in \mathbb{R}^m$, on $m \leq n$ (normalment $m \ll n$, i n pot ser molt gran, de l'ordre de milions). Les observacions es calculen a través de $b = Ax$, on $A \in \mathbb{R}^{m\times n}$ i suposarem que té rang complet. El problema es formula com

$$\min_{x\in\mathbb{R}^n} \rho\|x\|_1 + \frac{1}{2}\|Ax-b\|_2^2,$$

on $0 < \rho \in \mathbb{R}$ és un paràmetre positiu de ponderació, i $\|\cdot\|_1$ i $\|\cdot\|_2$ denoten la norma 1 i norma 2 d'un vector. Com que $\|x\|_1$ no és diferenciable, podem aproximar $|x_i|$ per l'anomenada funció pseudo-Huber:

$$\phi_\delta(x_i) = +\sqrt{\delta^2 + x_i^2} - \delta,$$

on $0 < \delta \in \mathbb{R}$ és un paràmetre fixat i proper a zero. Com més proper a zero sigui δ, millor aproximem el valor absolut amb la funció ϕ_δ. El problema d'optimització resultant és ara:

$$(P) \qquad \min_{x\in\mathbb{R}^n} f(x) \triangleq \rho\sum_{i=1}^n \phi_\delta(x_i) + \frac{1}{2}\|Ax-b\|_2^2.$$

i. Podeu garantir que un òptim local de (P) serà també òptim global? Justifiqueu-ho.

ii. Suposeu per simplificar els càlculs que $A^TA = I$. Calculeu l'expressió de la direcció del mètode de Newton aplicat a (P) en un punt x qualsevol. Simplifiqueu l'expressió de la direcció de Newton per a les components $x_i = 0$.

iii. Calculeu la fita superior de la velocitat de convergència del mètode del gradient en aquest problema quan $\delta \to 0$. Suposeu per simplificar que $A^TA = I$; suposeu també que a la solució x^* els dos subconjunts de la partició de components x_i^* que són zero i diferents de zero són no-buits, i que existeix un $\bar{\delta}$ tal que aquesta partició no varia per a $\delta < \bar{\delta}$. Creieu que el mètode del gradient seria un mètode efectiu en aquest cas?

2.26. [Problema examen gener 2016]

i. Demostreu que el mètode de Newton és *invariant* a transformacions lineals no-singulars $x = Ay$, on $A \in \mathbb{R}^{n\times n}$ és no-singular. És a dir, que donats el problema $\min f(x), x \in \mathbb{R}^n$, on $f : \mathbb{R}^n \to \mathbb{R}$, $f \in \mathcal{C}^2$, i $\min g(y), y \in \mathbb{R}^n$, on $g(y) = f(Ay)$, i suposant que els punts inicials dels dos problemes verifiquen $x^0 = Ay^0$, i la longitud de pas és en els dos problemes la mateixa i igual a α, llavors es verifica que $x^k = Ay^k$, per a tot $k \geq 0$. Això ens indica que el mètode de Newton no està afectat per canvis de coordenades.

ii. Demostreu que, sota les mateixes hipòtesis del problema anterior, el mètode de màxim descens o del gradient NO és invariant.
Segons la demostració que heu fet, indiqueu quina propietat li hauríeu d'imposar a la matriu A per a que el mètode de màxim descens fos invariant.
Creieu que tindria interès fer una transformació lineal amb una matriu que tingués aquesta propietat en el mètode de màxim descens? Per què? (responeu en una o dues línees).

2.27. [Problema examen gener 2017]

i. Donada la funció $\varphi : [l, u] \subset \mathbb{R} \to \mathbb{R}$ definida com

$$\varphi(b) = \begin{array}{ll} \min & f(x) \\ \text{s. a} & g(x) \leq b, \end{array}$$

on $b \in [l, u] \subset \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$ és una funció convexa i $g : \mathbb{R}^n \to \mathbb{R}$ també és una funció convexa, demostreu que φ és una funció convexa (és a dir, el resultat d'un problema d'optimització convex és una funció convexa del terme de la dreta de les restriccions). Suposeu que per a tot $b \in [l, u]$ el problema d'optimització té solució (és factible i no il·limitat). (Nota: aquest resultat és també cert si $b \in \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^m$, i $m > 1$; aquí ens limitem al cas $m = 1$ per simplificar la notació).

ii. Ara volem solucionar el problema sense restriccions

$$\min_{b\in\mathbb{R}^m} \varphi(b)$$

(suposem que b no està restringida a $[l, u]^m$, i $\varphi(b)$ continua essent convexa). Volem usar el mètode del gradient o màxim descens, i ens trobem al punt b_k. Doneu l'expressió per calcular b_{k+1} en funció de b_k. Suposeu que la funció és diferenciable a b_k, i que teniu un procediment per calcular la longitud de pas α_k. (**Nota**: la resposta ha de ser breu, unes poques línies).

2.28. [Problema examen gener 2018]

i. Donada la funció $\phi : C \subseteq \mathbb{R}^n \to \mathbb{R}$ definida com

$$\phi(c) = \begin{array}{ll} \min & c^T x \\ \text{s. a} & g(x) \leq b, \end{array}$$

on $c \in C \subseteq \mathbb{R}^n$, $b \in \mathbb{R}^m$, i $g : \mathbb{R}^n \to \mathbb{R}^m$ és una funció convexa (és a dir cada component $g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$ de g és una funció convexa), demostreu que ϕ és una funció còncava (una funció ϕ és còncava si $-\phi$ és convexa). Suposeu que per a tot $c \in C$ el problema d'optimització té solució (és factible i no il·limitat).

ii. Ara volem solucionar el problema sense restriccions

$$\max_{c \in \mathbb{R}^n} \phi(c)$$

(suposem que, tot i que c no està restringida a C, ϕ és igualment còncava). Volem usar el mètode del gradient o màxim ascens, i ens trobem al punt c_k. Doneu l'expressió per calcular c_{k+1} en funció de c_k. Suposeu que la funció ϕ és diferenciable a c_k; que en un entorn B_k de c_k la solució $x^*(c), c \in B_k$, del problema d'optimització no varia ($x^*(c)$ és constant per a $c \in B_k$); i que teniu un procediment per calcular la longitud de pas α_k. (**Nota**: la resposta ha de ser molt breu, unes poques línies).

2.29. [Problema examen gener 2019] Considerem una funció $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1$. Si el gradient és Lipschitz continu de constant $L > 0$, és a dir,

$$\|\nabla f(y) - \nabla f(x)\| \leq L\|x - y\| \quad \forall\, x, y \in \mathbb{R}^n,$$

es pot demostrar (aquí no ho farem, ho donem per fet) que la funció anterior verifica, per a tot $x, y \in \mathbb{R}^n$, la següent desigualtat:

$$|f(y) - f(x) - \nabla f(x)^T(y - x)| \leq \frac{L}{2}\|y - x\|^2. \tag{2}$$

Volem solucionar $f(x^*) = \min_{x \in \mathbb{R}^n} f(x)$, usant el mètode del gradient $x^{k+1} = x^k - \alpha \nabla f(x^k)$. L'objectiu d'aquest exercici es provar que el nombre d'iteracions del mètode del gradient que cal fer per tenir una solució on $\|\nabla f(x)\| \leq \epsilon$, essent ϵ una tolerancia fixada, és $O\left(\frac{1}{\epsilon^2}\right)$ (resultat molt interessant, perquè indica que el nombre d'iteracions depèn de la tolerància ϵ i no de n, el nombre de variables del problema).

i. Proveu, usant (2), que a cada iteració del mètode del gradient, per un α determinat i independent del número d'iteració, es verifica la següent desigualtat:

$$f(x^{k+1}) \leq f(x^k) - \frac{1}{2L}\|\nabla f(x^k)\|^2. \tag{3}$$

Això ens dóna una idea del rendiment del mètode del gradient a cada iteració.

ii. Usant (3), proveu que per a qualsevol número d'iteració N es verifica la següent desigualtat:

$$\frac{1}{2L}\sum_{k=0}^{N}\|\nabla f(x^k)\|^2 \leq f(x^0) - f(x^*). \tag{4}$$

iii. Definint el mínim dels gradients trobats fins a la iteració N com

$$g_N = \min_{0 \leq k \leq N} \|\nabla f(x^k)\|,$$

i usant (4), proveu que el nombre d'iteracions N que cal fer per garantir que $g_N \leq \epsilon$ és

$$N \geq \frac{2L}{\epsilon^2}\left(f(x^0) - f(x^*)\right) - 1.$$

Això prova que el nombre d'iteracions del mètode del gradient és $N = O\left(\frac{1}{\epsilon^2}\right)$: depèn de ϵ i no de la dimensio n.

2.30. [Problema examen gener 2020] Considereu les condicions d'Armijo-Wolfe:

$$g(\alpha) \;\leq\; g(0) + \alpha c_1 g'(0) \qquad 0 < c_1 < 1 \tag{AW1}$$
$$g'(\alpha) \;\geq\; c_2 g'(0) \qquad 0 < c_1 < c_2 < 1, \tag{AW2}$$

on $g(\alpha) = f(x^k + \alpha d^k)$ i $f : \mathbb{R}^n \to \mathbb{R} \in C^1$ és la funció objectiu a minimitzar d'un problema d'optimització sense restriccions.

Considereu les següents hipòtesis:

- d^k és direcció de descens de f al punt x^k.
- f està afitada inferiorment al llarg de $\{x^k + \alpha d^k : \alpha > 0\}$.

i. Proveu que existeix (al menys) un $\widehat{\alpha} > 0$ tal que la funció $g(\alpha)$ intersecta amb la línia $l(\alpha) = g(0) + \alpha c_1 g'(0)$ (és a dir, $g(\widehat{\alpha}) = l(\widehat{\alpha})$).

ii. Proveu que, si $0 < c_1 < c_2 < 1$, llavors existeix un interval de valors dintre de l'interval $(0, \bar{\alpha})$ que verifica les condicions (AW1) i (AW2), on $\bar{\alpha} = \min\{\widehat{\alpha} > 0 : g(\widehat{\alpha}) = l(\widehat{\alpha})\}$ (és a dir, $\bar{\alpha}$ és, de tots els valors $\widehat{\alpha} > 0$ que verifiquen $g(\widehat{\alpha}) = l(\widehat{\alpha})$, el més petit).

El resultat anterior garanteix que sempre podem fer una exploració lineal i satisfer les condicions (AW1) i (AW2) en mètodes d'optimització sense restriccions.

3. Problemes d'optimització amb restriccions

3.1. Sigui $v : \mathbb{R}^n \to \mathbb{R}^m$ una funció vectorial suau ($v_i \in \mathcal{C}^2, i = 1, \dots, m$). Considereu els dos problemes d'optimització $\min f(x)$ on

$$f(x) = \|v(x)\|_\infty \qquad\qquad f(x) = \max_{i=1,\dots,m} v_i(x).$$

Reformuleu els dos problemes com a problemes d'optimització "suaus" (sense que apareguin el "max" ni la "$\|.\|_\infty$" a la funció objectiu).

3.2. En $\mathbb{R}^2$ considereu la regió factible definida per la restricció de desigualtat

$$\|x\|_1 = |x_1| + |x_2| \leq 1.$$

Aquest restricció no és "suau" (és a dir, si la escrivim com $g(x) \leq 0$ tenim que $g(x) = |x_1| + |x_2| - 1$ i $g \notin \mathcal{C}^2$). Reformuleu la mateixa regió factible usant restriccions "suaus".

3.3. Considereu el problema amb una restricció d'igualtat $\min f(x)$ s. a $h(x) = 0$. Proveu que si la condició necessària de primer ordre $\nabla f(x) + \lambda h(x) = 0$ no es verifica al punt x, llavors la direcció

$$d = -\left(I - \frac{\nabla h(x) \nabla h(x)^T}{\|\nabla h(x)\|^2}\right) \nabla f(x)$$

és factible i de descens.

3.4. Considereu el problema amb una restricció de desigualtat $\min f(x)$ s. a $g(x) \leq 0$. Proveu que si la restricció és inactiva i $\nabla f(x) \neq 0$ llavors la direcció

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

és factible i de descens.

3.5. Donat el semipla $H = \{x \in \mathbb{R}^n : a^T x + \alpha \geq 0\}$ trobeu el punt de H de norma Euclídea mínima.

3.6. Donat un conjunt polièdric definit per $C = \{x \in \mathbb{R}^n : Ax = b, A \in \mathbb{R}^{m \times n}, m \leq n\}$ (suposem que A té rang complet), la projecció d'un punt $x_0 \in \mathbb{R}^n \setminus C$ sobre C es defineix com el punt x de C més proper a x_0. Trobeu una expressió analítica de x formulant i solucionant (usant les condicions KKT) el corresponent problema d'optimització amb restriccions. És un problema d'optimització convex?

3.7. Considereu el següent problema d'optimització no lineal amb restriccions:

$$\begin{aligned} \min \;\; & f(x_1, x_2) = x_1 x_2 + (x_1 - x_2)^3 \\ & x_1 \geq 0, \;\; x_2 \geq 0 \end{aligned}$$

Determineu si les direccions $d_1 = (-1, 1)^T$, $d_2 = (1, 2)^T$ i $d_3 = (3/2, 2)^T$ són factibles i de descens a partir del punt $x_0 = (0, 1)^T$. Justifiqueu la resposta.

3.8. Considereu el següent problema de programació no lineal:

$$\begin{aligned} &\min\ f(x_1,x_2,x_3) = (x_1x_2-3)^2 - 4x_1x_2 + 2(x_1x_3-1)^3 - \frac{15}{2}x_2 + 5x_3 \\ &6x_1^2 - 3x_2x_3 - x_3^2 \le 5 \\ &x_1^2 + \frac{1}{2}x_2 + x_3^2 \le 4 \end{aligned}$$

i. Formuleu les condicions KKT.
ii. Comproveu si el punt $x = (0,-2,1)^T$ pot ser un mínim local del problema.

3.9. Considereu el següent problema de programació no lineal amb restriccions:

$$\begin{aligned} &\min\ f(x_1,x_2) = x_1x_2 + (x_1-x_2)^3 \\ &(x_1+1)(x_2+1) \le 3 \\ &x_1 \ge 0,\ \ x_2 \ge 0 \end{aligned}$$

i. Comproveu a partir de les condicions KKT si el punt $x = (1,1/2)^T$ és un mínim local del problema.
ii. Tenint en compte la regió factible del problema, podem dir que es tracta d'un problema de programació convexa? Justifiqueu la resposta.

3.10. Considereu el següent problema de programació no lineal amb restriccions:

$$\begin{aligned} &\min\ f(x_1,x_2) = (x_1-1)^2x_2 + (x_2-x_1)^2 \\ &(x_1+1)(x_2+1) + 2x_1^2 + 2x_2^2 \le 4 \\ &x_1 \ge 0,\ \ x_2 \ge 0 \end{aligned}$$

i. Formuleu les condicions KKT.
ii. Comproveu, a partir de les condicions KKT, si el punt $x = (0,1)^T$ és un mínim local del problema.

3.11. Considereu el problema

$$\begin{aligned} &\min\ f(x_1,x_2,x_3) = -\frac{1}{3}x_1^3 + \frac{1}{2}x_2^2 - 2x_2x_3 + 4x_3 \\ &\frac{1}{2}x_1^2 - 2x_3^2 - 2x_3 = \frac{1}{2} \\ &\frac{1}{2}x_1^2 - \frac{1}{2}x_2^2 + 2x_2 + 2x_3^3 + x_1 + 2x_3 \le \frac{3}{2} \\ &x_1 + x_3 \le 0 \end{aligned}$$

Comproveu que $x = (1,0,-1)^T$ verifica les condicions KKT.

3.12. Trobeu un punt que verifiqui les condicions KKT per al problema

$$\begin{aligned} &\min\ f(x_1,x_2) = \frac{1}{2}\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 9x_1 - 5x_2 \\ &x_1 - x_2 \ge -1 \\ &x_2 \le \frac{3}{2} \\ &x_1 + x_2 \le 3 \\ &x_1 \le \frac{5}{2} \end{aligned}$$

3.13. Considereu el problema

$$\begin{aligned} &\min\ f(x_1,x_2) = -(x_1-1)^2 - (x_2-1)^2 \\ &x_1^2 + x_2^2 \ge 1 \\ &x_1^2 + x_2^2 \le 4 \\ &x_1,\ x_2 \ge 0 \end{aligned}$$

i. Dibuixeu el domini del problema i les corbes de nivell de la funció objectiu. Obteniu gràficament la solució òptima del problema.
ii. Formuleu les condicions KKT i comproveu que el punt $x_a = (0,1)^T$ les verifica. Són suficients les condicions KKT per al problema? Per què?
iii. Considereu ara el punt $x_b = (0, 3/2)^T$:
   A. Caracteritzeu les direccions factibles en el punt x_b.
   B. Caracteritzeu les direccions de descens en el punt x_b.
   C. Comproveu que $d = (0,1)^T$ satisfà les condicions (i.) i (ii.) i obteniu el valor de λ que dóna el punt mínim en aquesta direcció en el domini factible.

3.14. Contesteu les preguntes següents justificant les respostes:

i. Per a un problema de minimització no lineal, pot un punt KKT ser un màxim local?
ii. Sigui f una funció diferenciable, X un conjunt convex i sigui $\bar{x} \in X$ tal que $\nabla f(\bar{x})^T(x - \bar{x}) > 0$, $\forall x \in X$, $x \neq \bar{x}$. És necessariament $\bar{x}$ un mínim local?

3.15. Considereu el següent problema:

$$\begin{aligned} \max\ & x_1^2 + 4x_1x_2 + x_2^2 \\ & x_1^2 + x_2^2 = 1 \end{aligned}$$

i. Utilitzant les condicions KKT, trobeu un solució òptima per al problema.
ii. Comproveu les condicions d'optimalitat de segon ordre.
iii. Té el problema una única solució òptima?

3.16. Considereu el següent problema d'optimització lineal:

$$\begin{aligned} \max\ & 2x_1 + 3x_2 \\ & x_1 + x_2 \geq 8 \\ & -x_1 + 2x_2 \geq 4 \\ & x_1,\ x_2 \geq 0 \end{aligned}$$

i. Escriviu les condicions KKT.
ii. Per a cada extrem, verifiqueu si es compleixen les condicions KKT, tant de forma algebraica com geomètrica. A partir d'aquí, trobeu la solució òptima.

3.17. Considereu el següent problema, on c és un vector fila no nul donat:

$$\begin{aligned} \max\ & cd \\ & d^Td \leq 1 \end{aligned}$$

i. Proveu que $\bar{d} = c/\|c\|$ és un punt KKT. Proveu també que $\bar{d}$ és l'única solució òptima global.
ii. Utilitzant l'apartat anterior, proveu que la direcció de màxim ascens de f en un punt x ve donada per l'expressió $\nabla f(x)/\|\nabla f(x)\|$, suposant que $\nabla f(x) \neq 0$.

3.18. Considereu el següent problema de programació no lineal (amb funció quadràtica i restriccions lineals)

$$\begin{aligned} \min\ & x_1^2 + x_2^2 + x_3^2 - x_1x_3 - 3x_1 + 2x_2 \\ & x_1 + x_2 + x_3 = 5 \\ & 2x_1 + x_2 - x_4 = 1 \\ & x_i \geq 0\ ,\ \ i = 1, ..., 4 \end{aligned}$$

i. Comproveu que el punt $x^* = (3, 0, 2, 5)^T$ i els multiplicadors $\lambda^* = (-1, 0)$ i $\mu^* = (0, 1, 0, 0)^T$ (per a les restriccions d'igualtat i desigualtat respectivament), satisfan les condicions necessàries i suficients d'optimalitat del problema, per la qual cosa podem concloure que $x^* = (3, 0, 2, 5)^T$ és l'òptim del nostre problema.

3.19. **[Problema examen gener 2012]** *El problema planimètric de Kepler.* El problema planimètric de Kepler estableix que de tots els rectangles inscrits en un cercle, el que té àrea màxima és un quadrat. Haureu de demostrar aquest resultat formulant el problema planimètric de Kepler com a problema d'optimització no lineal amb restriccions.

i. En $\mathbb{R}^2$, considerant un cercle de centre $(0,0)$ i radi $\sqrt{2}$, directament observant la Figura 1 s'obté que el problema pot ser formulat com

$$\begin{array}{rl} \max & 4x_1x_2 \\ \text{s.a} & x_1^2 + x_2^2 = 2 \\ & x_1 \geq 0, x_2 \geq 0 \end{array}$$

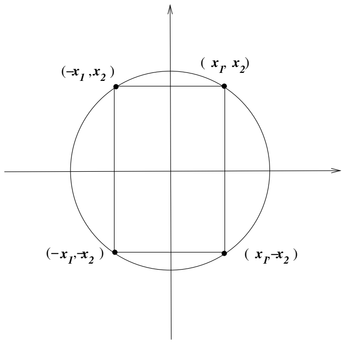

Figura 1: El rectangle inscrit té àrea $4x_1x_2$

Solucioneu el problema anterior usant les condicions d'optimalitat d'un problema amb restriccions (necessàries de primer ordre, i suficients de segon ordre), i comproveu que la solució correspon a un quadrat.

ii. El problema anterior (un cop transformat a problema de minimització) és un problema d'optimització convex? Justifiqueu-ho. Sabem que si el problema és convex, l'òptim obtingut és global. Però el quadrat proporciona l'òptim global al problema anterior. Hi ha llavors una contradicció?

iii. Demostreu el resultat general en $\mathbb{R}^n$ (que de tots els hiperrectangles inscrits en una hiperesfera de centre 0 i radi $\sqrt{n}$, el de volum màxim és un hipercub) usant el problema:

$$\begin{array}{rl} \max & 2^n \displaystyle\prod_{i=1}^{n} x_i \\ \text{s.a} & \displaystyle\sum_{i=1}^{n} x_i^2 = n \\ & x_i \geq 0 \quad i = 1, \ldots, n. \end{array}$$

Solucioneu el problema anterior usant les condicions d'optimalitat d'un problema amb restriccions (necessàries de primer ordre, i suficients de segon ordre), i comproveu que la solució correspon a un hipercub (és a dir, que totes les x_i tenen el mateix valor a l'òptim).

3.20. **[Problema examen juny 2012]** *Demostració de la llei de la refracció en òptica.* Considereu una corba en $\mathbb{R}^2$ definida per $h(x) = 0$, on $h : \mathbb{R}^2 \to \mathbb{R}$ és diferenciable. Tal i com mostra la Figura 2 la corba separa $\mathbb{R}^2$ en dos medis Y i Z. Suposem que la llum es propaga en ells amb velocitats v_Y i v_Z respectivament. Per anar d'un punt $y \in \mathbb{R}^2$ del medi Y a un punt $z \in \mathbb{R}^2$ del medi Z la llum passa per un punt x sobre la corba (y, z i x es mostren a la Figura 2). La llum segueix la trajectòria de temps mínim, i d'aquest fet resulta la llei de la refracció en òptica:

$$\frac{\sin \alpha_y}{\sin \alpha_z} = \frac{v_Y}{v_Z},$$

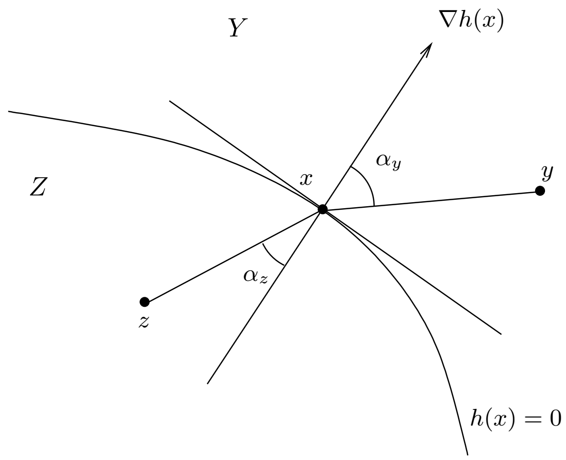

Figura 2: La corba $h(x) = 0$ separa els dos medis Y i Z.

és a dir, el quocient dels sinus dels angles és constant i igual al quocient de les velocitats, independentment dels punts y, z considerats.

i. Es demana que deduïu la llei de refracció anterior usant les condicions d'optimalitat del problema d'optimització amb restriccions següent, on la funció objectiu indica el temps d'anar de y a x i de x a z:

$$\begin{array}{ll} \min\limits_{x} & f(x) = \dfrac{||x-y||}{v_Y} + \dfrac{||x-z||}{v_Z} \\ \text{s.a} & h(x) = 0. \end{array}$$

(Nota: Ajudeu-vos de la Figura 2 per interpretar geomètricament el què imposen les condicions d'optimalitat, i d'aquí deduïu la llei de la refracció.)

ii. A l'apartat anterior hem de suposar que $\nabla h(x^*) \neq 0$ al punt solució del problema d'optimització. Per què?

3.21. [Problema examen gener 2013] Es vol dissenyar una llauna cilíndrica per a la nova beguda refrescant Caca-Cola, de superfície mínima i que el volum sigui de com a mínim v centilitres. Es planteja el problema d'optimització (P), usant les variables radi r i alçada h de la llauna:

$$(P) \quad \begin{array}{rl} 2\pi \cdot \min & r^2 + rh \\ \text{s.a} & \pi r^2 h \geq v \\ & r \geq 0 \\ & h \geq 0. \end{array}$$

i. Fent el canvi de variables $r = e^{x_1}$ i $h = e^{x_2}$, comproveu que el problema (P) és equivalent al problema $(P2)$:

$$(P2) \quad \begin{array}{rl} 2\pi \cdot \min & e^{2x_1} + e^{x_1+x_2} \\ \text{s.a} & 2x_1 + x_2 \geq \ln\left(\dfrac{v}{\pi}\right). \end{array}$$

ii. A. És (P) un problema d'optimització convex? Justifiqueu-ho.
B. Comproveu que $(P2)$ és un problema d'optimització convex.

iii. Calculeu el mínim global de (P) usant $(P2)$. Indiqueu la relació entre r i h a l'òptim, i els valors òptims de r i h en funció de v.

3.22. [Problema examen juliol 2013] Un el·lipsoide és una generalització a $\mathbb{R}^n$ de la idea d'el·lipse. Un el·lipsoide és un conjunt convex que es pot definir com $E(c, Q, r) = \{x \in \mathbb{R}^n : (x-c)^T Q (x-c) \leq r^2\}$, on c és el seu centre, Q és una matriu definida positiva (podem suposar simètrica), i $r > 0$ és el radi (usualment s'usa sempre $r = 1$, però aquí per comoditat considerem un r qualsevol positiu). Si Q és diagonal, l'el·lipsoide està alineat amb els eixos coordenats; si Q no és diagonal, l'el·lipsoide no està alineat amb els eixos coordenats (està "inclinat"). Fixeu-vos que si $Q = I$ (I és la matriu identitat) $E(c, I, r)$ és una bola de centre c i radi r.

Volem calcular la distància mínima entre dos el·lipsoides, o equivalentment, els punts més propers entre dos el·lipsoides. Formulem el problema d'optimització no lineal amb restriccions

$$\begin{array}{rl} \min\limits_{x_1 \in \mathbb{R}^n, x_2 \in \mathbb{R}^n} & \|x_1 - x_2\|^2 = (x_1 - x_2)^T (x_1 - x_2) \\ \text{s. a} & (x_1 - c_1)^T Q_1 (x_1 - c_1) \leq r_1^2 \\ & (x_2 - c_2)^T Q_2 (x_2 - c_2) \leq r_2^2, \end{array}$$

on x_i són els punts que busquem de cada el·lipsoide $E_i(c_i, Q_i, r_i)$, $i = 1, 2$.

i. Usant les condicions d'optimalitat, solucioneu el problema, ajudant-vos de la interpretació geomètrica pels diferents casos si us cal. **Nota**: pel cas que les dues restriccions siguin actives **no heu de solucionar** el sistema d'equacions resultant, només plantejar-lo.
ii. Ara considereu que $Q_1 = Q_2 = I$ (teniu dues boles). Solucioneu el sistema d'equacions que s'obté quan les dues restriccions són actives. Ajudeu-vos de la interpretació geomètrica d'aquest problema per solucionar el sistema, si us facilita la feina (quins són els punts més propers de dues boles que no es toquen, o només es toquen en un punt?). Quina implicació té de cara als multiplicadors de Lagrange que només es toquin en un punt?

3.23. [Problema examen gener 2014] *El problema d'Arquímedes del casquet esfèric de màxim volum.* Arquímedes va plantejar el següent problema: *De tots els casquets esfèrics de superfície (lateral, sense la base) $a > 0$, quin és el que té volum màxim?* La Figura 3 mostra un casquet esfèric en una esfera de radi r. El casquet ve determinat per l'alçada h, $0 \leq h \leq 2r$.

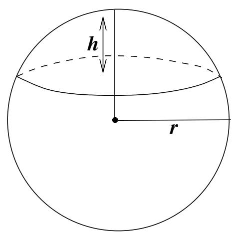

Figura 3: Casquet esfèric d'alçada h en una esfera de radi r.

Les fórmules de la superfície S (lateral, sense la base) i volum V del casquet esfèric són $S = 2\pi rh$ i $V = \pi h^2(r - h/3)$. Per solucionar el problema d'Arquímedes plantegem el problema d'optimització:

$$\begin{array}{rl} \max & \pi h^2 (r - h/3) \\ \text{s. a} & 2\pi r h = a \\ & 0 \leq h \leq 2r \\ & r \geq 0. \end{array}$$

Solucioneu aquest problema d'optimització amb restriccions utilitzant les condicions de primer i segon ordre, i comproveu que el casquet òptim correspon a una semiesfera. **Nota:** Per evitar tenir tants multiplicadors de Lagrange, usant l'equació d'igualtat podeu eliminar r, i formular un problema amb funció objectiu no lineal i només límits a h (elimineu r encara que h sigui 0).

3.24. [Problema examen gener 2014] *Condicions d'optimalitat de problemes amb només límits.* Considereu el problema

$$\begin{array}{rl} \min & f(x) \\ \text{s. a} & l \leq x \leq u, \end{array}$$

on $f : \mathbb{R}^n \to \mathbb{R}$ pertany a $\mathcal{C}^2$ i és convexa, i suposem que $u_i > l_i$, $i = 1, \ldots, n$. Escriviu i simplifiqueu les condicions d'optimalitat fins obtenir una expressió que permeti determinar "fàcilment" si un punt x^* és òptim. Apliqueu-ho al cas on $f(x) = c^T x$ és lineal per calcular x^* (càlcul força obvi, per altra banda).

3.25. [Problema examen juliol 2014] Volem solucionar el següent problema d'optimització

$$(P) \quad \begin{array}{ll} \min\limits_{x} & f(x) \\ \text{s. a} & b \leq g(x) \leq B \\ & l \leq x \leq u, \end{array}$$

on $f : \mathbb{R}^n \to \mathbb{R}$ i $g : \mathbb{R}^n \to \mathbb{R}^m$ pertanyen a $\mathcal{C}^2$, $-\infty < b_j < B_j < +\infty$ $j = 1, \ldots, m$, i $-\infty < l_i < u_i < +\infty$ $i = 1, \ldots, n$.
Però tenim un paquet d'optimització que només resol problemes en la forma estàndard següent:

$$(\tilde{P}) \quad \begin{array}{ll} \min\limits_{\tilde{x}} & \tilde{f}(\tilde{x}) \\ \text{s. a} & \tilde{g}(\tilde{x}) = \tilde{b} \\ & 0 \leq \tilde{x} \leq \tilde{u}. \end{array}$$

Per aquest motiu transformem el problema (P) en $(\tilde{P})$ fent els següents canvis de funcions i variables:

$$\tilde{x} = \begin{pmatrix} x - l \\ s \end{pmatrix}, \quad \tilde{u} = \begin{pmatrix} u - l \\ B - b \end{pmatrix}, \quad \tilde{b} = B, \quad \tilde{g}(\tilde{x}) = g(x) + s, \quad \tilde{f}(\tilde{x}) = f(\tilde{x}_{1:n} + l) = f(x),$$

on $\tilde{x}, \tilde{u} \in \mathbb{R}^{n+m}$, $\tilde{b} \in \mathbb{R}^m$, $\tilde{f} : \mathbb{R}^{n+m} \to \mathbb{R}$, $\tilde{g} : \mathbb{R}^{n+m} \to \mathbb{R}^m$, i $\tilde{x}_{1:n}$ representa les n primeres components de $\tilde{x}$.

i. Comproveu que (P) i $(\tilde{P})$ són equivalents.
ii. A partir de la solució de $(\tilde{P})$ (variables $\tilde{x}$ i multiplicadors de Lagrange de $(\tilde{P})$) obtingueu i demostreu/justifiqueu quina seria la solució de (P) (variables x i tots els multiplicadors de Lagrange de (P)). (Useu només les condicions de primer ordre, com faria el paquet d'optimització, no considereu les de segon ordre.)

3.26. [Problema examen gener 2015] Considereu la Figura 4, dintre de l'espai euclidià $\mathbb{R}^2$ (seria igualment vàlid per a $\mathbb{R}^n$). Ens trobem al punt a i volem arribar el més ràpid possible al punt b per socórrer a un amic que ens demana ajut. Pel mig, però, tenim una zona amb un medi diferent (aigua, fang,...) delimitada per les línies paral·leles $c^T x = d_1$, $c^T x = d_2$, $c \neq 0$. Suposem que ni a ni b es troben a sobre d'aquestes línies. Ens traslladem pels dos medis a unes velocitats $v_1 \in \mathbb{R}$ i $v_2 \in \mathbb{R}$.

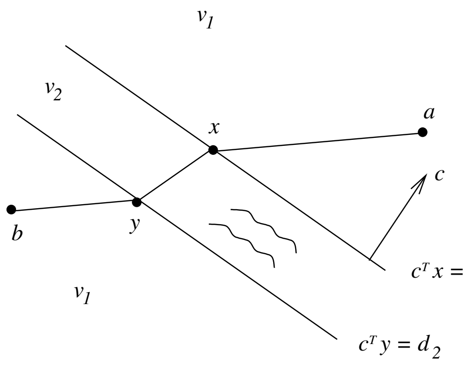

Figura 4: La trajectòria més ràpida per anar del punt a al b.

Usant les condicions d'optimalitat del problema d'optimització següent, que minimitza la suma de temps als tres trams:

$$\begin{array}{ll} \min\limits_{x,y} & f(x,y) \triangleq \dfrac{||x-a||}{v_1} + \dfrac{||x-y||}{v_2} + \dfrac{||y-b||}{v_1} \\ \text{s.a} & c^T x = d_1 \\ & c^T y = d_2, \end{array}$$

proveu que $x - a$ i $b - y$ es troben en la mateixa direcció. (Nota: això pot provar-se d'altres formes, usant per exemple la llei de la refracció de la llum —el que faria la llum és similar al que faríem nosaltres. Però ho heu de provar usant les condicions d'optimalitat del problema d'optimització anterior, sinó la solució no serà considerada vàlida)

3.27. [Problema examen gener 2015] Considereu el problema amb restriccions d'igualtat

$$\begin{array}{rl} \min & f(x) \\ \text{s.a} & h(x) = 0, \end{array}$$

on $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^m$, $f, h_i, i = 1, \ldots, m, \in \mathcal{C}^2$. Proveu que si x^* és un punt solució que satisfà les condicions KKT, i **és un punt regular**, llavors el vector de multiplicadors de Lagrange λ és únic. És a dir, la regularitat garanteix la unicitat del vector de multiplicadors de Lagrange. (Nota: aquest resultat és vàlid també per a problemes amb desigualtats, però considerem només igualtats per simplicitat.)

3.28. [Problema examen juliol 2015] Volem minimitzar una forma quadràtica definida per la matriu simètrica $Q \in \mathbb{R}^{n \times n}$ dintre dels punts de l'esfera unitat $\|x\|^2 \leq 1$:

$$\begin{array}{rl} \min\limits_{x} & x^T Q x \\ \text{s. a} & x^T x \leq 1. \end{array}$$

Solucioneu el problema anterior, usant les condicions d'optimalitat de problemes amb restriccions, i tenint en compte les possibles definicions de Q ((semi)definida positiva, (semi)definida negativa, indefinida). A què corresponen la solució òptima i multiplicadors òptims del problema? (Nota: recordeu que si Q és simètrica; (1) tots els seus valors propis són reals; (2) la matriu és diagonalitzable usant la matriu ortogonal formada pels seus vectors propis).

3.29. [Problema examen gener 2016] Donada una matriu Hessiana $A = \begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix}$ ($a_i \in \mathbb{R}, i = 1, \ldots, 3$) volem buscar la matriu $X = \begin{bmatrix} x_1 & x_3 \\ x_3 & x_2 \end{bmatrix}$ ($x_i \in \mathbb{R}, i = 1, \ldots, 3$) semidefinida positiva que sigui més propera a A segons una norma de Frobenius (la norma de Frobenius d'una matriu quadrada M es defineix com $\|M\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n m_{ij}^2}$). Plantegem el següent problema:

$$\begin{array}{rl} \min\limits_{x} & \|X - A\|_F^2 \\ \text{s. a} & X \succeq 0, \end{array}$$

que pot ser formulat com

$$\begin{array}{rll} \min\limits_{x} & \|X - A\|_F^2 & \\ \text{s. a} & x_1 \geq 0 & [r_1] \\ & x_2 \geq 0 & [r_2] \\ & x_1 x_2 - x_3^2 \geq 0 & [r_3]. \end{array}$$

i. Escriviu les condicions KKT d'aquest problema i la Hessiana de la Lagrangiana respecte les variables x.

ii. Suposeu que us diuen que a l'òptim les restriccions r_1 i r_2 són actives i que $\mu_3 = 0$ (μ_3 és el multiplicador de Lagrange de la restricció r_3). Solucioneu el problema en aquest cas, indicant quins valors a_1, a_2, a_3 donarien lloc a aquesta solució.

iii. Per al cas particular $A = \begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}$, suposeu que us diuen que a l'òptim les restriccions r_1 i r_2 són inactives i la restricció r_3 és activa. Solucioneu el problema en aquest cas.

3.30. [Problema examen gener 2017] *Desigualtat de les mitjanes geomètrica i aritmètica.* La desigualtat de les mitjanes geomètrica i aritmètica de n nombres reals qualsevols no negatius $a_1, \ldots, a_n$ estableix que

$$\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n},$$

i és té una igualtat només si tots els a_i són iguals.

Demostrarem aquest resultat usant un problema d'optimització auxiliar de la forma següent:

(i) Definim $S = \sum_{i=1}^{n} a_i$, i $x_i = \dfrac{a_i}{S}$, $i = 1, \ldots, n$, de forma que

$$\sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0 \ i = 1, \ldots, n.$$

(ii) Solucionem el problema d'optimització següent

$$\begin{array}{ll} \max & x_1 x_2 \cdots x_n \\ \text{s. a} & x_1 + \cdots + x_n = 1 \\ & x_i \geq 0 \ i = 1, \ldots, n \end{array} \tag{5}$$

per comprovar que

$$x_1 x_2 \cdots x_n \leq \frac{1}{n^n}. \tag{6}$$

(iii) A partir de (6) i desfent el canvi de variable $x_i = \dfrac{a_i}{S}$, $i = 1, \ldots, n$, obtenim fàcilment la desigualtat de les mitjanes geomètrica i aritmètica.

Es demana que feu el següent:

i. Solucioneu el problema d'optimització (5) usant les condicions d'optimalitat de primer i de segon ordre.

ii. A partir de (6) proveu la desigualtat de les mitjanes geomètrica i aritmètica (és a dir, finalitzeu el punt (iii) anterior).

3.31. [Problema examen gener 2018] *Problema de la Olimpiada Matemàtica de 1984.* Donats x, y, z reals no negatius, tals que $x + y + z = 1$, proveu que

$$0 \leq xy + xz + yz - 2xyz \leq 7/27.$$

(Per provar això plantegeu el corresponent problema d'optimització amb restriccions d'igualtat i desigualtat, i solucioneu-lo usant les condicions KKT de primer ordre, i les de segon ordre.)

3.32. [Problema examen gener 2019] Considerem un fenòmen aleatori (com llençar un dau, o característiques físiques de persones com alçada, coeficient intel.lectual, etc) que pot prendre els valors $x_1, \ldots, x_n$ amb probabilitats respectives $p_1, \ldots, p_n$, de forma que $p_i \geq 0, i = 1, \ldots, n$, i $\sum_{i=1}^{n} p_i = 1$. D'això se'n diu *variable aleatòria discreta.* En Teoria de la Informació s'anomena *entropia* d'aquesta variable aleatòria al valor definit com $-\sum_{i=1}^{n} p_i \ln p_i$.

i. Calculeu els valors (probabilitats) p_i que maximitzen l'entropia solucionant el següent problema d'optimització (usant les condicions d'optimalitat que calguin):

$$\begin{array}{ll} \max\limits_{p_i} & -\sum\limits_{i=1}^{n} p_i \ln p_i \\ \text{s. a} & \sum\limits_{i=1}^{n} p_i = 1 \\ & p_i \geq 0 \quad i = 1, \ldots n. \end{array}$$

Nota: Usem $p_i \geq 0$ encara que el domini de ln són els reals positius. Podríem haver formulat $p_i \geq \epsilon$ essent $\epsilon \approx 0$ però això complicaria innecessàriament la notació. Si us cal useu que $\ln 0 = -\infty$.

ii. La mitjana (esperança) de la variable aleatòria es calcula com $\sum_{i=1}^{n} p_i x_i$. Suposeu que s'ha observat que el fenòmen estudiat té una mitjana de m. Calculeu ara els valors de p_i

solucionant:

$$\begin{aligned}
\max_{p_i} \quad & -\sum_{i=1}^{n} p_i \ln p_i \\
\text{s. a} \quad & \sum_{i=1}^{n} p_i = 1 \\
& \sum_{i=1}^{n} p_i x_i = m \\
& p_i \geq 0 \quad i = 1, \ldots n.
\end{aligned}$$

Nota: en aquest cas només doneu l'expressió de p_i (en funció dels multiplicadors); no calculeu els valors dels multiplicadors, perquè en general, aquests es troben de forma numèrica.

3.33. [Problema examen gener 2020] Nicolo Tartaglia (1500–1557) va formular el següent problema: dividir el nombre 8 en dues parts de forma que es maximitzi el producte de les dues parts multiplicat per la seva diferència. Els dos problemes d'optimització següents són formulacions vàlides d'aquest problema:

$$(A) \quad \begin{aligned}
\max_{x_1, x_2} \quad & x_1 x_2 (x_1 - x_2) \\
\text{s. a} \quad & x_1 + x_2 = 8 \\
& x_1 \geq x_2 \\
& x_2 \geq 0.
\end{aligned}
\qquad\qquad
(B) \quad \begin{aligned}
\max_{x_1, x_2} \quad & \ln x_1 + \ln x_2 + \ln(x_1 - x_2) \\
\text{s. a} \quad & x_1 + x_2 = 8 \\
& x_1 \geq x_2 \\
& x_2 \geq 0.
\end{aligned}$$

i. Quina de les dues formulacions (A) o (B) és un problema d'optimització convex (un cop transformat com a problema de minimització)?

ii. Resoleu el problema d'optimització ((A) o (B)) que sigui convex, usant les condicions d'optimalitat de problemes amb restriccions. **Nota 1:** no podeu eliminar d'entrada una variable a partir de la restricció d'igualtat, heu de considerar un problema de dues variables. **Nota 2:** calculeu la solució com a expressió algebràica, sense usar decimals (us surtiran arrels quadrades, però són senzilles de manipular).