Col·lecció de problemes resolts d'Aplicacions de la Programació No Lineal

(v.09.01)

Jordi Castro Pérez, F. Javier Heredia Cervera, N. Nabona

Dept. d'Estadística i Investigació Operativa
Secció d'Informàtica
Universitat Politècnica de Catalunya

Índex

1 Problemes de PNL sense constriccions.

1.1 Convexitat.

1. Conjunts convexos *

Raoneu si els següents conjunts són o no convexos:
a) $C = \{x \in \mathbb{R}^n : Ax = b,\ A \in \mathbb{R}^{m\times n} \quad b \in \mathbb{R}^m\}$
b) $C = \{x \in \mathbb{R}^n : Ax \leq b,\ A \in \mathbb{R}^{m\times n} \quad b \in \mathbb{R}^m\}$
c) $C = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$
d) $C = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1 + x_2| \leq 1\}$
e) $C = C_1 \cup C_2$, on C_1 i C_2 són conjunts convexos.
f) $C = C_1 \cap C_2$, on C_1 i C_2 són conjunts convexos.

2. Funcions convexes *

Aplicant la definició de funció convexa, comproveu que:
a) Si f_1 i f_2 són funcions convexes, $f_1 + f_2$ és també una funció convexa.
b) Si f és una funció convexa, af és convexa $\forall a \geq 0$.

3. Restriccions no lineals d'igualtat

Sabem que si $g(x)$ és una funció convexa, aleshores $g(x) \leq c$ defineix un conjunt de punts convex per a tot $c \in \mathbb{R}$. Creieu que això també és cert si la restricció fos $g(x) = c$. Raoneu-ho trobant algun contraexemple.

4. Terme diagonal negatiu *

Mostreu que si una matriu quadrada H té un terme diagonal negatiu aleshores no pot ser semidefinida positiva. Recordeu que una matriu H és semidefinida positiva si $\forall x\ x^T H x \geq 0$.

5. Determinació d'uns paràmetres per garantir la convexitat *

Determineu quins valors poden prendre els paràmetres p i q de la següent funció:

$$f(x, y) = q \ln(x^2) + \sin(py)$$

per tal que garantim que sigui convexa en algun subconjunt de $\mathbb{R}^2$.

6. El mínim global *

Si apliquéssim un determinat algorisme d'optimització al problema:

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

i aquest ens donés un determinat punt solució, podríem garantir que aquest correspon a un mínim global?

1.2 Optimalitat i direccions de descens.

7. Direccions de descens

Considereu el problema d'optimització no lineal amb constriccions següent :

$$\textbf{(PNL)} \quad \min \quad f(x_1, x_2) = x_1 x_2 + (x_1 - x_2)^3$$
$$\text{s.a}: \quad x_1 \geq 0 \,,\; x_2 \geq 0$$

Dieu si les direccions $d^1 = [\,-1 \quad 1\,]'$, $d^2 = [\,1 \quad 2\,]'$, $d^3 = [\,3/2 \quad 2\,]'$ són factibles i de descens a partir del punt $x^0 = [\,0 \quad 1\,]'$ pel problema **(PNL)**. Justifiqueu la vostra resposta.

8. Direccions de descens*

Considereu el problema d'optimització no lineal amb constriccions següent :

$$\textbf{(PNL)} \quad \min \quad f(x_1, x_2) = (x_1 - 1)^2 x_2 + (x_2 - x_1)^2$$
$$\text{s.a}: \quad x_1 \geq 0 \,,\; x_2 \geq 0$$

Dieu si les direccions $d^1 = [\,2 \quad 0\,]'$, $d^2 = [\,-2 \quad 1\,]'$, $d^3 = [\,2 \quad 2\,]'$. són factibles i de descens a partir del punt $x^0 = [\,1 \quad 0\,]'$ pel problema (PN2). Justifiqueu la vostra resposta.

9. Direccions de descens*

Donada la funció $f(X) = x_1^2 + 2x_2^2 + 3x_3^2 + x_1x_2 - x_2x_3 - 5x_2 - 10x_3$:

- Determineu i justifiqueu si aquesta funció és estrictament convexa.
- A través de la propietat de l'òptim X^* d'una funció $f(X)$ en un domini, per la qual $\nabla f(X^*)D \geq 0$ per a tota direcció D, a partir de X^*, factible respecte al domini, trobeu (si existeix/teixen) el/els punt/s del domini definit per:

  $$x_1 \geq 0 \quad , \quad x_2 \geq 0 \quad , \quad x_3 \geq 0$$

  que minimitzin la funció.

10. Direccions de descens

Donada la funció $f(x) = x_1^2 + x_2^2 + x_1x_2 - 6x_1$, trobeu (si existeix) el mínim de la funció subjecte a $x_1 \geq 2$, $x_2 \geq 0$.
(Empreu la propietat de l'òptim x^* $\nabla f(x^*)D \geq 0$ per a tota direcció factible D respecte al domini de definició).

1.3 Condició d'òptim per a problemes sense constriccions.

11. Mínim analític*

Trobeu analíticament la solució dels següents problemes d'optimització sense constriccions. Indiqueu si existeix solució, si aquesta és única, i el seu caràcter (minim local, global, estricte,...):
a) $\min f(x) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_2x_3 - x_1x_3$
b) $\min f(x) = e^{x_1 - x_2} + e^{x_2 - x_1} + e^{x_1^2} + x_3^2$
c) $\min f(x) = e^{x_1 - x_2} + e^{x_2 - x_1}$
d) $\min f(x) = e^{x_1 - x_2} + e^{x_1 + x_2}$

12. Punts estacionaris *

Estudieu els següents problemes de minimització sense restriccions: trobeu els punts estacionaris i comproveu si són o no mínims locals fent ús de les condicions suficients i necessàries d'optimalitat.

$$
\begin{array}{ll}
\text{a)} & \min\ x^2 + y^2 - 4x - 8y + xy \\
\text{b)} & \min\ x^2 - y^2 - 4x + 6y - 5 \\
\text{c)} & \min\ -e^{-(x^4+y^4)} \\
\text{d)} & \min\ x^2 - y^4
\end{array}
$$

13. Balanç de masses *

Tenim un procés alimentat amb dos productes d'entrada A i B que donen lloc a un producte de sortida C. El procés es modelitza mitjançant l'equació $M_A + M_B = M_C$ (balanç de masses). Les masses dels fluxos dels productes B i C poden ser determinats mitjançant uns sensors. Per determinar la massa del flux de A, però, no disposem de cap sensor (ens hem volgut estalviar la seva instal·lació), i volem determinar-lo a partir dels valors mesurats per B i C en n instants de temps diferents $M_{B_i}, M_{C_i}, i = 1, ..., n$. Com obtindríem el millor valor de M_A? (entenent per millor, aquell que ens minimitza els errors totals en les n mesures segons l'equació del procés).

14. Resolució del problema de balanç de masses

Solucioneu el problema d'optimització d'una variable que obtinguereu en formular el problema 13 (podeu usar la solució proposada, si voleu). Observeu com la massa òptima M_A obtinguda correspon a la mitjana dels termes $M_{B_i} - M_{C_i}$ (és a dir, $\sum_{i=1}^{n}(M_{B_i}/n - M_{C_i}/n)$).

1.4 Mètodes no derivatius: l'algorisme de Nelder i Mead

15. Casos*

S'aplica la metodologia de Nelder-Mead a minimitzar una funció en un espai de dimensió 2, i tenim un poliedre amb tres vèrtexs x_a, x_b i x_c on la funció objectiu pren respectivament els valors $f(x_a)$=3, $f(x_b)$=4 i $f(x_c)$=5. Un cop realitzada l'operació de reflexió obtenim un nou punt x_R on la funció val $f(x_R)$. Indiqueu els valors que hauria de prendre $f(x_R)$ per tal que el que seguís en l'aplicació de l'algorisme fos cadascuna de les operacions:

- una expansió
- una contracció
- una reducció a partir de x_a
- que prengessim x_R i descartessim x_c.

16. Una iteració

Sigui la funció objectiu $f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2$. En la resolució del problema $\min_{x \in \Re^3} f(x)$ pel métode de Nelder-Mead s'ha arribat a la iteració k-éssima amb el següent poliedre:

$$\begin{array}{rc} & \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ P = & \left[\begin{array}{cccc} 0 & -1.495 & 0.546 & -0.747 \\ 0 & 1.157 & 2.314 & 1.462 \\ 0 & 1.157 & 0.546 & -1.189 \\ 2 & 122.463 & 406.791 & 8.563 \end{array}\right] \\ f(x_j) = & \end{array}$$

Apliqueu una passa del métode de Nelder-Mead a partir de P amb $\alpha = 1$, $\beta = 0.5$ i $\gamma = 2$

1.5 Exploració lineal.

17. Cerca lineal usant el mètode de Fibonacci *

Donada la funció

$$f(x) = x^2 - 3x + 5 + \ln(\frac{1}{x+1})$$

comproveu que aquesta té un únic mínim a l'interval [0,5]. Un cop comprovat això, apliqueu el mètode de Fibonacci per tal de donar un interval de longitud igual o menor a 1.0 dins el qual es trobi aquest valor mínim.

18. I ara usant el mètode de la secció àuria *

Considereu de nou la funció del problema 17 anterior, i el mateix interval inicial [0,5]. Doneu un interval d'incertesa de longitud menor o igual a 1, usant ara el mètode de la secció àuria.

19. I com que no pot haver dos sense tres... *

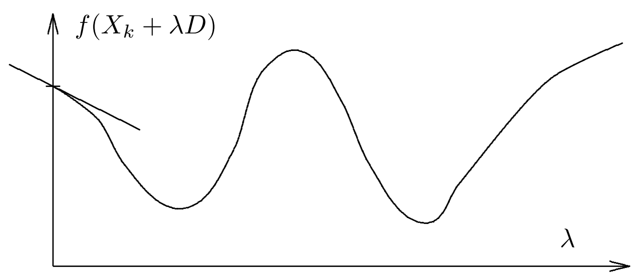
Considereu de nou la funció del problema 17. Trobeu, usant el mètode de Newton primer i el de la secant després, una aproximació (de 6 xifres decimals correctes) del punt òptim. Per al mètode de Newton podeu usar com a punt inicial $x_0 = 4$, mentre que per al de la secant podeu provar amb $x_0 = 4$, $x_1 = 3$.

20. El mètode de Newton no té convergència global *

Un dels desavantatges del mètode de Newton és que no té convergència global (és a dir, podem generar una seqüència de punts que no arribi mai al punt òptim). Comproveu com això succeeix realment amb la funció $f(x) = 2/3\ |x|^{3/2}$, prenent com punt inicial $x_0 = 1$ (aquesta funció te el punt mínim global a $x^* = 0$).

21. Ajustos de corbes i Armijo-Goldstein *

Donada la funció de la figura adjunta, la qual expressa la variació del valor de $f(X_k + \lambda D)$ en funció de λ, sent $X_k, D \in \mathbb{R}^n$ i $f(X)$: $\mathbb{R}^n \rightarrow \mathbb{R}$, copieu la figura i indiqueu les zones de l'eix de les abscisses on es satisfà la $1^{\underline{a}}$ condició d'Armijo-Goldstein i, de forma diferenciada, les zones on es satisfà la $2^{\underline{a}}$ condició d'Armijo-Goldstein. Indiqueu els valors numèrics que utilitzeu per definir els paràmetres que caracteritzen les esmentades condicions.



22. Ajustos de corbes i Armijo-Goldstein *

Donada la funció objeciu $f(x) = x_1^2 + x_1x_2 + \dfrac{2x_1}{x_2}$, el punt iterat $X_0 = [1 \quad -2]'$ i la direcció de cerca $D = [2 \quad 1]'$:

a) Feu exploració lineal a partir de X_0 al llarg de D amb ajust quadràtic emprant el valor de la derivada de la funció objectiu a X_0. Comproveu si el punt obtingut satisfà la 1^a condició d'Armijo-Goldstein.

b) Feu exploració lineal a partir de X_0 al llarg de D amb ajust cúbic emprant el valor de la derivada de la funció objectiu a X_0. Comproveu si el punt obtingut satisfà la 2^a condició d'Armijo-Goldstein. En cas de no fer l'ajust cúbic, comproveu aquesta condició sobre el punt associat a $\lambda = 1/8$.

23. Ajustos de corbes i Armijo-Goldstein *

Per a la funció objectiu $f(x) = x_1^3 + x_1x_2 + 2x_1/x_2$:

a) Verifiqueu si la direcció $D = [0 \quad -1]'$ és de descens per a la funció $f(x)$ a partir del punt inicial $X_0 = [1 \quad 1]'$. Si no ho fos, trobeu-ne una que ho sigui.

b) Partint del punt X_0 donat, efectueu exploració lineal per ajust quadràtic en la direcció de descens que tingueu. Efectueu a continuació exploració lineal per ajust cúbic del tipus $y = a\lambda^3 + b\lambda^2 + c\lambda + d$ emprant els valors de la funció a X_0 i al punt trobat a l'ajust quadràtic i el valor de la derivada a X_0.

c) Efectueu per als punts trobats en les exploracions lineals per ajust quadràtic X_q i ajust cúbic X_c, una comprovació d'acceptabilitat que no sigui la simple verificació de que $f(X_q)$ i $f(X_c)$ valen menys que $f(X_0)$.

1.6 Mètode del Gradient.

24. Direccions perpendiculars al mètode del gradient *

El mètode del gradient genera una seqüència de punts $x_{k+1} = x_k - \alpha \nabla f(x_k)^T$ per tal de minimitzar la funció $f(x)$. En el cas que la funció a minimitzar sigui quadràtica (és a dir, $f(x) = 1/2\ x^T Q x - b^T x$) es pot calcular anàliticament la longitud de pas α, i aquesta ve determinada per $\alpha = (\nabla f(x_k) \nabla f(x_k)^T)/(\nabla f(x_k) Q \nabla f(x_k)^T)$. Per a aquest cas particular, comproveu com dues direccions successives són perpendiculars (és a dir, comproveu que $\nabla f(x_k)$ i $\nabla f(x_{k+1})$ són perpendiculars).

25. Aplicació del mètode del gradient *

Donades les dues funcions següents:

$$f_1(x) = x_1^2 + x_2^2 + x_1 x_2 - x_1 - 2x_2 \qquad\qquad f_2(x) = 5x_1^2 + x_2^2 - 2x_2$$

i) Comproveu que les dues tenen el mateix punt mínim.
ii) Digueu quina funció tindrà una convergència més lenta si apliquem el mètode del gradient. Raoneu la resposta.
iii) Realitzeu dues iteracions del mètode del gradient amb cada funció, prenent en ambdós casos com punt inicial $x_0 = (2 \;\; 2)^T$.
iv) Comproveu que les direccions que heu obtingut (per a cada problema) són perpendiculars.

26. El mètode del gradient amb funcions quadràtiques *

Obtenció d'un punt des d'on el mètode del gradient arriba en una sola passa al mínim d'una funció objectiu quadràtica.

$$\frac{1}{2} X' \begin{bmatrix} 7/2 & 1 & 0 & -1/2 \\ 1 & 11/2 & -1/2 & -2 \\ 0 & -1/2 & 11/2 & -1 \\ -1/2 & -2 & -1 & 15/2 \end{bmatrix} X - [\,21/2 \quad 47 \quad -58 \quad 41/2\,]\, X \quad ; \quad X^* = \begin{bmatrix} 1 \\ 9 \\ -9 \\ 4 \end{bmatrix}$$

Valors propis de Q : $\lambda_i = \{9, 6, 4, 3\}$

27. El mètode del gradient amb funcions quadràtiques *

Realitzeu la primera iteració del métode del gradient aplicat a la resolció del problema:

$$\min\ f(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2 + x_1 x_2 x_3$$

Preneu com a punt inicial $x^0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, com a tolerància de detecció d'òptim $\epsilon = 0.1$ i feu exploració lineal aproximada per Fibonacci amb $N = 4$ i interval d'incertessa inicial [0,1].

28. El mètode del gradient amb funcions qualssevol

Donada la funció $f(x_1, x_2) = 50x_1^2 + 3x_1x_3 + 4x_2^2 + 15x_3^2 - 6x_1 + 8x_2 - 60x_3$ es demana que:

a) apliqueu una passa completa del mètode del gradient per a minimitzar aquesta funció a partir de $X_0 = [0 \quad -1 \quad 0]'$.

b) comproveu, a través de la funció d'error $\epsilon(X) = (X - X^*)'Q(X - X^*)/2$, que la convergència local de l'algorisme entre X_0 i X_1 (que haureu trobat) està dins del límit teòric.

29. Convergència del mètode del gradient

La taula que ve a continuació mostra el valor de la funció objectiu al llarg de tres iteracions, a prop de l'optim, del procés de resolució d'un problema d'optimització no lineal sense constriccions $\min_{x \in \Re^n} f(x)$ que té un minim local amb valor $f(x^*) = 1$:

$f(x^k) =$	0.9991
$f(x^{k+1}) =$	0.999325
$f(x^{k+2}) =$	0.99949375

a) Raoneu, a partir dels valors de la taula, si el mètode d'optimització aplicat pot ser el mètode de Newton.

b) Indiqueu, raonadament, si els valors que s'indiquen a la taula poden correspondre a l'aplicació del métode del Gradient sobre la funció objectiu $f(x) = x_1^3 + x_2^2 - 3x_1 - 2x_2 + 4$

30. Mètode del gradient i exploració lineal de Fibonacci

Realitzeu la primera iteració del métode del gradient aplicat a la resolució del problema:

$$\textbf{(PNL)} \quad \min\ f(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2 + x_1x_2x_3$$

Preneu com a punt inicial $x^0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, com a tolerància de detecció d'òptim $\epsilon = 0.1$ i feu exploració lineal aproximada per Fibonacci amb $N = 4$ i interval d'incertessa inicial [0,1].

31. Mètode del gradient i exploració lineal de Fibonacci*

Realitzeu la primera iteració del mètode del gradient aplicat a la resolució del problema:

$$\textbf{(PNL)} \quad \min\ f(x) = x_1^2 + x_2^2 - (x_2 - x_3)^3 + x_1x_2x_3$$

Preneu com a punt inicial $x^0 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$, com a tolerància de detecció d'òptim $\epsilon = 0.1$ i feu exploració lineal aproximada per Fibonacci amb $N = 4$ a l'interval inicial d'incertessa [0,0.5].

32. Mètode del gradient i exploració lineal de Fibonacci

Considereu el següent problema d'optimització no lineal sense constriccions :

$$\textbf{(PNL)} \quad \min\ f(x_1,x_2) = x_1x_2 + (x_1 - x_2)^3$$

Feu la primera iteració de l'algorisme del gradient aplicat sobre el problema **(PNL)**, prenent com a punt inicial $x^0 = [\,1 \quad 1\,]'$, com a tolerància de detecció d'òptim $\epsilon = 0.5$ i feu exploració lineal aproximada per Fibonacci amb $N = 3$ i interval inicial d'incertesa [0,3].

33. Mètode del gradient i exploració lineal de Fibonacci†

Considereu el següent problema d'optimització no lineal sense constriccions :

$$\textbf{(PNL)} \quad \min\ f(x_1,x_2) = (x_1 - 1)^2 x_2 + (x_2 - x_1)^2$$

Feu la primera iteració de l'algorisme del gradient aplicat sobre el problema **(PNL)**, prenent com a punt inicial $x^0 = [\,1 \quad 2\,]'$, com a tolerància de detecció d'òptim $\epsilon = 0.5$ i feu exploració lineal aproximada per Fibonacci amb $N = 3$ i interval inicial d'incertesa [0,1/2].

34. Mètode del gradient i exploració lineal de Fibonacci

Considereu el següent problema d'optimització no lineal sense constriccions:

$$\textbf{(PNL)} \quad \min\ f(x_1,x_2) = e^{x_1} + e^{x_2} - 3x_1 - 2x_2$$

a) Resoleu el problema **(PNL)** aplicant l'algorisme del gradient amb tolerància de l'òptim $\epsilon = 0.5$, interval inicial d'incertessa [0,1] i exploració lineal de Fibonacci amb N=4. Preneu com a punt inicial $x^0 = [\,1 \quad 1\,]'$

b) Calculeu la fita superior de la taxa de convergència β de l'algorisme del gradient aplicat al problema **(PNL)** i comproveu, a partir dels resultats de l'apartat anterior, que es satisfà.

35. Mètode del gradient i exploració lineal exacta i per Fibonacci*

Donat el següent problema d'optimització no lineal sense constriccions:

$$\textbf{(PNL)} \quad \min f(x) = x_1^2 + \frac{3}{2}x_2^2 + 2x_1x_2 - 2x_1 - 2x_2$$

a) Efectueu una passa completa del mètode del gradient, partint de $x^0 = [\,-1 \quad 0\,]'$, amb tolerància de detecció d'òptim $\epsilon = 0.5$ i fent exploració lineal exacta.

b) Feu exploració lineal de fibonacci a partir del punt $x = [\,-1 \quad 0\,]'$ en la direcció $d = [\,4 \quad 4\,]'$ amb N=4.

36. Mètode del gradient i taxa de convergència†

Considereu el problema **(PNL)** consistent en la minimització de la funció $f(x) = 2x_1^2 - x_2^2 + x_1x_3 - 3x_2 + x_3$.

a) Efectueu una passa del métode del gradient a partir de $x^0 = [\,1 \quad 1 \quad 1\,]'$ amb exploració lineal exacta.

b) Considereu que l'aplicació del mètode del gradient en la resolució de **(PNL)** a partir de x^0 convergeix a un punt extrem x^*. És x^* la solució del problema **(PNL)**? Per què? Quina és la solució del problema **(PNL)**?

c) Calculeu la fita superior a la taxa de convergència β del mètode del gradient aplicat sobre **(PNL)**.

37. Mètode del gradient i taxa de convergència†

Cosidereu els següents dos problemes **(PNS)** de minimització de funcions quadràtiques:

$$\textbf{(PNS1)}\ \min_{x\in\Re^3} f(x) = \frac{1}{2}\,[\,x_1 \quad x_2 \quad x_3\,] \begin{bmatrix} 1 & 1/2 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - [\,3/2 \quad 3/2 \quad 1\,] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\textbf{(PNS2)}\ \min_{x\in\Re^3} f(x) = \frac{1}{2}\,[\,x_1 \quad x_2 \quad x_3\,] \begin{bmatrix} 1 & 1 & 0 \\ 1 & 10 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - [\,2 \quad 11 \quad 1\,] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

a) Efectueu una iteració del mètode del Gradient aplicat sobre el problema **(PNS1)** a partir del punt $x^0 = [\,0 \quad 0 \quad 0\,]'$. Preneu com a tolerància d'optimalitat $\epsilon = 10^{-6}$ i feu exploració lineal per Fibonacci amb $N = 3$ dins l'interval d'incertessa [0,2].

b) Si apliquem el mètode del Gradient als dos problemes a partir del mateix punt, amb quin s'arribaria a la solució òptima en menys iteracions? Justifiqueu la vostra resposta.

1.7 Mètode de Newton.

38. El mètode de Newton aplicat a funcions quadràtiques *

Una de les propietats del mètode de Newton per minimitzar funcions és que, quan s'aplica a funcions quadràtiques, troba el punt òptim en una iteració. Això és degut a que el mètode de Newton es basa en fer una aproximació quadràtica de la funció a minimitzar a cada iteració, trobant aleshores el punt que anul·la el gradient d'aquesta aproximació (és a dir, trobant el punt òptim exacte de l'aproximació). Quan la funció original ja és quadràtica, l'"aproximació" ja no és tal, ja que coincideix amb la funció original. En aquest cas és clar que, quan trobem el punt que anul·la el gradient, estem calculant el punt òptim. Per comprovar experimentalment aquest fet, apliqueu el mètode de Newton a les dues funcions del problema 25 anterior, i observeu com trobeu l'òptim en una iteració.

39. Mètode de Newton, variant de Luenberger *

La funció

$$f(X) = x_1^3 + x_2^2 + \frac{11}{2}x_3^2 + x_4^3 + \frac{5}{2}x_5^2 + x_1x_2^2 + 6x_1x_3 + 3x_1x_5 + 2x_2x_4 + x_2^2x_5 + 6x_3x_5 + x_4x_5$$

té l'hessiana indefinida al punt $X_0 = [\,1/2 \quad 0 \quad -1 \quad 1/2 \quad 0\,]'$

Trobeu la direcció de minimització a partir de X_0 pel mètode de Newton en la variant de Luenberger i comproveu que la direcció P obtinguda és de descens. Efectueu exploració lineal per ajust quadràtic en la direcció P i en la $[2 \quad 0 \quad -1 \quad 0 \quad 0]'$.

40. Mètode de Newton, variant de Luenberger

Efectueu una passa del mètode de Newton en la variant de Luenberger aplicat sobre la funció $f(x) = x_1^2 + x_2^2 + x_1x_2x_3^2$ a partir del punt $X_0 = [0 \quad 1/2 \quad 1]'$

41. Mètode de Newton, variant de Luenberger

Considereu la funció $f(x) = -1/3(x_1 - 3)^3 + (x_2 - 4)^2 + (x_3 - 2)^2 + x_1x_2x_3$ i el punt $x_0 = (2, -3, 0)'$.

a) Indiqueu, raonadament, si seria correcte aplicar una passa del mètode de Newton sense modificar.

b) Calculeu la direcció de descens de $f(x)$ a partir de x_0 pel mètode de Newton en la variant de Luenberger amb $\delta = 0.5$.

b) Si es calcula la direcció de descens pel mètode de Newton sense modificar s'obté la direcció $p = (1.3, 6.4, 0.5)'$, que és de descens. Comenteu aquest fet i la seva relació amb l'aplicació d'un mètode de Newton modificat.

42. Mètode de Newton, variant de Luenberger

Considereu el problema d'optimització no lineal sense constriccions:

$$\textbf{(P)} \qquad \min_{x \in \Re^3} f(x) = e^{2x_1 - x_2} + e^{x_2 - x_1} - x_3^2$$

Es vol resoldre aquest problema a partir del punt inicial $x^0 = [0 \quad 0 \quad 1]'$.

a) Indiqueu, raonadament, si seria correcte aplicar una passa del mètode de Newton sense modificar a partir de x^0.

a) Calculeu la direcció de descens de $f(x)$ a partir de x^0 pel mètode de Newton modificat en la variant de Luenberger amb $\epsilon = 0.5$.

b) Considereu que heu aplicat el mètode de de Newton en la variant de Luenberger al problema **(P)** i n'heu obtingut com a resultat el punt x^*. Indiqueu raonadament si x^* pot ser una solució (mínim local) del problema **(P)**.

43. Mètode de Newton, variant de Gill-Murray *

La funció $f(X) = x_1^3 - 3x_1x_2^2 + 2x_1^2x_4 + x_2x_3x_4 + 6x_2$ té gradient zero i hessiana indefinida en el punt $X_0 = [1 \quad 1 \quad -2 \quad 0]'$.

a) Trobeu una direcció de descens D a partir de X_0 sabent que la triangularització de Gill-Murray de l'hessiana $\nabla^2 f(X_0)$ és:

$$\begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ 0 & 0 & 1 & \\ 2/3 & 1/6 & 6 & 1 \end{bmatrix} \begin{bmatrix} 6 & & & \\ & 12 & & \\ & & 1/6 & \\ & & & 9 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 2/3 \\ & 1 & 0 & 1/6 \\ & & 1 & 6 \\ & & & 1 \end{bmatrix}$$

b) Per al problema anterior, efectueu exploració lineal per ajust cúbic del tipus $y = a\lambda^3 + b\lambda^2 + c\lambda + d$ emprant la informació dels valors de la funció objectiu i el seu gradient als punts X_0 i $X_0 + D$. Indiqueu si observeu alguna circumstància atípica en l'ajust cúbic en la direcció D.

b) Efectueu la mateixa exploració que l'apartat anterior prenent la direcció $P = [\,7/3 \quad 1 \quad 0 \quad 0\,]'$.

44. Mètode de Newton, variant de Dennis-Schnabel

La funció:

$$\begin{aligned} f(x) = & x_1^3 + x_2^2 + (11/2)x_3^2 + x_4^3 + (5/2)x_5^2 + x_1x_2^2 + 6x_1x_3 + 3x_1x_5 + \\ & + 2x_2x_4 + x_2^2x_5 + 6x_3x_5 + x_4x_5 \end{aligned}$$

té l'hessià indefinit al punt $X_0 = [1/2, 0, -1, 1/2, 0]'$. Trobeu la direcció de minimització a partir de X_0 pel mètode de Dennis-Schnabel, sabent que la triangularització de Gill-Murray de l'hessià a X_0 afegeix a l'hessià la matriu:

$$E = \begin{pmatrix} 0.272727 & & & & \\ & 0 & & & \\ & & 0.022727 & & \\ & & & 0 & \\ & & & & 18.7 \end{pmatrix}$$

45. Mètode de Newton, variant de Dennis-Schnabel

Considereu al següent problema d'optimització sense constriccions:

$$\min f(x) = x_1^2 + 2x_2^2 + 3x_1x_2 + 3x_2x_3^2 + x_1x_2x_3$$

La triangularització de Gill-Murray de $\nabla^2 f(x)$ sobre el punt $x^0 = [-1, 1, 0]'$ és:

$$\begin{pmatrix} 1. & 0 & 0 \\ 1.5 & 1. & 0 \\ 0.5 & -2.4 & 1. \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1.0417 & 0 \\ 0 & 0 & 0.5 \end{pmatrix} \begin{pmatrix} 1. & 1.5 & 0.5 \\ 0 & 1. & -2.4 \\ 0 & 0 & 1. \end{pmatrix}$$

a) Calculeu la direcció de descens de f(x) a partir de x^0 segons el mètode de Dennis-Schnabel.

b) Considereu ara una nova funció objectiu $\tilde{f}(x)$ obtinguda afegint a $f(x)$ els termes $-x_1 - x_2 + x_3$, és a dir:

$$\tilde{f}(x) = x_1^2 + 2x_2^2 + 3x_1x_2 + 3x_2x_3^2 + x_1x_2x_3 - x_1 - x_2 + x_3$$

Calculeu una direcció de descens de $\tilde{f}(x)$ a partir de x^0. Comproveu que la direcció obtinguda és de descens.

46. Mètode de Newton, variant de Dennis-Schnabel

Considereu al següent problema d'optimització sense constriccions:

$$\min_{x \in \Re^3} f(x) = x_1^2 + 3x_2^2 + x_1x_2/x_3$$

La correcció diagonal de la triangularització de Gill-Murray de la Hessiana de $f(x)$ sobre el punt $X_k = [1/4, 1/5, -1/2]'$ és:

$$E = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 3.86 \end{pmatrix}$$

Efectueu una iteració del mètode de Newton modificat en la variant de Dennis-Schnabel. Feu exploració lineal per ajust quadràtic entre $\lambda = 0$ i $\lambda = 1$. Preneu com a criteri d'acceptabilitat de la longitud de pas la primera condició d'Armijo-Goldstein amb $\alpha = 0.1$.

2 Problemes de PNL amb constriccions.

2.1 Condicions d'òptim per a problemes amb constriccions.

47. Corbes diferenciables*.

Considereu la constricció no lineal d'igualtat $h(x) = 0$ amb $h(x) = x_1^2 - x_2$ i la funció objectiu $f(x) = e^{x_1(x_2+1)}$. Sigui $\mathcal{S}$ la superfície de $\mathbb{R}^2$ definida per $h(x) = 0$ i $x(t)$ la corba diferenciable de $\mathcal{S}$ que té $x_1(t) = t$.

a) A quín punt de $\mathbb{R}^2$ correspon $x(0)$?

b) Calculeu el valor de $\dfrac{d}{dt} f(x(t))$ per a $t = 0$.

c) A la vista del resultat anterior, pot ser $x(0)$ un punt estacionari del problema $\begin{cases} \min f(x) \\ \text{s. a: } h(x) = 0 \end{cases}$?

48. Formulació de les condicions de Kuhn i Tucker*

Formuleu les condicions necessàries de primer ordre de mínim local del següent problema d'optimització no lineal :

$$\begin{aligned} &\min f(x_1, x_2) \\ \text{s.a. :}\, &g_1(x_1, x_2) \leq b_1 \\ &g_2(x_1, x_2) \geq b_2 \\ &x_1, x_2 \geq 0 \end{aligned}$$

amb $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $g_1: \mathbb{R}^2 \rightarrow \mathbb{R}$ i $g_2: \mathbb{R}^2 \rightarrow \mathbb{R}$

49. Condicions de Kuhn i Tucker*

Considereu el següent problema de programació no lineal:

$$
\textbf{(PNL)}\begin{cases}
\min \quad f(x) = (x_1-1)^2 + (x_2-2)^2 \\
\text{subj.a}: \quad -x_1 + x_2 \geq -1 \\
\qquad\qquad x_1 + x_2 \leq 2 \\
\qquad\qquad x_1 \, , \, x_2 \geq 0
\end{cases}
$$

a) Suposem que s'ha trobat una solució factible **(PNL)** que satisfà les condicions necessàries de primer ordre (condicions de Kuhn i Tucker). Què podem dir sobre el caràcter de mínim local o global d'aquesta solució? Justifiqueu la vostra resposta.

b) Formuleu les condicions de Khun i Tucker del problema **(PNL)**.

c) Resoleu el problema **(PNL)** usant les condicions de Kuhn i Tucker, sabent que a l'òptim es satisfà que $x_1 > 0$ i $x_2 > 0$.

50. Condicions de Kuhn i Tucker*

Considereu el següent problema de programació no lineal:

$$
\textbf{(PNL)}\begin{cases}
\min \quad f(x) = (x_1x_2-3)^2 - 4x_1x_2 + 2(x_1x_3-1)^3 - \frac{15}{2}x_2 + 5x_3 \\
\text{s.a}: \quad 6x_1^2 - 3x_2x_3 - x_3^2 - 5 \leq 0 \\
\qquad\quad x_1^2 + \frac{1}{2}x_2 + x_3^2 - 4 \leq 0
\end{cases}
$$

a) Formuleu les condicions necessàries de primer ordre de mínim local (condicions de de Karush-Khun-Tucker) pel problema **(PNL)**.

b) Comproveu si el vector $x' = [\,0 \quad -2 \quad 1\,]$ pot ser un mínim local del problema **(PNL)**.

51. Condicions de Kuhn i Tucker

Considereu el següent problema de programació no lineal amb constriccions :

$$
\begin{array}{lll}
\textbf{(PNL)} & \min & f(x_1,x_2) = x_1x_2 + (x_1-x_2)^3 \\
 & \text{s.a}: & (x_1+1)(x_2+1) \leq 3 \\
 & & x_1 \geq 0 \, , \, x_2 \geq 0
\end{array}
$$

a) Comproveu, a partir de les condicions de Kuhn-Tucker si el punt $x^0 = [\,1 \quad 1/2\,]'$ és un mínim local del problema **(PNL)**.

b) Tenint en compte la regió factible de **(PNL)**, podem dir que es tracta d'un problema de programació convexa? Justifiqueu la vostra resposta.

52. Condicions de Kuhn i Tucker i programació convexa*

Considereu el següent problema de programació no lineal amb constriccions :

$$
\textbf{(PNL)}\begin{cases} \min & f(x_1,x_2)=(x_1-1)^2x_2+(x_2-x_1)^2 \\ \text{s.a}: & (x_1+1)(x_2+1)+2x_1^2+2x_2^2\le 4 \\ & x_1\ge 0\ ,\ x_2\ge 0 \end{cases}
$$

a) Formuleu les condicions de Kuhn-Tucker pel problema **(PNL)**.
b) Comproveu, a partir de les condicions de Kuhn-Tucker si el punt $x^0=[\,0 \quad 1\,]'$ és un mínim local del problema **(PNL)**.
c) És **(PNL)** un problema de programació convexa? Justifiqueu la vostra resposta.

53. Condicions de Kuhn i Tucker i pla tangent*

Considereu el problema d'optimització no lineal següent :

$$
\textbf{(PNL)}\begin{cases} \min & f(x)=e^{x_1(x_2+1)} \\ \text{s.a.} & x_1^2-x_2\le 0 \\ & 2x_1^3-2x_1^2-x_2+1\le 0 \end{cases}
$$

a) Formuleu les condidicions necessàries de primer ordre de **(PNL)**. Indiqueu el procediment que s'hauria de seguir si es volgués resoldre **(PNL)** a partir d'aquestes condicions. Quina seria la dificultat més rellevant d'aquest procés?.
   Considereu que GINO mostra la següent informació associada al punt $\tilde{x}$:

```
              OBJECTIVE FUNCTION VALUE

        1)           .529462

  VARIABLE        VALUE          REDUCED COST
        X1         -.455418          .000000
        X2          .396285         -.000007

       ROW   SLACK OR SURPLUS          PRICE
        2)          .188880          .000000
        3)         -.000009         -.241119
```

b) Usant la informació proporcionada per GINO, indiqueu si el punt $\tilde{x}=[\,-.455418 \quad .396285\,]'$ és un mínim local de **(PNL)**.
c) Trobeu una base del subespai tangent de les constriccions actives sobre $\tilde{x}$.

54. Condicions de segon ordre

Considereu el següent problema de programació no lineal amb constriccions :

$$
\textbf{(PNL)}\begin{cases} \min & f(x_1,x_2)=(x_1-1)^2x_2+(x_2-x_1)^2 \\ \text{s.a}: & (x_1+1)(x_2+1)+2x_1^2+2x_2^2\le 4 \\ & x_1\ge 0\ ,\ x_2\ge 0 \end{cases}
$$

a) Formuleu les condicions de Kuhn-Tucker pel problema **(PNL)**.
b) Comproveu, a partir de les condicions de Kuhn-Tucker si el punt $x^0 = [\,0 \quad 1\,]'$ és un mínim local del problema **(PNL)**.
c) És **(PNL)** un problema de programació convexa? Justifiqueu la vostra resposta.

55. Condicions de segon ordre*

Considereu el problema d'optimització no lineal següent :

$$
\textbf{(PNL)} \begin{cases} \min & f(x) = e^{x_1(x_2+1)} \\ \text{s.a.} & x_1^2 - x_2 \geq 0 \\ & 2x_1^3 - 2x_1^2 - x_2 + 1 \geq 0 \end{cases}
$$

Comproveu si el punt $x^* = [\,-1/2 \quad 1/4\,]$ és solució del problema **(PNL)**.

56. Condicions de segon ordre*

Considereu el segënt problema d'optimització amb constriccions:

$$
\textbf{(PNL)} \quad \begin{cases} \min\limits_{x \in \mathbb{R}^3} & f(x) = \dfrac{x_1 x_2}{x_3 + 1} \\ \text{subj. a :} & h(x) = x_1^2 + x_2^2 - x_3 - 1 = 0 \end{cases}
$$

La solució $x^* = [\,1 \quad 1 \quad 1\,]'$ és un punt estacionari de **(PNL)** amb multiplicador de Lagrange associat $\lambda^* = [\,-1/4\,]$. Comproveu si x^* és màxim, mínim o punt de sella de **(PNL)**.

57. Condincions de mínim i GINO

S'ha definit el següent problema de programació no lineal dins del paquet GINO:

```
MODEL:
   1) MIN= X1 ^ 2 - X1 / ( X2 + 1 ) ;
   2) X1 - EXP( X1 - X2 ) = 1 ;
   3) X1 - X1 * X2 < 0 ;
END
```

i el resultat proporcionat pel programa és:

```
SOLUTION STATUS:  OPTIMAL TO TOLERANCES.  DUAL CONDITIONS:  SATISFIED.

              OBJECTIVE FUNCTION VALUE

         1)          .866781

    VARIABLE         VALUE          REDUCED COST
         X1        1.013351             .000000
         X2        5.329532             .000008

        ROW   SLACK OR SURPLUS           PRICE
         2)          .000000         -1.894009
         3)         4.387335            .000000
```

a) Coneixent el valor de les variables a l'òptim x^* proporcionat per GINO a la taula anterior, calculeu, usant les condicions de Karush-Kuhn-Tucker, el valor dels multiplicadors de Lagrange associats a les constriccions del problema.
b) Useu les condicions de segon ordre de mínim local per a determinar si el punt x^* calculat per GINO és efectivament mínim local del problema plantejat.

58. Assumpció de complementarietat estricte

Considereu el següent problema d'optimització no lineal de dos variables:

$$\textbf{(PNL)} \quad \begin{cases} \min\limits_{x \in \mathbb{R}^2} & f(x) = \dfrac{1}{2}(x_1^2 - x_2^2) \\ \text{subj. a :} & x_2 \leq 0 \end{cases}$$

Comproveu que $x^* = [\,0 \quad 0\,]'$ i $\mu^* = 0$ satisfan les condicions suficients de segon ordre de mínim local, excepte l'assumpció de complementarietat estricte. És x^* un mínim local de **(PNL)**?

59. Àrea de dos quadrats *

Volem determinar l'àrea mínima de dos quadrats, els costats dels quals tenen una longitud de x_1 i x_2 respectivament, de forma que la suma d'aquestes longituds $(x_1 + x_2)$ sigui igual a un determinat valor a. Trobeu els valors x_1 i x_2 que minimitzen l'àrea que ocuparan els dos quadrats. Feu-ho de dues formes: i) primer considerant un problema sense restriccions (fent el canvi de variable $x_2 = a - x_1$); ii) en segon lloc, aplicant el mètode dels multiplicadors de Lagrange. Observeu com en ambdós casos s'obté el mateix resultat. (Nota: no cal afegir les restriccions de no-negativitat $x_1 \geq 0, \quad x_2 \geq 0$, ja que, com s'observarà, són inactives al punt òptim, i d'aquesta forma simplifiquem el problema).

60. Volum màxim d'un cilindre *

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

i) Plantegeu el problema amb restriccions que cal solucionar.
ii) Solucioneu el problema usant el mètode dels multiplicadors de Lagrange. Podeu eliminar les restriccions de no-negativitat sobre les dimensions del cilindre, ja que a l'òptim no són actives, com podreu observar. D'aquesta forma evitem treballar amb restriccions de desigualtat, tot simplificant el problema.
iii) Solucioneu el problema amb l'ajut del paquet Lingo (o un altre d'equivalent), considerant ara les restriccions de no-negativitat de les variables. Comenteu el que observeu (Nota: la sortida de Lingo —i de la majoria de paquets d'optimització— a part del valor de les variables i restriccions, dóna també els valors dels multiplicadors λ de les restriccions d'igualtat i μ de les de desigualtat. Lingo anomena als multiplicadors "dual price", i es troben a la vora del valor de cada restricció).

61. Obtenció de les condicions d'òptim al mètode del símplex *

(Nota: aquest problema comporta una certa dificultat. Si teniu dubtes, consulteu la solució proposada).

El problema de programació lineal $\{\min \ c^T x \ \text{ subj. a } Ax = b, \ x \geq 0\}$ (suposarem que A

és de rang complet, és a dir, les restriccions són linealment independents) pot ser solucionat mitjançant l'algorisme del símplex, el qual es basa en l'obtenció de successives solucions bàsiques $x^T = (x_B^T \;\; x_N^T)$, $x_B \in \mathbb{R}^m, x_N \in \mathbb{R}^{n-m}$ fins a obtenir una d'òptima. Considerant una solució bàsica qualsevol no degenerada (és a dir, $x_B > 0$ i $x_N = 0$), deduiu, usant les condicions de Kuhn-Tucker, les condicions necessàries que hauria de satisfer si fos un punt òptim, i comproveu que equivalen a les ja conegudes del mètode del símplex.

62. La funció de cost d'un reactor *

S'ha estimat que el cost diari (comptabilitzant amortització, manteniment, etc.) d'un tipus de reactor és funció d'un grau de conversió que oscil·la entre 0 i 1 (com més proper a 1, millor és la qualitat del reactor) i del seu volum (en unes unitats qualsevols). Aquesta funció de cost és $cost(x, y) = e^x y^4$, on x és el grau de conversió i y és el volum. Per altra banda, el benefici del producte obtingut diàriament també és funció d'aquests dos paràmetres i ve donat per $guanys(x, y) = 100y^2(1 + x)$. Plantegem el següent problema per obtenir els valors de x i y (qualitat i volum del reactor) que ens permeten obtenir el màxim benefici amb el mínim cost:

$$\begin{aligned} \min_{x,y} \quad & e^x y^4 - 100y^2(1+x) \\ & 0 \leq x \leq 1 \\ & 0 \leq y \end{aligned}$$

Responeu les següents qüestions:

a) És convexa la funció objectiu del problema? Podem, doncs, garantir que l'òptim trobat és un òptim global?

b) Considerant que estem al punt $(x, y) = (0, 1)$, comproveu que la direcció de moviment $(d_x, d_y) = (1, -0.5)$ és de descens (és a dir, que verifica la condició de descens).

c) Podem escriure el problema de la següent forma equivalent

$$\min_{x,y} \quad e^x y^4 - 100y^2(1+x) \quad \text{subjecte a} \quad x \leq 1, \;\; -x \leq 0, \;\; -y \leq 0$$

tenint ara tres restriccions de desigualtat. Sabent que a l'òptim $x > 0$ i $y > 0$ (és a dir, la 2a i 3a restriccions són inactives), trobeu quin punt és candidat a ser solució del problema (qualitat i volums òptims). **Nota:** com que únicament heu d'indicar quin punt és candidat, només heu d'utilitzar les condicions necessàries. I d'aquestes, considereu només les de 1r ordre per simplificar la resolució.

2.2 Optimització amb constriccions lineals

63. Aplicació del mètode del gradient reduït *

Donat el problema de programació no lineal (amb funció quadràtica i restriccions lineals)

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

realitzeu els següents apartats:

i) Comproveu que el punt $x^* = (3\ 0\ 2\ 5)^T$ i els multiplicadors $\lambda^* = (-1\ 0)^T$ 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, amb el qual podem concloure que $x^* = (3\ 0\ 2\ 5)^T$ és l'òptim del nostre problema.

ii) Feu dues iteracions del mètode del gradient reduït començant a iterar des del punt $x^0 = (0\ 2\ 3\ 1)^T$.

iii) Feu una iteració del mètode del gradient reduït començant a iterar des del punt $x^0 = (1\ 1\ 3\ 2)^T$, i considerant que x_1 i x_2 són les variables bàsiques (o dependents).

64. Transport d'un fluid *

Disposem d'una xarxa de transport d'un fluid tal i com mostra la Fig. 1.

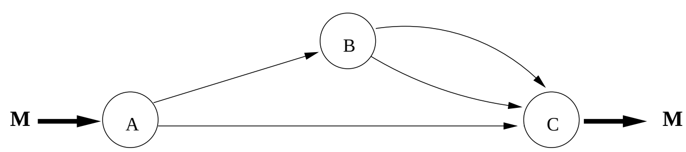


Figura 1. Xarxa de transport d'un fluid

Al node A s'injecta un flux de M lb$_m$/s, el qual ha de sortir pel node C. Suposem que el cost de transport de x lb$_m$/s per una canonada és de x^2 mils de pts, i que la capacitat de cada canonada és de M lb$_m$/s. Contesteu les següents qüestions

a) Plantegeu el problema que cal solucionar per trobar la política de distribució òptima per les quatre canonades.

b) Raoneu per què podem eliminar una restricció i els límits de les variables. Escriviu la nova formulació havent eliminat la restricció associada al node C, i sense límits a les variables.

c) Solucioneu el problema obtingut a l'apartat b).

d) Podem garantir que l'òptim trobat és un mínim global? Raoneu-ho.

e) Realitzeu dues iteracions del mètode del gradient reduït amb el problema formulat a l'apartat b), considerant ara els límits inferiors de les variables $x \geq 0$, i que $M = 10$ lb$_m$/s. Preneu al punt inicial els fluxos entre els nodes (A,C) i (A,B) com variables dependents (o bàsiques). Considereu que en aquest punt inicial els fluxos actuals factibles són de 10 per a la canonada que comunica A i C, i de 0 per a la resta (obvieu el fet que la base inicial és degenerada).

65. Algorisme de Murtagh i Saunders *

Donada la minimització:

$$\min\ f(X) = \frac{x_1 {x_3}^2}{2} - x_1x_5 + 3x_1 - 6x_2 - 4x_3 - 4x_4x_6 + {x_6}^2 x_7 x_8$$
$$\text{subj.}\ \begin{bmatrix} 3 & 1 & 0 & -3 & 3 & 0 & 1 & 3 \\ -6 & 1 & 1 & 8 & 6 & 4 & 0 & -3 \\ 3 & -1 & 0 & -3 & 2 & -3 & 1 & 0 \end{bmatrix} X = \begin{bmatrix} 23 \\ 41 \\ 9 \end{bmatrix}$$
$$\underline{0} \le X \le \overline{X} \quad \text{amb} \quad \overline{X} = [\,5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5\,]'$$

determineu el gradient projectat pel procediment de Murtagh-Saunders a partir del punt factible $X_k = [\,2 \quad 0 \quad 1 \quad 3 \quad 5 \quad 1 \quad 5 \quad 2\,]'$, reordenant, si cal, les variables i comproveu que la direcció de menys el gradient projectat a partir de X_k és una direcció factible de descens.

66. Algorisme de Murtagh i Saunders *

En la minimització de:

$$\min\ f(X) = \frac{{x_1}^2 x_5}{10} + x_2x_3 - 3x_4{x_8}^2 + 2x_5x_7 - 2x_6 + 3x_8$$
$$\text{subj.}\ \begin{bmatrix} 3 & 1 & 0 & -3 & 3 & 0 & 1 & 3 \\ -6 & 1 & 1 & 8 & 6 & 4 & 0 & -3 \\ 3 & -1 & 0 & -3 & 2 & -3 & 1 & 0 \end{bmatrix} X = \begin{bmatrix} 31 \\ 92/3 \\ 11 \end{bmatrix}$$
$$\underline{0} \le X \le \overline{X} \quad \text{amb} \quad \overline{X} = [\,5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5\,]'$$

pel procediment de Murtagh-Saunders, s'ha trobat que en el punt X_0=$[\,3 \quad 0 \quad 3 \quad \frac{1}{3} \quad 5 \quad 4 \quad 5 \quad 1\,]'$ i partint de la subdivisió en variables bàsiques: X_B=$\{x_1, x_3, x_6\}$, superbàsiques: X_S=$\{x_4, x_8\}$ i no bàsiques: X_N=$\{x_2, x_5, x_7\}$, el gradient projectat val zero: $Z'\nabla f(X_0)'$=$\underline{0}$. Indiqueu si X_0 pot ser un òptim, i si no ho podés ser, quines variables haurien de canviar de valor i com calcularieu el seu canvi.

67. Algorisme de Murtagh i Saunders

Considereu el següent problema d'optimització no lineal:

$$\textbf{(P)} \quad \begin{cases} \displaystyle\min_{x \in \Re^7} & f(x) \\ \text{subj. a :} & Ax = b \\ & [\,0\,] \le x \le \bar{x} \end{cases}$$

amb:

$$f(x) = 3x_1^2 + x_2^2 - 4x_3^2 - 2x_4^2 + x_5^2 - 2x_6^2 + 3x_7^2 + x_1x_2 + x_3x_4 - x_3x_7 + 4x_5 - 2x_6$$
$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 2 & -3 \\ -1 & 0 & 4 & -2 & 2 & 4 & -4 \\ 0 & -3 & 0 & 1 & 3 & -3 & -2 \end{bmatrix} \quad ; \quad \bar{x}' = [\,3 \quad 2 \quad 3 \quad 2 \quad 3 \quad 3 \quad 2\,]$$

Durant la resolució d'aquest problema amb l'algorisme de Murtagh-Saunders s'ha obtingut l'iterat $x^{k'} = [\,2 \quad 1 \quad 2 \quad 2 \quad 2 \quad 2 \quad 0\,]$ amb la partició de variables següent: ${x_{\mathcal{B}}}^{k'} = [\,x_1 \quad x_2 \quad x_5\,]$, ${x_{\mathcal{S}}}^{k'} = [\,x_3 \quad x_6\,]$, ${x_{\mathcal{N}}}^{k'} = [\,x_4 \quad x_7\,]$.

a) Realitzeu una nova iteració de l'algorisme Murtagh-Saunders a partir de x^k.

b) És x^k un minim local de problema? En cas que no ho sigui, podrieu indicar algún procediment que permetés trobar una direcció de descens i continuar iterant?

68. Algorisme de Murtagh i Saunders

Considereu el següent problema d'optimització no lineal:

$$
\textbf{(P)} \quad \left\{ \begin{array}{ll} \displaystyle\min_{x \in \Re^5} & f(x) = \dfrac{x_1 x_2}{x_3 + 1} - x_4 x_5 \\ \text{subj. a :} & \begin{bmatrix} 2 & 4 & 4 & 0 & 3 \\ 1 & 9 & 2 & 5 & 0 \end{bmatrix} x = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ & [\,0\,]' \le x \le \bar{x} = [\,4 \quad 4 \quad 3 \quad 4 \quad 4\,] \end{array} \right.
$$

Aplicant la fase I del símplex s'ha obtingut la solució inicial factible $x^0 = [\,1 \quad 4 \quad 3 \quad 0 \quad 2\,]'$. Efectueu la primera iteració de l'algorisme de Murtagh i Saunders aplicat sobre **(P)** a partir de x^0. En concret:

a) Determineu els conjunts $\mathcal{B}^0$, $\mathcal{S}^0$ i $\mathcal{N}^0$.

b) Comproveu la condició d'aturada de l'algorisme sobre x^0, realitzant, si cal, les actualitzacions de $\mathcal{B}^0$, $\mathcal{S}^0$ i $\mathcal{N}^0$.

c) Calculeu la direcció de descens.

d) Efectueu l'actualització de les variables prenent com a passa òptima α^* la passa màxima permessa $\bar{\alpha}$, indicant com s'actualitzarien les llistes $\mathcal{B}^0$, $\mathcal{S}^0$ i $\mathcal{N}^0$.

69. Algorisme de Murtagh i Saunders

Considereu el següent problema d'optimització no lineal amb constriccions lineals:

$$
\textbf{(NCL)} \left\{ \begin{array}{lcc} \min & f(x) & \\ \text{subj. to :} & Ax = b & ; \\ & 0 \le x \le u & \end{array} \begin{array}{c} f\colon \mathbb{R}^n \to \mathbb{R}\,,\ f \in \mathcal{C}^2 \\ A \in \mathbb{R}^{m \times n}\,,\ \text{rang}(A) = m \\ x, u \in \mathbb{R}^n\,,\ b \in \mathbb{R}^m \end{array} \right\}
$$

En la resolució d'aquest problema mitjançant l'algorisme de Murtagh i Saunders s'ha arribat a una solució factible X^k amb gradient reduït nul (${g_z}^k = 0$). Un cop avaluats els multiplicadors de Lagrange $\sigma_{\mathcal{N}}$ s'observa que la i-èssima variable no bàsica $x_{\mathcal{N}i} = 0$ té multiplicador de Lagrange associat $\sigma_{\mathcal{N}i} < 0$. Sigui p la direcció definida per la relació $\bar{A}^k p = e_{m+i}$, amb $\bar{A}^k$ matriu de constriccions actives a X^k i $e_{m+i} \in \mathbb{R}^{n-s}$ vector unitari amb $+1$ a la posició $m+i$. Demostreu que la direcció p és factible i de descens.

70. Algorisme de Murtagh i Saunders

Considereu el següent problema d'optimització no lineal:

$$
\textbf{(P)} \quad \begin{cases} \displaystyle\min_{x\in\Re^4} & f(x) = -x_1^2 + x_3x_2^2 - x_1x_4 \\ \text{subj. a :} & \begin{bmatrix} 2 & 2 & 1 & 4 \\ 3 & 0 & 0 & -1 \end{bmatrix} x = \begin{bmatrix} 10.75 \\ 1.25 \end{bmatrix} \\ & [\,0\,]' \le x \le \bar{x} = [\,2 \quad 2 \quad 2 \quad 2\,] \end{cases}
$$

a) Apliqueu dues iteracions de l'algorisme de Murtagh i Saunders a partir del punt $X_k = [1, 0, 1.75, 1.75]'$ amb la següent partició de variables: $X_B = \{x_1, x_3\}$, $X_S = \{x_4\}$ i $X_N = \{x_2\}$. Preneu com a longitud de pas òptima la longitud de pas màxima permessa.

a) Estudieu el càracter de mínim local de **(P)** del punt obtingut a l'apartat anterior.

71. Algorisme de Murtagh i Saunders

Considereu el problema d'optimització no lineal amb constriccions lineals definit per la funció objectiu $f(x) = x_1^2 + x_2^2 + x_3^2$ i el fitxer MPS:

```
NAME          PROVA
ROWS
 G  CONSLIN1
 G  CONSLIN2
 L  CONSLIN3
COLUMNS
    X1        CONSLIN1  2.0
    X1        CONSLIN2  1.0
    X2        CONSLIN1  4.0
    X2        CONSLIN2  1.0
    X2        CONSLIN3  2.0
    X3        CONSLIN1  -1.0
    X3        CONSLIN2  1.0
    X3        CONSLIN3  1.0
RHS
    TERMINDE  CONSLIN1  1.
    TERMINDE  CONSLIN2  1.
    TERMINDE  CONSLIN3  2.
BOUNDS
 LO LIMSIMP   X1        0.
 LO LIMSIMP   X2        0.
 UP LIMSIMP   X3        1.
ENDDATA
```

En resoldre aquest problema amb el paquet MINOS, la solució obtinguda després de tres iteracions es la que reflecteix el següent llistat:

```
 EXIT -- TOO MANY ITERATIONS
 No. of iterations                  3     Objective value      3.5595567867E-01
 No. of major iterations            1     Linear objective     0.0000000000E+00
 Penalty parameter         100.000000     Nonlinear objective  3.5595567867E-01
 No. of calls to FUNOBJ            11     No. of calls to FUNCON              0
 No. of superbasics                 2     Norm of reduced gradient    3.684E-01
 No. of basic nonlinears            2     Norm RG / Norm PI           3.684E-01
 No. of degenerate steps            0     Percentage                       0.00
 Norm of X                  1.654E+00     Norm of PI                  1.000E+00
 Norm of X  (unscaled)      1.654E+00     Norm of PI (unscaled)       1.000E+00
1
 NAME           PROVA                 Objective value       3.5595567867E-01
 Status         EXCESS ITNS           Iteration      3    Superbasics      2
 OBJECTIVE      FUNOBJ   (MIN)
 RHS            TERMINDE
 RANGES
 BOUNDS         LIMSIMP

 SECTION 1 - ROWS
 NUMBER ...ROW..STATE...ACTIVITY...SLACK ACTIVITY..LOWER LIMIT...UPPER LIMIT..DUAL ACTIVITY..I
      4 CONSLIN1  SBS    1.23684        -0.23684      1.00000         NONE         0.00000  1
      5 CONSLIN2   LL    1.00000         0.00000      1.00000         NONE         0.78947  2
      6 CONSLIN3   BS    0.81579         1.18421        NONE       2.00000         0.00000  3
1
 SECTION 2 - COLUMNS
 NUMBER .COLUMN.STATE...ACTIVITY....OBJ GRADIENT...LOWER LIMIT...UPPER LIMIT. REDUCED GRADNT M+J
      1 X1         BS    0.39474         0.78947      0.00000         NONE          0.00000   4
      2 X2      N SBS    0.21053         0.42105      0.00000         NONE         -0.36842   5
      3 X3         BS    0.39474         0.78947      0.00000       1.00000         0.00000   6
 FUNOBJ called with NSTATE =   5
 ENDRUN
```

a) Indiqueu, a partir de la informació continguda en el llistat de sortida de MINOS, i sense fer cap operació, el valor dels seguents paràmetres:
   - Conjunts $\mathcal{B}$, $\mathcal{S}$ i $\mathcal{N}$.
   - Matrius B, S i N.
   - Gradient reduit g_z.
   - Multiplicadors π i σ. itemIndiqueu clarament on es troba aquesta informació al llistat de MINOS. Si algún dels paràmetres demanats no els sabeu localitzar al llistat, calculeu-los. itemb)Realitzeu una iteració de l'algorisme de Murtagh & Saunders a partir de l'iterat indicat en el llistat de sortida anterior. Feu exploració lineal exacta i useu com a tolerància d'òptim del gradient reduit el valor $\epsilon = 0.1$.

72. Algorisme de Murtagh i Saunders

Considereu el problema d'optimització no lineal amb constriccions lineals definit per la funció objectiu $f(x) = x_1^2 - x_1 x_2^2 + x_4 x_5^2$ i el fitxer MPS:

```
NAME          PROVA
ROWS
 E  CONSLIN1
 E  CONSLIN2
 E  CONSLIN3
COLUMNS
    X1        CONSLIN1  1.0
    X1        CONSLIN3  1.0
    X2        CONSLIN2  2.0
    X2        CONSLIN3  3.0
    X3        CONSLIN1  2.0
    X3        CONSLIN3  2.0
    X4        CONSLIN1  1.0
    X4        CONSLIN2  2.0
    X4        CONSLIN3  4.0
    X5        CONSLIN2  2.0
RHS
    TERMINDE  CONSLIN1  4.
    TERMINDE  CONSLIN2  6.
    TERMINDE  CONSLIN3  10.
BOUNDS
 LO LIMSIMP   X1        0.
 LO LIMSIMP   X2        0.
 LO LIMSIMP   X3        0.
 LO LIMSIMP   X4        0.
 LO LIMSIMP   X5        0.
ENDDATA
```

Considereu el punt iterat $x^k = [\,1 \quad 1 \quad 1 \quad 1 \quad 1\,]'$.

a) Indiqueu raonadament si el punt x^k pot ser solució del problema plantejat

b) Realitzeu una iteració de l'algorisme de Murtagh & Saunders a partir de x^k. Preneu com a longitud de pas òptima la meitat de la longitud màxima permessa ($\alpha^{*k} = \frac{1}{2}\bar{\alpha}$) i useu $\epsilon = 10^{-3}$ com a tolerància a l'òptim.

c) Indiqueu com s'actualitzarien els conjunts $\mathcal{B}^k$, $\mathcal{S}^k$ i $\mathcal{N}^k$ si, a la iteració de l'apartat anterior, es pren $\alpha^{*k} = \bar{\alpha}$ en lloc de $\frac{1}{2}\bar{\alpha}$

2.3 Optimització amb constriccions qualssevol

73. Gradient reduït generalitzat*

Considereu el següent problema d'optimització amb constriccions:

$$
\textbf{(PNL)}\left\{\begin{array}{rl}
\min & \dfrac{x_1 x_3^2}{2} - x_1 x_5 + 3x_1 - 6x_2 - 4x_3 - 4x_4 \\
\text{subj. a:} & \dfrac{x_1^2}{2} + x_2 - x_3 + 2x_4 x_5 - x_5 = -2 \\
& \dfrac{3x_1^2}{2} + x_2 x_3 - 2x_2 + \dfrac{3x_4^2}{2} + x_5 = \dfrac{7}{2} \\
& \\
& \mathbf{0} \leq X \leq [\,4 \quad 4 \quad 4 \quad 4 \quad 4\,]'
\end{array}\right.
$$

Efectueu un pas complet del mètode del gradient reduït generalitzat a partir del punt $X_0 = [\,1 \quad 0 \quad 3 \quad 1 \quad 1/2\,]'$, trobant: el gradient reduït, la direcció d'exploració, la passa màxima i, prenent com a passa òptima la meitat de la passa màxima, un nou punt. A continuació efectueu una passa de projecció del nou punt sobre al hipersuperfície de les constriccions, segons el procediment del mètode. Comproveu que la direcció d'exploració obtinguda és de descens.

74. Gradient reduït generalitzat.

Considereu el següent problema d'optimització no lineal:

$$
\textbf{(PNL)}\left\{\begin{array}{rl}
\min\limits_{x \in \mathbb{R}^3} & f(x) = \dfrac{x_1 x_2}{x_3 + 1} \\
\text{subj. a :} & x_1^2 + x_2^2 - x_3 = 1 \\
& 3x_1 + \dfrac{3}{2x_3} \leq 5
\end{array}\right.
$$

a) Enunicieu la condició d'aturada de l'algorisme del Gradient Reduït Generalitzat. Comproveu aquesta condició pel problema **(PNL)** sobre la solució factible $x^k = [\,1 \quad 1 \quad 1\,]'$

b) Si s'aplica una passa del mètode del GRG a partir del punt factible $x^k = [\,0 \quad 2 \quad 3\,]'$ amb longitud de pas $\alpha^k = 2$ s'obté el punt iterat:

$$x^{k+1} \approx [\,-0.11419 \quad 1.98611 \quad 2.94444\,]'$$

Efectueu una iteració del procés de retorn a la regió factible a partir de x^{k+1}. El vector de variables dependents a x^k és $y^k = [\,{x_1}^k \quad {x_2}^k\,]$

75. Gradient reduït generalitzat amb dues variables.

Considereu el següent problema de programació no lineal:

$$
\textbf{(PNL)}\left\{\begin{array}{rl}
\min\limits_{x \in \Re^2} & f(x) = x_1^2 - x_1 x_2 + 3x_3 \\
\text{subj. a :} & x_1^2 - x_2 + x_3 = 3 \\
& x_1 \geq -1 \quad , \quad x_2 \leq 1 \quad , \quad x_3 \geq 0
\end{array}\right.
$$

Efectueu una iteració completa del métode del Gradient Reduit Generalitzat a partir del punt inicial $x = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$, fent exploració lineal exacta i amb una tolerància de proximitat a la regió factible $\epsilon = 10^{-6}$. A l'hora de seleccionar el conjunt de variables dependents les variables amb índex inferior tenen prioritat.

76. El mètode SUMT, funcions convexes, i el mètode de Newton *

Existeix un mètode de programació no lineal anomenat SUMT (Sequential Unconstrained Minimization Technique, popularitzat per Fiacco i McCormick) consistent en substituir un problema de programació no lineal amb restriccions per la resolució successiva de problemes sense restriccions. Així, donat el problema general;

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

construim el següent problema sense restriccions (anomenat funció de penalització):

$$P(x, \rho) = f(x) + \frac{1}{\sqrt{\rho}} \sum_{i=1}^{m} h_i^2(x) + \rho \sum_{j=m+1}^{p} \frac{1}{g_j(x)}$$

on $\rho \in \mathbb{R}$. La idea del mètode és inicialment considerar un valor de ρ gran i minimitzar la funció $P(x, \rho)$. Després es disminueix ρ i es torna a solucionar el nou problema, prenent com punt inicial el punt solució del problema anterior, i així fins a arribar a un valor de ρ molt proper a 0. Fixem-nos que en el nou problema sense restriccions, un punt que no verifiqui que $h_i(x) = 0$ penalitzarà molt el valor de $P(x, \rho)$ quan $\rho \to 0$. Igualment, el terme $\frac{1}{g_j(x)}$ de $P(x, \rho)$ evita que ens atansem a $g_j(x) = 0$, i així garantim que sempre tindrem punts que satisfaran que $g_j(x) \geq 0$. Això fa que, en el límit, les solucions dels successius problemes sense restriccions satisfacin les restriccions del problema original, i siguin, doncs, factibles. Aquests termes $h_i^2(x)$ i $\frac{1}{g_j(x)}$ en que es transformen les restriccions originals s'anomenen penalitzacions.

Tenint en compte l'exposat anteriorment, i donat el problema de programació no lineal

$$\begin{array}{ll} \min\limits_{x_1, x_2} & \frac{x_1^2}{10} - \ln(x_1 + x_2 + 2) \\ \text{subj. a} & \\ & x_1^2 + x_2^2 \leq 10 \\ & x_1 \geq 1 \\ & x_2 \geq 0 \end{array}$$

contesteu a les següents qüestions:

a) Comproveu si la funció objectiu és o no convexa?

b) Formuleu la funció de penalització corresponent a la primera iteració del mètode SUMT de Fiacco i McCormick, amb $\rho = 1$ i penalitzacions inversament proporcionals ($\frac{1}{g_j(x)}$).

c) Considereu ara una nova penalització (anomenada *logarísmica*) que s'aplica a restriccions del tipus $g(x) \geq 0$, l'expressió de la qual és $-\ln g(x)$. A l'igual que a l'apartat anterior, escriviu la nova funció de penalització del mètode SUMT usant la nova penalització logarísmica que acabem de definir per a totes les restriccions de desigualtat. Raoneu també per què creieu que aquesta funció $-\ln g(x)$ es comporta com una penalització, i compareu-la amb la penalització inversament proporcional $\frac{1}{g_j(x)}$.

d) Comproveu que la funció de penalització del mètode SUMT formulada a l'apartat c) és una funció convexa **dins de la regió factible**. Donada la seva convexitat, quin mètode diferencial de minimització sense restriccions creieu convenient usar, i per què? (NOTA: per comprovar la convexitat de la funció de penalització us pot ser útil tenir en compte que si tenim una funció f tal que $f = f_1 + f_2$, aleshores $\nabla^2 f = \nabla^2 f_1 + \nabla^2 f_2$, on ∇^2 denota la matriu hessiana, i que si H_1 i H_2 són matrius (semi)definides positives aleshores $H_1 + H_2$ és també (semi)definida positiva).

e) Realitzeu una iteració del mètode de Newton amb la funció de penalització del SUMT de l'apartat c), prenent com punt inicial $x_1 = 2$ i $x_2 = 1$. Comproveu que el nou punt millora el valor de funció objectiu.

3 Solucions dels problemes de PNL sense constriccions.

3.1 Convexitat.

Solució del problema 1.

a) Sí és convex. Es pot comprovar de diverses formes. Una d'elles és aplicant la definició de convexitat "C és convex si $\forall u, v \in C \quad \alpha u + (1-\alpha)v \in C, \quad 0 \leq \alpha \leq 1$":

$$A(\alpha u + (1-\alpha)v) = \alpha Au + (1-\alpha)Av = \alpha b + (1-\alpha)b = b$$

Per tant $\alpha u + (1-\alpha)v \in C$

b) Sí és convex. Es pot comprovar de forma anàloga al cas anterior.

c) Sí és convex. Només cal veure que els punts de $\mathbb{R}^2$ tals que $x_1^2 + x_2^2 \leq 1$ defineixen un cercle centrat al (0,0) de radi 1. I un cercle és un conjunt convex.

d) Sí és convex. Aquest conjunt està format per tots els punts de $\mathbb{R}^2$ que satisfan que $-1 \leq x_1 + x_2 \leq 1$. Aquests són els punts que es troben entre les rectes paral·leles $x_1 + x_2 = -1$ i $x_1 + x_2 = 1$, i clarament son un conjunt convex (tot i que no afitat).

e) No és conjunt convex. Podem trobar fàcilment un contraexemple. Suposem que C_1 és l'interval real (i convex) $[0,1]$, i que C_2 és l'interval $[2,3]$. Clarament, els punts entre 1 i 2 no pertanyen a la unió, però poden ser obtinguts de forma $\alpha 1 + (1-\alpha)2 \quad 0 \leq \alpha \leq 1$. Per tant la unió de C_1 i C_2 no és un conjunt convex.

f) Sí és conjunt convex. Es demostra fàcilment directament aplicant la definició de convexitat. Cal veure si $\alpha u + (1-\alpha)v \in C_1 \cap C_2, \quad 0 \leq \alpha \leq 1$ on $u, v \in C_1 \cap C_2$. Com que $u \in C_1$ i C_1 és convex, tenim que $\alpha u + (1-\alpha)v \in C_1$. Aplicant el mateix raonament podem concloure que $\alpha u + (1-\alpha)v \in C_2$. Per tant $\alpha u + (1-\alpha)v \in C_1 \cap C_2$.

Solució del problema 2.

a) Hem de veure si $\forall x, y$ i $0 \leq \alpha \leq 1$ es satisfà

$$(f_1 + f_2)(\alpha x + (1-\alpha)y) \leq \alpha(f_1 + f_2)(x) + (1-\alpha)(f_1 + f_2)(y)$$

Sabent que $(f_1+f_2)(x) = f_1(x) + f_2(x)$ i usant que f_1 i f_2 són convexes tenim:

$$\begin{aligned}(f_1+f_2)(\alpha x+(1-\alpha)y) &= f_1(\alpha x+(1-\alpha)y) + f_2(\alpha x+(1-\alpha)y) \leq \\ &\leq \alpha f_1(x) + (1-\alpha)f_1(y) + \alpha f_2(x) + (1-\alpha)f_2(y) = \\ &= \alpha(f_1+f_2)(x) + (1-\alpha)(f_1+f_2)(y)\end{aligned}$$

b) Hem de veure si $\forall x, y$ i $0 \leq \alpha \leq 1$ es satisfà

$$af(\alpha x+(1-\alpha)y) \leq \alpha a f(x) + (1-\alpha)af(y)$$

Com que f és convexa tenim directament que:

$$f(\alpha x+(1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$

i multiplicant a ambdós costats per a (com que $a \geq 0$ ho podem fer sense alterar el sentit de la desigualtat) tenim directament el resultat desitjat.

Solució del problema 3.

Si la restricció és $g(x) = c$, en general no es verifica que el conjunt de punts definit és convex. Per exemple, si la restricció és $x^2 = 9$ (on x^2 és una funció convexa), el conjunt de punts factibles són $\{+3, -3\}$, que clarament no és un conjunt convex.

Solució del problema 4.

Si considerem que H té el terme diagonal de la posició i negatiu ($H_{ii} < 0$) aleshores podem veure com per algun vector x no es satisfà la condició $x^T Hx \geq 0$. Concretament, si prenem com x la i-èssima columna de la matriu identitat (columna que denotarem per e_i, i que té totes les components a 0, excepte la de la posició i on té un 1), aleshores tenim directament que

$$e_i^T He_i = H_{ii} < 0$$

Per tant, H no és semidefinida positiva.

Solució del problema 5.

Per determinar la convexitat de $f(x,y)$ en funció dels paràmetres p i q caldrà estudiar els signes dels determinants dels menors principals de la hessiana $\nabla^2 f(x,y)$:

$$\nabla f(x,y) = \begin{pmatrix} 2q/x & p\cos(py) \end{pmatrix} \qquad \nabla^2 f(x,y) = \begin{pmatrix} -2q/x^2 & 0 \\ 0 & -p^2\sin(py) \end{pmatrix}$$

I ara, per garantir que $\nabla^2 f(x,y)$ sigui semidefinida positiva (amb el qual $f(x,y)$ serà convexa), s'ha de verificar que el determinants dels dos menors principals siguin no negatius:

$$\begin{aligned}\det(\Delta_1) &= -2q/x^2 \geq 0 \\ \det(\Delta_2) &= 2p^2q/x^2\sin(py) \geq 0\end{aligned}$$

Hem d'observar en primer lloc que quan $x = 0$ el valor del primer determinant tendeix a $\pm\infty$. Això és degut a que la funció $\ln(x)$ tendeix a $-\infty$ en aquest punt. De fet hauríem d'imposar que $x > 0$ per tenir ben definit el signe del $\det(\Delta_1)$. Un cop comentat això, per garantir que

$\det(\Delta_1) \geq 0$ cal que $q \leq 0$.

Estudiem ara el segon determinant. El terme $2p^2q/x^2$ sempre és no positiu (donat que abans hem fixat que $q \leq 0$). Per tant si volem que $\det(\Delta_2) \geq 0$ cal que $\sin(py) \leq 0$. Això es verificarà sempre que $py \in \bigcup_{k=0}^{\infty} \left([(2k+1)\pi, (2k+2)\pi] \cup [-2k\pi, , -(2k+1)\pi]\right)$. En aquest cas veiem que p pot prendre qualsevol valor, i, segons el valor que prengui, la regió on $f(x,y)$ és convexa variarà.

Resumint, la funció $f(x,y)$ serà convexa per valors de $q \leq 0$, per a qualsevol valor de p, i dins el subconjunt de $\mathbb{R}^2$ següent:

$$\{(x,y) \mid x > 0,\ y \in \bigcup_{k=0}^{\infty} \left([(2k+1)/p\pi, (2k+2)/p\pi] \cup [-2k/p\pi, -(2k+1)/p\pi]\right)\}$$

La Fig. 2 mostra la gràfica de $f(x,y)$ per $q = -1$ i $p = 1$.

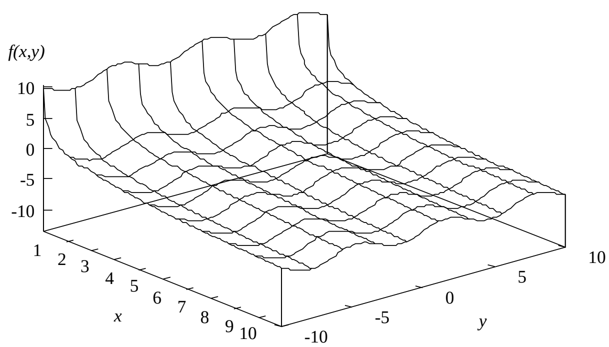


Figura 2. Gràfica de la funció $f(x,y) = q\ln(x^2) + \sin(py)$ quan $q = -1$ i $p = 1$

Solució del problema 6.

Per poder garantir que el punt solució obtingut és un mínim global s'ha de satisfer i) que el conjunt de punts Ω que defineix la regió factible sigui un conjunt convex, i ii) que la funció objectiu $f(x,y) = x^4 + 3y^2 - 2xy - 4x + 2y$ sigui convexa dins de Ω.

Anem a veure si es satisfà i). Tenim quatre restriccions. Sabem que el conjunt convex $\{z \in C \mid g_1(z) \leq c_1, \ldots, g_m(z) \leq c_m\}$ és convex si cada una de les restriccions $g_i(z)$ és convexa dins C (on C és convex). En aquest cas podem suposar que C correspon als valors de x i y majors que 1/2:

$$C = \{(x,y) \mid x \geq 1/2\ ,\ y \geq 1/2\}$$

Hem de veure llavors que ocorre amb les dues restriccions que ens queden. En primer lloc escrivim de forma $g_1(x) \leq c_1$ la primera:

$$-\ln(x+y) \leq 0$$

Ara hem de veure si $g_1(x,y)$ és una funció convexa dins C (observem que en aquest cas hem d'explicitar que és "dins C" donat que per valors tals que $x+y \leq 0$ la funcio ln no està definida; si això no passés ens podríem haver estalviat la introducció de C, i podríem haver considerat els límits simples $x \geq 1/2$ i $y \geq 1/2$ com dues restriccions qualsevols més). Per veure si és convexa cal veure si la matriu hessiana és semidefinida positiva:

$$\nabla g_1(x,y) = \begin{pmatrix} -1/(x+y) & -1/(x+y) \end{pmatrix} \qquad \nabla^2 g_1(x,y) = \begin{pmatrix} 1/(x+y)^2 & 1/(x+y)^2 \\ 1/(x+y)^2 & 1/(x+y)^2 \end{pmatrix}$$

Ara cal comprovar el signe del determinant dels menors principals de la matriu $\nabla^2 g_1(x,y)$:

$$\det(\Delta_1) = 1/(x+y)^2 \geq 0$$
$$\det(\Delta_2) = 1/(x+y)^4 - 1/(x+y)^4 = 0$$

Com que els dos determinants són ≥ 0 tenim que $-\ln(x+y) \leq 0$ defineix un conjunt convex.

Ara hem de fer el mateix estudi amb la segona restricció $x^2 + y^2 \leq 20$:

$$\nabla g_2(x,y) = \begin{pmatrix} 2x & 2y \end{pmatrix} \qquad \nabla^2 g_2(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\det(\Delta_1) = 2 \geq 0$$
$$\det(\Delta_2) = 4 \geq 0$$

Per tant $x^2 + y^2 \leq 20$ també defineix un conjunt convex dins C (en aquest cas defineix un conjunt convex a tot $\mathbb{R}^2$ de fet).

Ara finalment ens queda veure si es satisfà ii), és a dir, si la funció objectiu és convexa a $\Omega = \{(x,y) \mid \ln(x+y) \geq 0,\ x^2+y^2 \leq 20,\ x \geq 1/2,\ y \geq 1/2\}$, on Ω ara ja sabem que és un conjunt convex pel que hem vist anteriorment. Estudiem la funció objectiu f:

$$\nabla f(x,y) = \begin{pmatrix} 4x^3 - 2y - 4 & 6y - 2x + 2 \end{pmatrix} \qquad \nabla^2 f(x,y) = \begin{pmatrix} 12x^2 & -2 \\ -2 & 6 \end{pmatrix}$$

$$\det(\Delta_1) = 12x^2 \geq 0$$
$$\det(\Delta_2) = 6 \cdot 12x^2 - 4 = 72x^2 - 4$$

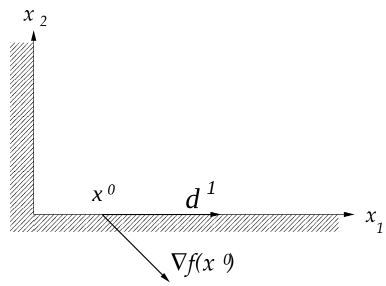
Observem com el determinant de $\nabla^2 f(x,y)$ és igual a $72x^2 - 4$. Per alguns valors (per. ex., $x = 0$) aquest determinant no pren valors positius. Tanmateix, per garantir que f és convexa dins Ω només hem d'assegurar-nos de que $72x^2 - 4$ sigui positiu dins Ω. I dins Ω els valors possibles de x sempre són superiors o iguals a $1/2$. En el cas extrem, per $x = 1/2$, tenim que $72x^2 - 4 = 72/4 - 4 = 18 - 4 \geq 0$. Per tant podem concloure que f és convexa dins Ω.

A la vista dels resultats anteriors podem garantir que l'òptim proporcionat per l'algorisme d'optimització correspon a un mínim global.

3.2 Optimalitat i direccions de descens.

Solució del problema 8.

- $d^1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. La representació gràfica d'aquest primer cas és:



∗ *Factibilitat:* Els punts a partir de x^0 al llarg de d^1 es representen per:

$$x(\alpha) = x^0 + \alpha d^1 = \begin{bmatrix} 1+2\alpha \\ 0 \end{bmatrix}$$

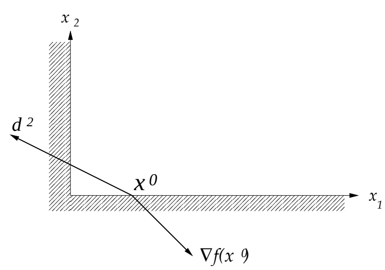
donat que $x(\alpha) \geq 0$ per a tot $\alpha \geq 0$ la direcció és factible.

∗ *Descens:* Fem el producte escalar amb el vector gradient:

$$\nabla f(x^0)d^1 = [\,2 \quad -2\,] \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 4 > 0$$

donat que $\nabla f(x^0)d^1 > 0$, la direcció no és de descens.

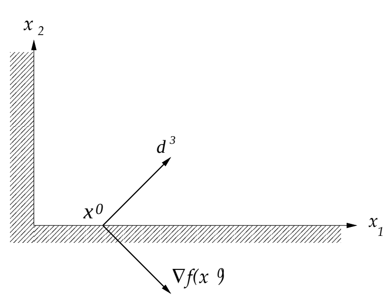
- $d^2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.



∗ *Factibilitat:* $x(\alpha) = x^0 + \alpha d^2 = \begin{bmatrix} 1-2\alpha \\ \alpha \end{bmatrix}$. Per a qualssevol valor de α entre zero i $\alpha = 1/2$ es satisfà $x(\alpha) \geq 0$, sent doncs d^1 factible.

∗ Descens: $\nabla f(x^0)d^2 = [\,2 \quad -2\,] \begin{bmatrix} -2 \\ 1 \end{bmatrix} = -6 < 0$. d^2 és de descens.

- $d^3 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$



$*$ Factibilitat: $x(\alpha) = \begin{bmatrix} 1+2\alpha \\ 2\alpha \end{bmatrix} \geq 0 \forall \alpha > 0$, sent doncs factible.

$*$ Descens: $\nabla f(x^0)d^2 = [\,2 \quad -2\,] \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 0$. En aquest cas, per tal de saber si d^3 és de descens hem d'estudiar les segones derivades de $f(x)$ al llarg de d^3:

$$\nabla^2 f(x^0) = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

La corbatura de $f(x)$ al llarg de d^3, $g''(\alpha)$ ens pot indicar, si la direcció és de descens ($g''(\alpha) < 0$) o d'ascens ($g''(\alpha) > 0$). En el nostre cas és:

$$g''(\alpha) = d^{3'} \nabla f(x^0) d^3 = [\,2 \quad 2\,] \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 0$$

Així doncs, l'estudi de la corbatura tampoc ens permet dir si la direcció és de descens.

Solució del problema 9.

VECTOR GRADIENT I MATRIU HESSIANA.

$$\nabla f(X)' = \begin{bmatrix} 2x_1 + x_2 \\ 4x_2 + x_1 - x_3 - 5 \\ 6x_3 - x_2 - 10 \end{bmatrix} \quad ; \quad \nabla^2 f(X) = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 6 \end{bmatrix}$$

CONDICIÓ DE $f(X)$ ESTRICTAMENT CONVEXA.

$\nabla^2 f(X)$ és definit positiu perquè els menors principals són tots positius:

$$2 > 0 \quad , \quad \det \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} = 7 > 0 \text{ , i } \det \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 6 \end{bmatrix} = 40 > 0$$

Així doncs, $\nabla^2 f(X)$ és definit positiu per a $\forall X$ (perquè $\nabla^2 f(X)$ és constant), i això és una condició suficient perquè $f(X)$ sigui estrictament convexa.

MÍNIM DE $f(X)$ A $\mathbb{R}^n$

$f(X)$ és una funció quadràtica amb $Q = \nabla^2 f(X)$ definida positiva. Primer comprovarem si el mínim d'aquesta funció quadràtica està dins del domini definit per $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$.

Per a una funció quadràtica $\frac{1}{2}X'QX - B'X$ el gradient és $G(X) = QX - B$. Per a $X = \underline{0}$ es té:

$$G(\underline{0}) = -B \quad , \quad G(\underline{0}) = \nabla f(\underline{0})' = \begin{bmatrix} 0 \\ -5 \\ -10 \end{bmatrix} \Rightarrow B = \begin{bmatrix} 0 \\ 5 \\ 10 \end{bmatrix}$$

d'on $f(X) = \frac{1}{2}X' \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 6 \end{bmatrix} X - [\,0 \quad 5 \quad 10\,]\, X$. X^* és tal que $QX^* = B$. Resolent el sis-

tema $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 6 \end{bmatrix} X = \begin{bmatrix} 0 \\ 5 \\ 10 \end{bmatrix}$ amb $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 6 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 1/2 & 1 & \\ 0 & -2/7 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ & 7/2 & -1 \\ & & 40/7 \end{bmatrix}$
s'obté:

$$\begin{bmatrix} 1 & & \\ 1/2 & 1 & \\ 0 & -2/7 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ & 7/2 & -1 \\ & & 40/7 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 1 & & \\ 1/2 & 1 & \\ 0 & -2/7 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 10 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 80/7 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 0 \\ & 7/2 & -1 \\ & & 40/7 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 80/7 \end{bmatrix} \Rightarrow \boxed{\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}}$$

que es troba fora de $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$.

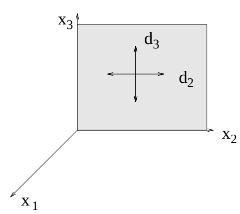
MÍNIM DE f(X) AL DOMINI DEFINIT PER $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$.

El mínim al domini $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$ haurà de satisfer $\nabla(X^*)D \geq 0\, \forall D$ factible, i estarà sobre la frontera. S'harà doncs de satisfer:

$$\begin{bmatrix} 2x_1 + x_2 & 4x_2 + x_1 - x_3 - 5 & 6x_3 - x_2 - 10 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \geq 0$$

Segons a quina faceta ens trobem de la frontera definida per $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$, es poden donar els casos següents:

a) $\underline{x_1 = 0,\ d_2 \text{ i } d_3 \text{ qualsevol i } d_1 \geq 0}$:



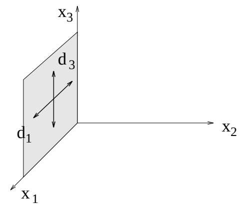
$$\begin{bmatrix} x_2 & 4x_2 - x_3 - 5 & 6x_3 - x_2 - 10 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \geq 0 \Rightarrow \begin{cases} x_2 \geq 0 & \text{(perquè } d_1 \geq 0) \\ 4x_2 - x_3 - 5 = 0 & \text{(perquè } d_2 \text{ pot ser molt negatiu)} \\ 6x_3 - x_2 - 10 = 0 & \text{(perquè } d_2 \text{ pot ser molt negatiu)} \end{cases}$$

d'on el sistema

$$\begin{cases} 4x_2 - x_3 \\ -x_2 + 6x_3 = 10 \end{cases} \Rightarrow \begin{matrix} x_2 = 40/23 \geq 0 \\ x_3 = 45/23 \geq 0 \end{matrix} \Rightarrow \boxed{\tilde{X}^* = \begin{bmatrix} 0 \\ 40/23 \\ 45/23 \end{bmatrix}}$$

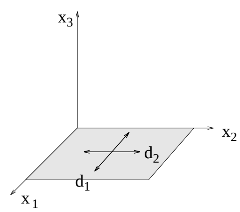
on $\tilde{X}^*$ és un mínim a $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$.

b) $\underline{x_2 = 0,\ d_1 \text{ i } d_3 \text{ qualsevol i } d_2 \geq 0}$:



$$[\,2x_1 \quad x_1 - x_3 - 5 \quad 6x_3 - 10\,] \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \geq 0 \Rightarrow \begin{cases} x_2 = 0 & d_1 \text{ i } d_3 \text{ qualssevol, i } d_2 \geq 0 \\ x_1 = 0 & \text{(perquè } d_1 \text{ pot ser molt negatiu)} \\ -x_3 - 5 \geq 0 & \text{(perquè } d_2 \geq 0) \Rightarrow \boxed{x_3 \leq -5 \text{ IMPOSSIBLE}} \end{cases}$$

c) $\underline{x_3 = 0,\ d_1 \text{ i } d_2 \text{ qualsevol i } d_3 \geq 0}$:



$$[\,2x_1 + x_2 \quad 4x_2 + x_1 - 5 \quad -x_2 - 10\,] \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \geq 0 \Rightarrow \begin{cases} 2x_1 + x_2 = 0 \text{ (perquè } d_1 \text{ pot ser molt negatiu). Llavors, si } x_1 \geq 0 \Rightarrow x_2 \leq 0 \text{ i viceversa} \Rightarrow \boxed{\text{IMPOSSIBLE}} \end{cases}$$

L'únic mínim és doncs $\boxed{\tilde{X}^* = [\,0 \quad 40/23 \quad 45/23\,]'}$.

3.3 Condició d'òptim per a problemes sense constriccions.

Solució del problema 11.

a) Calcularem primer els punts estacionaris:

$$\nabla f(x) = \begin{bmatrix} 2x_1 - x_2 - x_3 \\ -x_1 + 2x_2 + x_3 \\ -x_1 + x_2 + 2x_3 \end{bmatrix} \quad , \quad \nabla f(x^*) = [0] \quad , \quad \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\det \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} \neq 0 \Rightarrow x^* = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ únic punt estacionari.}$$

Estudiem ara la convexitat de $f(x)$ sobre el punt trobat:

$$\nabla^2 f(x) = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

la matriu Hessiana és constant: es tracta d'una funció quadràtica. Comprovem la definició de l'Hessiana:

$$\Delta_1 = 2 > 0 \quad , \quad \Delta_2 = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 3 > 0 \quad , \quad \Delta_3 = \det \nabla^2 f(x) = 4 > 0$$

$$\Delta_k > 0 \ , \ k = 1, 2, 3 \Rightarrow \nabla^2 f(x) \text{ def} + \forall x \in \mathbb{R}^3$$

l'Hessiana és definida positiva arreu ($f(x)$ estrictament convexa): el punt $x^* = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ és un mínim global estricte de $f(x)$.

b) Punts estacionaris de f(x):

$$\nabla f(x) = \begin{bmatrix} e^{x_1 - x_2} - e^{x_2 - x_1} + 2x_1 e^{x_1^2} \\ -e^{x_1 - x_2} + e^{x_2 - x_1} \\ 2x_3 \end{bmatrix}$$

$$\nabla f(x^*) = [0] \Rightarrow \begin{cases} x_3^* = 0 \\ e^{x_2^* - x_1^*} = e^{x_1^* - x_2^*} \Rightarrow x_1^* - x_2^* = x_2^* - x_1^* \Rightarrow \boxed{x_1^* = x_2^*} \\ \left.\begin{matrix} e^{x_1^* - x_2^*} - e^{x_2^* - x_1^*} + 2x_1^* e^{x_1^{*2}} = 2x_1^* e^{x^{*2}_1} = 0 \\ e^{x_1^{*2}} > 0 \end{matrix}\right\} \boxed{x_1^* = 0} \Rightarrow \boxed{x_2^* = 0} \end{cases}$$

Comprovem la definició de la Hessiana sobre x^*:

$$\nabla^2 f(x) = \begin{bmatrix} e^{x_1 - x_2} + e^{x_2 - x_1} + 4x_1^2 e^{x_1^2} + 2e^{x_1^2} & -e^{x_1 - x_2} - e^{x_2 - x_1} & 0 \\ -e^{x_1 - x_2} - e^{x_2 - x_1} & e^{x_1 - x_2} + e^{x_2 - x_1} & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$\Delta_1 > 0$ doncs és suma de termes positius

$$\Delta_2 = (e^{x_1 - x_2} + e^{x_2 - x_1})^2 + (e^{x_1 - x_2} + e^{x_2 - x_1})(e^{x_1 - x_2} + e^{x_2 - x_1} + 4x_1^2 e^{x_1^2} + 2e^{x_1^2}) - (e^{x_1 - x_2} + e^{x_2 - x_1})^2 =$$
$$= (e^{x_1 - x_2} + e^{x_2 - x_1})(e^{x_1 - x_2} + e^{x_2 - x_1} + 4x_1^2 e^{x_1^2} + 2e^{x_1^2}) > 0$$

$\Delta_3 = 2\Delta_2 > 0$

$$\Delta_k > 0 \ , \ k = 1, 2, 3 \Rightarrow \nabla^2 f(x) \text{ def} + \forall x \in \mathbb{R}^3$$

b) l'Hessiana és definida positia arreu ($f(x)$ estrictament convexa): el punt $x^* = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ és un mínim global estricte de $f(x)$.

c) Es calculen els punts estacionaris de $f(x)$:

$$\nabla f(x) = \begin{bmatrix} e^{x_1 - x_2} - e^{x_2 - x_1} \\ -e^{x_1 - x_2} + e^{x_2 - x_1} \end{bmatrix}$$

$$\nabla f(x^*) = [0] \Rightarrow e^{x_1^*-x_2^*} = e^{x_2^*-x_1^*} \ , \ x_1^* - x_2^* = x_2^* - x_1^* \ , \ 2x_1^* = 2x_2^* \ , \ \boxed{x_1^* = x_2^*}$$

En aquest cas ens trobem amb un conjunt de punts estacionaris $\mathcal{X}^*$ determinat pels punts de $\mathbb{R}^2$ amb $x_1 = x_2$. Comprovem ara la definició de la matriu Hessiana:

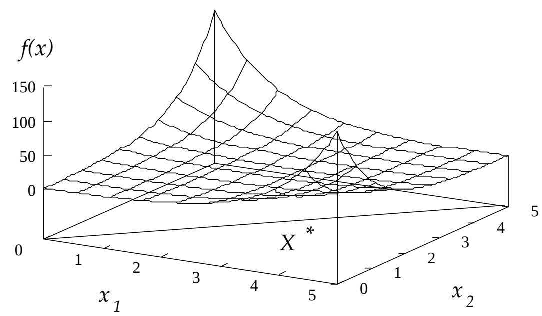
$$\nabla^2 f(x) = \begin{bmatrix} e^{x_1-x_2} + e^{x_2-x_1} & -e^{x_1-x_2} - e^{x_2-x_1} \\ -e^{x_1-x_2} - e^{x_2-x_1} & e^{x_1-x_2} + e^{x_2-x_1} \end{bmatrix}$$

$\Delta_1 = e^{x_1-x_2} + e^{x_2-x_1} > \forall x \in \mathbb{R}^2$

$\Delta_2 = (e^{x_1-x_2} + e^{x_2-x_1})^2 - (e^{x_1-x_2} + e^{x_2-x_1})^2 = 0$

$$\Delta_1 > 0 \ , \ \Delta_2 = 0 \Rightarrow \nabla^2 f(x) \text{ semidef} + \forall x \in \mathbb{R}^2$$

La matriu Hessiana és semidefinida positiva arreu ($f(x)$ és convexa): el conjunt de punt $\mathcal{X}^*$ són mínims global febles. La següent figura mostra la respresentació gràfica de la funció $f(x)$ i del conjunt $\mathcal{X}^*$:



d) Càlcul de punts estacionaris:

$$\nabla f(x) = \begin{bmatrix} e^{x_1-x_2} + e^{x_1+x_2} \\ -e^{x_1-x_2} + e^{x_1-x_2} \end{bmatrix}$$

$$\nabla f(x^*) = [0] \Rightarrow \begin{cases} e^{x_1+x_2} = e^{x_1-x_2} \ , \ x_1 + x_2 = x_1 - x_2 \ , \ x_2 = -x_2 \Rightarrow \boxed{x_2 = 0} \\ \left. \begin{array}{r} e^{x_1-x_2} = -e^{x_1+x_2} \\ x_2 = 0 \end{array} \right\} \Rightarrow e^{x_1} = -e^{x_1} \Rightarrow \boxed{\nexists x^*} \end{cases}$$

En aquest cas no existeixen punts estacionaris, i el conjunt solució $\mathcal{X}^*$ és buit (el problema no té solució). La funció $f(x)$ no està afitada inferiorment: per a un valor donat de x_2, la funció objectiu decreix a mida que ho fa x_1.

Solució del problema 12.

Per trobar els punts estacionaris cal solucionar $\nabla f(x) = 0$ per a cada problema, i tot seguit estudiar si $\nabla^2 f(x)$ és definida positiva als punts trobats.

a) La funció objectiu és $f(x, y) = x^2 + y^2 - 4x - 8y + xy$. El gradient i hessiana de f són:

$$\nabla f = (\, 2x - 4 + y \quad 2y - 8 + x \,) \qquad \nabla^2 f = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Observem com l'únic punt estacionari que hi ha és $(x, y) = (0, 4)$. Donat que la matriu hessiana és definida positiva $\forall (x, y)$, podem concloure que el nostre punt és un mínim (i a més global, donat que la funció original és convexa a tot $\mathbb{R}^2$).

b) La funció objectiu és $f(x,y) = x^2 - y^2 - 4x + 6y - 5$. El gradient i hessiana de f són:

$$\nabla f = (\, 2x - 4 \quad -2y + 6 \,) \qquad\qquad \nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

L'únic punt estacionari que hi ha és $(x,y) = (2,3)$. La matriu hessiana, però, no és definida positiva. Per tant no podem garantir que sigui un mínim. De fet, no ho és. Un punt estacionari on la funció té hessiana indefinida (és a dir, $\exists x, y : x^T H x > 0 \quad y^T H y < 0$) s'anomena punt de sella, i correspon a punts on tant podem avançar en direccions que ens disminueixen com incrementen el valor de la funció objectiu. La Fig. 3 ens mostra l'aspecte de la funció $f(x,y)$ on queda clar el concepte de punt de sella (rep aquest nom perquè recorda a una "sella" de muntar).

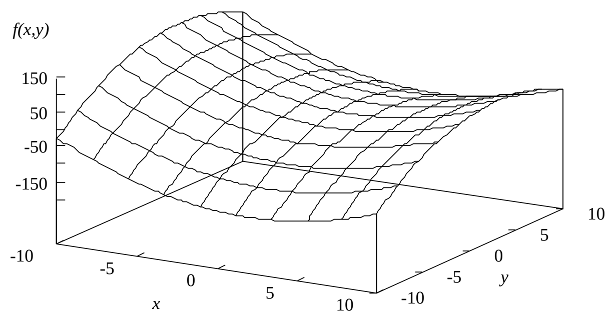


Figura 3. Gràfica de la funció $f(x,y) = x^2 - y^2 - 4x + 6y - 5$

c) La funció objectiu és $f(x,y) = -e^{-(x^4+y^4)}$. El gradient i hessiana de f són:

$$\nabla f = (\, 4x^3 e^{-(x^4+y^4)} \quad 4y^3 e^{-(x^4+y^4)} \,)$$

$$\nabla^2 f = \begin{pmatrix} -16x^6 e^{-(x^4+y^4)} + 12x^2 e^{-(x^4+y^4)} & -16x^3y^3 e^{-(x^4+y^4)} \\ -16x^3y^3 e^{-(x^4+y^4)} & -16y^6 e^{-(x^4+y^4)} + 12y^2 e^{-(x^4+y^4)} \end{pmatrix}$$

L'únic punt estacionari que hi ha és $(x,y) = (0,0)$ (ja que el terme $e^{-(x^4+y^4)}$ sempre és positiu, i únicament podem anul·lar les components del gradient fent que $x = y = 0$). La matriu hessiana avaluada al punt $(0,0)$ és una matriu de zeros, i, per tant, és semidefinida positiva. Donat que la condició suficient per garantir que $(0,0)$ és un mínim és que $\nabla^2 f(0,0)$ sigui definida positiva, no podem garantir que aquest punt sigui un mínim local. En aquest cas, però, sí que és mínim, i a més, és mínim global. Això es pot comprovar observant la gràfica d'aquesta funció a la Fig. 4, la qual té forma de campana invertida.

d) La funció objectiu és $f(x,y) = x^2 - y^4$. El gradient i hessiana de f són:

$$\nabla f = (\, 2x \quad -4y^3 \,) \qquad\qquad \nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -12y^2 \end{pmatrix}$$

L'únic punt estacionari que hi ha és $(x,y) = (0,0)$. La matriu hessiana avaluada al punt $(0,0)$ és semidefinida positiva. Com al cas c) anterior, no podem garantir que $(0,0)$ sigui un mínim

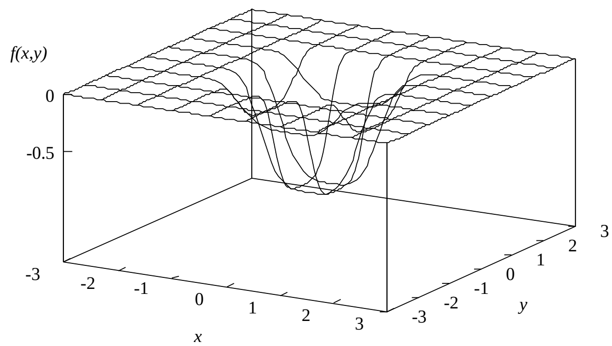


Figura 4. Gràfica de la funció $f(x,y) = -e^{-(x^4+y^4)}$

local (ja que $\nabla^2 f(0,0)$ no és definida positiva). En aquest cas, però, i a diferència de c), el punt $(0,0)$ no és cap mínim local (per exemple, el punt $(0,a)$ $a \in \mathbb{R}$ té un valor de funció objectiu menor).

Solució del problema 13.

Aquest és un problema on es pot aplicar la tècnica general dels mínims quadrats. Denotant per M_A la massa del flux del producte A, l'error que fem en un dels n instants de temps és $e_i = M_A + M_{B_i} - M_{C_i}$. Com que volem minimitzar la suma dels errors per als n intervals, només hem de sumar els valors de e_i, però elevant-los al quadrat (si no, podrien produir-se cancelacions per canvis de signe, les quals són eliminades en elevar al quadrat). El problema que haurem de solucionar finalment serà:

$$\min_{M_A} \sum_{i=1}^{n} (M_A + M_{B_i} - M_{C_i})^2$$

3.4 Mètodes no derivatius: l'algorisme de Nelder i Mead

Solució del problema 15.

EXPANSIÓ

Quan $f(x_r)$ té valor inferior al millor vèrtex $(f(x_a))$ i a més si determinant un vèrtex expandit x_e al llarg de la direcció de x_c a x_R, el vèrtex expandit és encara millor. D'on hi haurà Expansió si:

$$f(x_R) \leq 3 = f(x_a) \quad \text{i} \quad f(x_E) < f(x_a)$$

CONTRACCIÓ

Quan $f(x_R)$ té valor pitjor que tots els actuals menys el pitjor:

$$f(x_b) = 4 < f(x_R) \leq f(x_c) = 5$$

REDUCCIÓ

Quan $f(x_R)$ té valor pitjor que tots els actuals

$$f(x_R) > 5 = f(x_c)$$

SUBSTITUCIÓ DE x_c PER x_R

Quan $f(x_R)$ té valor entre el millor i el pitjor actuals i no és només millor que el pitjor actual (en el qual cas fóra contracció):

$$f(x_a) = 3 < f(x_R) \leq 4 = f(x_b)$$

També en el cas d'expansió infructuosa:

$$f(x_R) < 3 = f(x_a) \quad \text{i} \quad f(x_E) \geq f(x_R)$$

3.5 Exploració lineal.

Solució del problema 17.

La funció $f(x) = x^2 - 3x + 5 + \ln(\frac{1}{x+1})$ a l'interval [0,5] té una gràfica com la de la Fig. 5. Clarament s'observa com té un únic mínim en aquest interval. De fet la primera i segona derivades de $f(x)$ són:

$$f'(x) = 2x - 3 - \frac{1}{x+1} \qquad\qquad f''(x) = 2 + \frac{1}{(x+1)^2}$$

La segona derivada és sempre positiva, i per tant la funció $f(x)$ és convexa a [0,5] i només té un únic mínim en aquest interval (és a dir, és unimodal). Anul·lant la primera derivada podem observar com aquest mínim correspon al valor $x^* = (1+\sqrt{33})/4 = 1.6861$.

Un cop hem garantit que la funció $f(x)$ és unimodal a [0,5] podem aplicar la cerca de Fibonacci. En primer lloc cal determinar el nombre N d'avaluacions que ens cal fer per donar un interval d'incertesa de longitud menor o igual a 1, tal i com demana l'enunciat. L'interval original [0,5] té una longitud de $d_1 = 5$. Sabem que la longitud d_k de l'interval a cada avaluació k, al mètode de Fibonacci ve donada per

$$d_k = \frac{F_{N-k+1}}{F_N} d_1$$

A la darrera avaluació tenim que $k = N$ (per tant $F_{N-k+1} = F_1 = 1$), i que $d_k \leq 1$. Tot això fa que:

$$1 \geq d_k = \frac{F_1}{F_N} 5 = \frac{5}{F_N} \quad \Rightarrow \quad F_N \geq 5 \quad \Rightarrow \quad N \geq 4$$

Per tant prendrem $N = 4$, per tal de realitzar un nombre mínim d'avaluacions. Ara ja podem aplicar la cerca de Fibonacci. Ens serà útil conèixer el valor de la funció als extrems de l'interval: $f(0) = 5$, $f(5) = 13.208$.

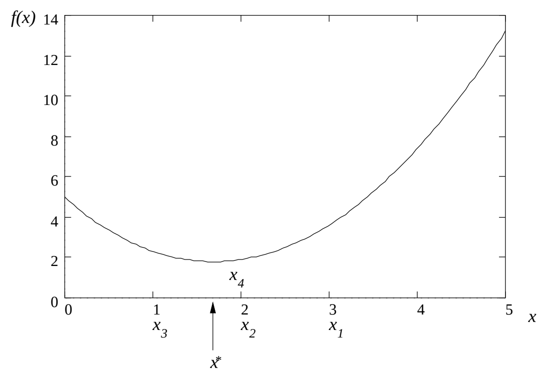


Figura 5. Gràfica de la funció $f(x) = x^2 - 3x + 5 + \ln(\frac{1}{x+1})$ amb la seqüència de punts obtinguts amb la cerca lineal de Fibonacci.

$k = 2)$ Trobem la longitud del segon interval:

$$d_2 = \frac{F_{N-2+1}}{F_N} d_1 = \frac{3}{5} 5 = 3$$

Els dos primers punts a avaluar són:

$$x_1 = 0 + d_2 = 3 \qquad f(3) = 3.6137$$
$$x_2 = 5 - d_2 = 2 \qquad f(2) = 1.9014$$

Com que $f(0) > f(2)$ i $f(2) < f(3)$ escollim l'interval $[0, x_1 = 3]$ (en comptes de $[x_2 = 2, 5]$). El nou interval és de longitud $d_2 = 3$.

$k = 3)$ Trobem la longitud del tercer interval:

$$d_3 = \frac{F_{N-3+1}}{F_N} d_1 = \frac{2}{5} 5 = 2$$

El nou punt a avaluar és:

$$x_3 = x_1 - d_3 = 3 - 2 = 1 \qquad f(1) = 2.3069$$

Adonem-nos de que l'altre punt simètric $0 + d_3 = 2$ no cal tornar-lo a avaluar perquè coincideix amb el punt x_2 que ja havíem trobat anteriorment. Com que $f(1) > f(2)$ i $f(2) < f(3)$ escollim l'interval $[x_3 = 1, x_1 = 3]$ (en comptes de $[0, x_2 = 2]$). El nou interval és de longitud $d_3 = 2$.

$k = 4)$ Trobem la longitud del quart i darrer interval:

$$d_4 = \frac{F_{N-4+1}}{F_N} d_1 = \frac{1}{5} 5 = 1$$

En aquest darrer interval sempre succeeix (quan s'usa el mètode de Fibonacci) que els dos punts situats a una distància $d_4 = 1$ dels extrems [1,3], són de fet el mateix punt (que a més ja hem avaluat: $x_2 = 2$). Per tant cal introduir una petita perturbació i avaluar un punt proper a l'anterior (per exemple, perturbem amb $\epsilon = 0.01$):

$$x_4 = x_1 - d_4 - \epsilon = 3 - 1 - 0.01 = 1.99 \qquad\qquad f(1.99) = 1.8948$$

Com que $f(1) > f(1.99)$ i $f(1.99) < f(2)$ l'interval que obtindrem finalment és $[x_3 = 1, x_2 = 2.0]$ (en comptes de $[x_4 = 1.99, x_1 = 3]$). El nou interval ara és de longitud $d_4 = 1$, amb el qual hem arribat a donar un interval d'incertesa de la longitud desitjada.

La Fig. 5 mostra els successius punts que s'han anat trobant en aplicar la cerca lineal de Fibonacci.

Solució del problema 18.

Al mètode de la secció àuria la longitud de l'interval d'incertesa número k és de

$$d_k = \tau^{k-1} d_1, \quad \text{on } d_1 = b - a \text{ i } \tau = \tfrac{2}{1+\sqrt{5}} \approx 0.618$$

essent $[a, b]$ l'interval d'incertesa original. En aquest problema concret $a = 0$ i $b = 5$. Com que volem un interval final de longitud menor o igual a 1, el valor k s'ha de determinar de forma:

$$1 \geq d_k = \tau^{k-1}(b-a) = \tau^{k-1}5 \quad \Rightarrow \quad \tau^{k-1}5 \leq 1/5 = 0.2 \quad \Rightarrow \quad k \geq 5$$

Per tant haurem de determinar 4 nous intervals (el valor de $k = 1$ correspon a l'interval original [0,5]), el que representarà 5 avaluacions de $f(x)$ a 5 punts diferents. Adonem-nos, però, de que no ens era necessari saber a priori quantes avaluacions hem de calcular per tal d'obtenir l'interval d'incertesa de mida necessària, i únicament s'ha fet per il·lustrar millor l'exercici. No ens era necessari perquè sabem a priori quant hem de reduir l'interval d'incertesa a cada iteració i quins nous punts hem d'anar calculant. Aquesta és una diferència important entre el mètode de la secció àuria i el de Fibonacci (en el de Fibonacci era imprescindible determinar a priori el nombre d'avaluacions N).

Sabent que $f(0) = 5$ i $f(5) = 13.208$, els intervals que anem generant són:

$k = 2$) Trobem la longitud del segon interval:

$$d_2 = \tau^{2-1} d_1 = \tau \cdot 5 = 3.0902$$

Els dos primers punts a avaluar són:

$$\begin{aligned} x_1 &= 0 + d2 = 3.0902 & \qquad f(3.0902) &= 3.8701 \\ x_2 &= 5 - d_2 = 1.9098 & \qquad f(1.9098) &= 1.8499 \end{aligned}$$

Com que $f(0) > f(x_2)$ i $f(x_2) < f(x_1)$ (escollim l'interval $[0, x_1 = 3.0902]$ (en comptes de $[x_2 = 1.9098, 5]$). El nou interval és de longitud d_2.

$k = 3$) Trobem la longitud del tercer interval:

$$d_3 = \tau^{3-1} d_1 = \tau^2 \cdot 5 = 1.9098$$

El nou punt a avaluar és:

$$x_3 = x_1 - d_3 = 1.1803 \qquad\qquad f(1.1803) = 2.0727$$

Adonem-nos de que l'altre punt simètric $0 + d_3 = 1.9098$ no cal tornar-lo a avaluar perquè coincideix amb el punt x_2 que ja havíem trobat anteriorment. Com que $f(x_3) >$

$f(x_2)$ i $f(x_2) < f(x_1)$ escollim l'interval $[x_3 = 1.1803, x_1 = 3.0902]$ (en comptes de $[0, x_2 = 1.9098]$). El nou interval és de longitud d_3.

$k = 4)$ Trobem la longitud del quart interval:

$$d_4 = \tau^{4-1} d_1 = \tau^3 \cdot 5 = 1.1803$$

El nou punt a avaluar és:

$$x_4 = x_3 + d_4 = 2.3607 \qquad f(2.3607) = 2.2786$$

Observem com l'altre punt simètric $x_1 - d_4 = 1.9098$ no cal tornar-lo a avaluar perquè coincideix amb el punt x_2 que ja havíem trobat anteriorment. Com que $f(x_3) > f(x_2)$ i $f(x_2) < f(x_4)$ escollim l'interval $[x_3 = 1.1803, x_4 = 2.3607]$ (en comptes de $[x_2 = 1.9098, x_1 = 3.0902]$). El nou interval és de longitud d_4.

$k = 5)$ Trobem la longitud del cinquè i darrer interval:

$$d_5 = \tau^{5-1} d_1 = \tau^4 \cdot 5 = 0.72949$$

El nou punt a avaluar és:

$$x_5 = x_4 - d_5 = 1.6312 \qquad f(1.6312) = 1.7998$$

L'altre punt simètric $x_3 + d_5 = 1.9098$ no cal tornar-lo a avaluar perquè coincideix amb el punt x_2 que ja havíem trobat anteriorment. Com que $f(x_3) > f(x_5)$ i $f(x_5) < f(x_2)$ escollim l'interval $[x_3 = 1.1803, x_2 = 1.9098]$ (en comptes de $[x_5 = 1.6312, x_4 = 2.3607]$). Aquest darrer interval és de longitud $d_5 < 1$ tal i com es demanava a l'enunciat.

Solució del problema 19.

Tal i com s'observa a la solució del problema 17, la primera i segona derivada de $f(x) = x^2 - 3x + 5 + \ln(\frac{1}{x+1})$ són:

$$f'(x) = 2x - 3 - \frac{1}{x+1} \qquad f''(x) = 2 + \frac{1}{(x+1)^2}$$

i el punt mínim (calculat analíticament) correspon al valor $x^* = (1 + \sqrt{33})/4 = 1.68614066$.

Si usem el mètode de Newton, tenim que la seqüència de punts es genera segons:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_{k+1})} = x_k - \frac{2x_k - 3 - \frac{1}{x_k+1}}{2 + \frac{1}{(x_k+1)^2}}$$

Es pot comprovar com la seqüència de punts generada és de $x_0 = 4$, $x_1 = 1.64705882$, $x_2 = 1.68610279$, $x_3 = 1.68614066$. Observem com hem assolit el punt òptim amb una precisió de 8 xifres correctes en només 3 iteracions (comprovant la ràpida convergència del mètode de Newton —és quadràtica— prop del punt òptim).

Usant ara el mètode de la secant, la seqüència de punts és generada segons:

$$\begin{aligned} x_{k+1} &= x_k - f'(x_k)\ \frac{x_{k-1} - x_k}{f'(x_{k-1}) - f'(x_k)} \\ &= x_k - \big(2x_k - 3 - \frac{1}{x_k + 1}\big)\ \Big[\frac{x_{k-1} - x_k}{2(x_{k-1} - x_k) - \frac{1}{x_{k-1}+1} + \frac{1}{x_k+1}}\Big] \end{aligned}$$

Es pot comprovar com la seqüència de punts ara generada és de $x_0 = 4, \quad x_1 = 3 \quad x_2 = 1.65853659, \ \ x_3 = 1.68674699, \ \ x_4 = 1.68614107, \ \ x_5 = 1.68614066$ (a partir d'aquesta iteració obtenim el mateix punt). Observem com per arribar al mateix punt ens han calgut quatre iteracions (una més que amb el mètode de Newton), mostrant a la pràctica com el mètode de la secant té una convergència més lenta cap al punt òptim.

Solució del problema 20.

La Fig. 6 mostra la gràfica de la funció $f(x) = 2/3\ |x|^{3/2}$. La primera i segona derivada de $f(x)$ són:

$$f'(x) = \begin{cases} \sqrt{x} & \text{si } x \geq 0 \\ -\sqrt{-x} & \text{si } x \leq 0 \end{cases} \qquad f''(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{si } x \geq 0 \\ \frac{1}{2\sqrt{-x}} & \text{si } x \leq 0 \end{cases}$$

La seqüència generada pel mètode de Newton és, per tant:

$$x_{k+1} = \begin{cases} x_k - \big(\frac{1}{2\sqrt{x_k}}\big)^{-1}\sqrt{x_k} = -x_k & \text{si } x_k \geq 0 \\ x_k - \big(\frac{1}{2\sqrt{-x_k}}\big)^{-1}(-\sqrt{-x_k}) = -x_k & \text{si } x_k \leq 0 \end{cases} \quad \Longrightarrow \quad x_{k+1} = -x_k$$

Per tant, si comencem al punt $x_0 = 1$, anirem generant la seqüència $1, -1, 1, -1, ...$, que no convergeix a $x^* = 0$.

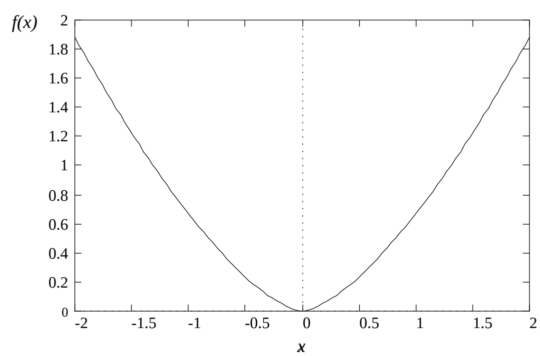


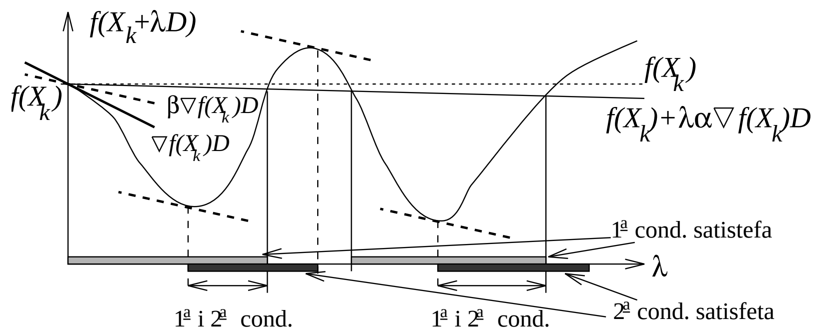
Figura 6. Gràfica de la funció $f(x) = 2/3\ |x|^{3/2}$ on es veu clarament que el mínim global correspon a $x^* = 0$.

Solució del problema 21.

1$^{\underline{\text{er}}}$ CRITERI D'ARMIJO-GOLDSTEIN

El 1$^{\underline{\text{er}}}$ criteri d'acceptabilitat d'una passa λ és que:

$$f(X_k+\lambda D) \leq f(X_k) + \lambda\alpha\nabla f(X_k)D \qquad 0<\alpha<1 \quad \text{generalment } \alpha{=}0{,}1$$



2$^{\underline{\text{on}}}$ CRITERI D'ARMIJO-GOLDSTEIN

Cal que:

$$\nabla f(X_k+\lambda D)D \geq \beta\nabla f(X_k)D \qquad \alpha<\beta<1 \quad \text{generalment } \beta{=}0{,}5$$

La figura mostra les zones on es satisfà el 1$^{\underline{\text{er}}}$ i el 2$^{\underline{\text{on}}}$ criteri d'Armijo-Goldstein d'acceptabilitat de la passa. En el dibuix s'han utilitzat (aproximadament) els valors α=0,1 i β=0,5

Solució del problema 22.

CÀLCULS PRELIMINARS :

$$f(x) = x_1^2 + x_1x_2 + \frac{2x_1}{x_2} \quad ; \quad \nabla f(x)' = \begin{bmatrix} 2x_1 + x_2 + \dfrac{2}{x_2} \\ x_1 - \dfrac{2x_1}{x_2^2} \end{bmatrix}$$

$$X_0 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad ; \quad f(X_0) = -2 \quad ; \quad \nabla f(X_0)' = \begin{bmatrix} -1 \\ \frac{1}{2} \end{bmatrix}$$

$$\nabla f(X_0)D = [\,-1 \quad \tfrac{1}{2}\,] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -\frac{3}{2} < 0 \Rightarrow \text{D de descens}$$

AJUST QUADRÀTIC :

Es pren el punt $X_0 + D$ com a punt addicional :

$$X_0 + D = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad ; \quad f(X_0+D) = 0$$

Ajust quadràtic : $q(\lambda) = a\lambda^2 + b\lambda + c$

$$\left.\begin{array}{lll} \lambda = 0 & : & q(0) = c = f(X_0) = -2 \Rightarrow \boxed{c = -2} \\ & & q(\lambda)' = 2a\lambda + b \quad ; \quad q(0)' = b = \nabla f(X_0)D = -\frac{3}{2} \Rightarrow \boxed{b = -\frac{3}{2}} \\ \lambda = 1 & : & q(1) = a + b + c = f(X_0 + D) = 0 \Rightarrow a + b + c = 0 \end{array}\right\} \Rightarrow \begin{array}{r} a - \frac{3}{2} - 2 = 0 \\ \boxed{a = \frac{7}{2}} \end{array}$$

Càlcul de λ_q : $q(\lambda_q)' = 2a\lambda_q + b = 0 \quad ; \quad \lambda_q = -\frac{b}{2a} = -\frac{-3/2}{7} \Rightarrow \boxed{\lambda_q = \frac{3}{14}}$ mínim, doncs $a > 0$.

COMPROVACIÓ DE LA 1^a COND. D'ARMIJO-GOLDSTEIN A $X_0 + \lambda_q D$:

Condició AG1 : $f(X_0 + \lambda D) \leq f(X_0) + \alpha\nabla f(X_0)D\lambda \quad ; \quad \alpha \in [0, 1]$. S'acostuma a prendre un escalar α molt petit. Usarem $\alpha = 0.05$.

$$\left.\begin{array}{l} X_0 + \lambda_q D = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \frac{3}{14}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{10}{7} \\ -\frac{25}{14} \end{bmatrix} \\ f(X_0 + \lambda_q D) = \underline{-2.1102} < \\ < f(X_0) + \alpha\nabla f(X_0)D\lambda_q = -2 + 0.05(-\frac{3}{2})\frac{3}{14} = \underline{-2.0161} \end{array}\right\} \Rightarrow \boxed{\text{Es satisfà AG1}}$$

AJUST CÚBIC :

Es prenen els punts $X_0 + D$ i $X_0 + \frac{5}{9}D$ com a punts addicionals :

$$X_0 + D = \begin{bmatrix} 3 \\ -1 \end{bmatrix}; f(X_0 + D) = 0; X_0 + \frac{5}{9}D = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \frac{5}{9}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{19}{9} \\ \frac{-13}{9} \end{bmatrix}; f(X_0 + \frac{5}{9}D) = -1.5156$$

Ajust cúbic : $u(\lambda) = a\lambda^3 + b\lambda^2 + c\lambda + d$

$$\left.\begin{array}{ll} \lambda = 0 & : u(0) = d = f(X_0) = -2 \Rightarrow \boxed{d = -2} \\ & u(\lambda)' = 3a\lambda^2 + 2b\lambda + c \\ & u(0)' = c = \nabla f(X_0)D = -\frac{3}{2} \Rightarrow \boxed{c = -\frac{3}{2}} \\ \lambda = 1 & : u(1) = a + b + c + d = f(X_0 + D) = 0 \Rightarrow \underline{a + b = \frac{7}{2}} \\ \lambda = \frac{5}{9} & : u(\frac{5}{9}) = a\frac{125}{729} + b\frac{25}{81} + c\frac{5}{9} - 2 = f(X_0 + \frac{5}{9}D) = -1.5156 \end{array}\right\} \left.\begin{array}{l} a + b = \frac{7}{2} \\ \frac{125}{729}a + \frac{25}{81}b = 1.3176 \end{array}\right\} (*)$$

$$(*) \Rightarrow \boxed{\begin{array}{l} a = -1.7307 \\ b = 5.2307 \end{array}}$$

Càlcul de λ_c :

$$u(\lambda_u)' = 3a\lambda_u^2 + 2b\lambda_u + c = 0 \quad ; \quad -5.1921\lambda_u^2 + 10.4614\lambda_u - 1.5 = 0$$

$$\lambda_u = \frac{-10.4614 + \sqrt{109.4408 - 31.1526}}{-10.3842} = 0.1553 \Rightarrow \boxed{\lambda_u = 0.1553}$$

COMPROVACIÓ DE LA 2^A COND.D'ARMIJO-GOLDSTEIN A $X_0+\lambda_u D$:

Condició AG2 : $\nabla f(X_0+\lambda D)D \geq \beta \nabla f(X_0)D \quad ; \quad \beta \in [\alpha, \frac{1}{2}]$. Usarem $\beta = 0.3$.

$$X_0+\lambda_u D = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + 0.1553 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.3107 \\ -1.8446 \end{bmatrix}$$

$$\nabla f(X_0+\lambda_u D)D = [\,-0.3074 \quad 0.5403\,] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -0.07448$$

$$\nabla f(X_0+\lambda_u D)D = \underline{-0.07448} > \beta \nabla f(X_0)D = 0.3(-\frac{3}{2}) = \underline{-0.45} \Rightarrow \boxed{\text{Es satisfà AG2}}$$

COMPROVACIÓ DE LA 2^A COND. D'ARMIJO-GOLDSTEIN A $X_0+\frac{1}{8}D$:

$$X_0+\lambda_u D = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \frac{1}{8} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.25 \\ -1.875 \end{bmatrix}$$

$$\nabla f(X_0+\lambda_u D)D = [\,-0.4416 \quad 0.5388\,] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -0.3444$$

$$\nabla f(X_0+\frac{1}{8}D)D = \underline{-0.3444} > \beta \nabla f(X_0)D = 0.3(-\frac{3}{2}) = \underline{-0.45} \Rightarrow \boxed{\text{Es satisfà AG2}}$$

Solució del problema 23.

GRADIENT I COMPROVACIÓ DE QUE D ÉS DE DESCENS: $\nabla f(X_0)D < 0$

$$\nabla f(X)' = \begin{bmatrix} 3x_1^2 + x_2 + \frac{2}{x_2} \\ x_1 - \frac{2x_1}{x_2^2} \end{bmatrix} \quad ; \quad \nabla f(X)' = \begin{bmatrix} 6 \\ -1 \end{bmatrix} \quad ; \quad \nabla f(X_0)D = [\,6 \quad -1\,] \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 1 > 0$$

$D = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ <u>no és de descens</u>. Una possible direcció de descens és $P = -D = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

EXPLORACIÓ LINEAL PER AJUST QUADRÀTIC EN LA DIRECCIÓ P.

$$q(X_0+\lambda P) = a\lambda^2 + b\lambda + c \quad , \quad q'(X_0+\lambda P) = 2a\lambda + b \quad , \quad q''(X_0+\lambda P) = 2a$$

$$\underline{\lambda = 0}: \quad \begin{cases} q(X_0) = f(X_0) = \boxed{c = 4} \\ q'(X_0) = \nabla f(X_0)P = \boxed{b = -1} \end{cases}$$

$$\underline{\lambda = 1}: \quad \begin{cases} X_0 + P = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow f(X_0+P) = 4 \\ q(X_0+P) = a + b + c = f(X_0+P) = 4 \Rightarrow \boxed{a = 1} \end{cases}$$

$$\min q(X_0+\lambda P) \Rightarrow \boxed{\lambda_q = -\frac{b}{2a} = -\frac{-1}{2\times 1} = \frac{1}{2}} \quad (q''(X_0+\lambda_q P) = 2 > 0 : \text{mínim})$$

$$X_q = X_0 + \lambda_q P = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3/2 \end{bmatrix} \Rightarrow f(X_0+\lambda_q P) = 23/6$$

EXPLORACIÓ LINEAL PER AJUST CÚBIC EN LA DIRECCIÓ P

$$u(X_0+\lambda P) = a\lambda^3 + b\lambda^2 + c\lambda + d \quad , \quad u'(X_0+\lambda P) = 3a\lambda^2 + 2b\lambda + c$$

$$\underline{\lambda = 0}: \quad \begin{cases} u(X_0) = f(X_0) = \boxed{d=4} \\ u'(X_0) = \nabla f(X_0)P = \boxed{c=-1} \end{cases}$$

$$\underline{\lambda = 1}: \quad u(X_0+P) = a+b+c+d = f(X_0+P) = 4 \Rightarrow \underline{a+b=1}$$

$$\underline{\lambda = \lambda_q = \frac{1}{2}}: \quad u(X_0+\lambda_q P) = \frac{a}{8} + \frac{b}{4} + \frac{c}{2} + d = f(X_0+\lambda_q P) = \frac{23}{6} \Rightarrow \underline{a+2b = \frac{8}{3}}$$

resolent $\begin{cases} a+b=1 \\ a+2b=8/3 \end{cases}$ s'obté $\boxed{a = -\frac{2}{3}\ ,\ b = \frac{5}{3}}$. Es calcula ara λ_u:

$$u'(X_0+\lambda_u P) = 3a\lambda_u^2 + 2b\lambda_u + c = 0 \quad , \quad -2\lambda_u^2 + 2\frac{5}{3}\lambda_u - 1 = 0$$

$$\boxed{\lambda_u = \frac{-5/3 + \sqrt{(5/3)^2 - 2}}{-2} = 0.392375} \quad \text{(signe + devant del radicand} \Rightarrow \text{mínim)}$$

$$X_u = X_0 + \lambda_u P = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0.392375 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.392375 \end{bmatrix} \Rightarrow f(X_0+\lambda_u P) = 3.82877$$

ACCEPTABILITAT DE X_q I DE X_u.

Podem aplicar la primera condició d'Armijo-Goldstein:

$$f(X_0+\lambda P) \le f(X_0) + \alpha\lambda\nabla f(X_0)P$$

amb $\alpha > 0$ però petita (p.ex.: $\alpha = 0.01$).

a) puntX_q : $\begin{cases} \lambda_q = \frac{1}{2}\ ;\ f(X_0+\lambda_q P) = \frac{23}{6} = 3.8\hat{3} \\ 3.8\hat{3} < 4 + 0.01 \times \frac{1}{2} \times (-1) = 3.995 \quad \Rightarrow \ \underline{\text{es satisfà la condició}} \end{cases}$

b) puntX_u : $\begin{cases} \lambda_u = 0.392375 \quad ; \quad f(X_0+\lambda_u P) = 3.82877 \\ 3.82877 < 4 + 0.01 \times 0.392375 \times (-1) = 3.996 \quad \Rightarrow \underline{\text{es satisfà la condició}} \end{cases}$

També es podria aplicar la segona condició d'Armijo-Goldstein:

$$\nabla f(X_0+\lambda P)P \ge \beta\nabla f(X_0)P$$

amb $1 > \beta > \alpha$ (p.ex.:$\beta = 0.5$):

a) punt X_q : $\begin{cases} \lambda_q = \frac{1}{2}\ ;\ \nabla f(X_0+\lambda_q P)' = \begin{bmatrix} 5.8\hat{3} \\ 0.\hat{1} \end{bmatrix} \\ [\,5.8\hat{3} \quad 0.\hat{1}\,] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.\hat{1} > \frac{1}{2} \times (-1) = -0.5 \quad \Rightarrow \ \underline{\text{es satisfà la condició.}} \end{cases}$

b) punt X_u : $\begin{cases} \lambda_u = 0.392375 \quad ; \quad \nabla f(X_0+\lambda_u P)' = \begin{bmatrix} 5.8288 \\ -0.0316 \end{bmatrix} \\ [\,5.8288 \quad -0.0316\,] \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -0.0316 \\ -0.0316 > \frac{1}{2} \times (-1) = -0.5 \Rightarrow \ \underline{\text{es satisfà la condició.}} \end{cases}$

3.6 Mètode del Gradient.

Solució del problema 24.

Per simplificar la notació escriurem el gradient $\nabla f(x_k)^T$ com g_k. El gradient de la funció quadràtica ve donat per $g_k = Qx_k - b$ Per comprovar que g_k i g_{k+1} són perpendiculars només cal veure que que $g_k^T g_{k+1} = 0$. Sabent que

$$x_{k+1} = x_k - \frac{g_k^T g_k}{g_k^T Q g_k} g_k$$

directament obtenim el resultat anterior fent:

$$\begin{aligned} g_k^T g_{k+1} &= g_k^T (Qx_{k+1} - b) = \\ &= g_k^T \big(Q(x_k - \frac{g_k^T g_k}{g_k^T Q g_k} g_k) - b\big) \\ &= g_k^T (Qx_k - b) - \frac{g_k^T g_k}{g_k^T Q g_k} g_k^T Q g_k \\ &= g_k^T g_k - g_k^T g_k = 0 \end{aligned}$$

Solució del problema 25.

Podem observar com ambdues funcions poden ser escrites de forma matricial com:

$$\begin{aligned} f_1(x) &= \frac{1}{2} x^T Q_1 x - b_1^T x = \frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ f_2(x) &= \frac{1}{2} x^T Q_2 x - b_2^T x = \frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

i els gradients vénen donats per:

$$\begin{aligned} \nabla f_1(x)^T &= Q_1 x - b_1 = \begin{pmatrix} 2x_1 + x_2 - 1 \\ x_1 + 2x_2 - 2 \end{pmatrix} \\ \nabla f_2(x)^T &= Q_2 x - b_1 = \begin{pmatrix} 10x_1 \\ 2x_2 - 2 \end{pmatrix} \end{aligned}$$

Passem ara a solucionar cada apartat.

i) Els punts mínims els trobarem anul·lant el vector gradient:

$$\nabla f_1(x)^T = 0 \quad \Longrightarrow \quad \begin{matrix} 2x_1 + x_2 = 1 \\ x_1 + 2x_2 + = 2 \end{matrix} \quad \Longrightarrow \quad x^* = (0 \;\; 1)$$

$$\nabla f_2(x)^T = 0 \quad \Longrightarrow \quad \begin{matrix} 10x_1 = 0 \\ 2x_2 = 2 \end{matrix} \quad \Longrightarrow \quad x^* = (0 \;\; 1)$$

S'observa com ambdues funcions tenen el mateix punt mínim.

ii) Sabem que la convergència del mètode del gradient serà millor quan més petit sigui el terme $\beta = [(\lambda_M - \lambda_m)/\lambda_M + \lambda_m)]^2$ (on λ_m és el valor propi menor i λ_M el valor propi major de la matriu Q). Podem trobar els valors propis per a Q_1 i Q_2 de la següent manera:

$$\det(Q_1 - \lambda) = 0 \quad \Leftrightarrow \quad \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3 = 0 \quad \Rightarrow \quad \lambda_m = 1, \quad \lambda_M = 3$$

$$\det(Q_2 - \lambda) = 0 \quad \Leftrightarrow \quad \begin{vmatrix} 10-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = (10-\lambda)(2-\lambda) \quad \Rightarrow \quad \lambda_m = 2, \quad \lambda_M = 10$$

Calculem ara els termes β_1 (per Q_1) i β_2 (per Q_2):

$$\beta_1 = \left(\frac{3-1}{3+1}\right)^2 = 1/4 = 0.25 \qquad \beta_1 = \left(\frac{10-2}{10+2}\right)^2 = 64/144 = 0.4444...$$

Per tant, amb la funció $f_1(x)$, a cada iteració del mètode del gradient reduim (com a poc) en una quarta part la nostra distància al punt òptim, mentre que amb la funció $f_2(x)$ reduim (com a mínim) en una mica més de la meitat aquesta mateixa distància a x^*.

iii) Realitzarem les dues iteracions del mètode del gradient per a cada $f_i(x)$. El vector gradient el denotarem, per comoditat, per $g_k = Q_i x_k - b_i$, on i serà 1 o 2 segons la funció considerada:

$f_1(x)$

$$\begin{aligned} x_0 &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ g_0 = Q_1 x_0 - b_1 &= \begin{pmatrix} 5 \\ 4 \end{pmatrix} \\ \alpha_0 = \tfrac{g_0^T g_0}{g_0^T Q_1 g_0} &= 0.33607 \\ x_1 = x_0 - \alpha_0 g_0 &= \begin{pmatrix} 0.31967 \\ 0.65574 \end{pmatrix} \\ g_1 = Q_1 x_1 - b_1 &= \begin{pmatrix} 0.29508 \\ -0.36885 \end{pmatrix} \\ \alpha_1 = \tfrac{g_1^T g_1}{g_1^T Q_1 g_1} &= 0.97619 \\ x_2 = x_1 - \alpha_1 g_1 &= \begin{pmatrix} 0.031616 \\ 1.015808 \end{pmatrix} \end{aligned}$$

$f_2(x)$

$$\begin{aligned} x_0 &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} \\ g_0 = Q_2 x_0 - b_2 &= \begin{pmatrix} 20 \\ 2 \end{pmatrix} \\ \alpha_0 = \tfrac{g_0^T g_0}{g_0^T Q_2 g_0} &= 0.10080 \\ x_1 = x_0 - \alpha_0 g_0 &= \begin{pmatrix} -0.015968 \\ 1.798403 \end{pmatrix} \\ g_1 = Q_2 x_1 - b_2 &= \begin{pmatrix} -0.15968 \\ 1.59681 \end{pmatrix} \\ \alpha_1 = \tfrac{g_1^T g_1}{g_1^T Q_2 g_1} &= 0.48095 \\ x_2 = x_1 - \alpha_1 g_1 &= \begin{pmatrix} 0.060831 \\ 1.030415 \end{pmatrix} \end{aligned}$$

Observem com en ambdós casos la seqüència generada s'apropa a $x^* = (0 \ \ 1)^T$.

iv) Per comprovar que les direccions obtingudes són perpendiculars, només cal veure que $g_1^T g_0 = 0$ (per a les dues funcions). Per a $f_1(x)$ tenim que

$$g_1^T g_0 = \begin{pmatrix} 0.29508 & -0.36885 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \end{pmatrix} = 5 \cdot 0.29508 - 4 \cdot 0.36885 \approx 0$$

Per la seva banda per a $f_2(x)$:

$$g_1^T g_0 = \begin{pmatrix} -0.15968 \\ 1.59681 \end{pmatrix} \begin{pmatrix} 20 \\ 2 \end{pmatrix} = -20 \cdot 0.15968 + 2 \cdot 1.59681 \approx 0$$

Solució del problema 26.

Prenem $\lambda_4 = 3$ i determinem U_4 : $Q - \lambda_4 \mathbf{I} = \begin{bmatrix} 1/2 & 1 & 0 & -1/2 \\ 1 & 5/2 & -1/2 & -2 \\ 0 & -1/2 & 5/2 & -1 \\ -1/2 & -2 & -1 & 9/2 \end{bmatrix}$

Per a resoldre el sistema homogeni $(Q - \lambda_4 \mathbf{I})U_4 = \mathbf{0}$ prenem $u_{44} = 1$ i considerem les tres primeres equacions :

$$\begin{bmatrix} 1/2 & 1 & 0 \\ 1 & 5/2 & -1/2 \\ 0 & -1/2 & 5/2 \end{bmatrix} \begin{bmatrix} u_{41} \\ u_{42} \\ u_{43} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 2 \\ 1 \end{bmatrix}$$

triangularitzant i resolent s'obté:

$$\begin{bmatrix} 1/2 & 1 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & -1/2 & 5/2 \end{bmatrix} \begin{bmatrix} u_{41} \\ u_{42} \\ u_{43} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ 1 \end{bmatrix} ; \begin{bmatrix} 1/2 & 1 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} u_{41} \\ u_{42} \\ u_{43} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ 2 \end{bmatrix} \Rightarrow \underline{U_4 = \begin{bmatrix} -5 \\ 3 \\ 1 \\ 1 \end{bmatrix}}$$

Qualsevol punt $X_0 = X^* + \alpha U_4$ satisfà la condició demanada. Podem prendre $X_0 = X^* + U_4$:

$$X_0 = X^* + U_4 = \begin{bmatrix} 1 \\ 9 \\ -9 \\ 4 \end{bmatrix} + \begin{bmatrix} -5 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 12 \\ -8 \\ 5 \end{bmatrix} \quad ; \quad \boxed{X_0 = \begin{bmatrix} -4 \\ 12 \\ -8 \\ 5 \end{bmatrix}}$$

Solució del problema 27.

PASSA PEL MÈTODE DEL GRADIENT A PARTIR DE X_0

Donada una funció quadràtica: $q(X) = \frac{1}{2}X'QX - B'X$ amb $X, B \in \mathbb{R}^n$, i $Q = Q' \in \mathbb{R}^{n \times n}$, l'expressió de l'algorisme del gradient és:

$$X_{k+1} = X_k - \alpha_k G_k \quad \text{amb } G_k = QX_k - B \text{ el gradient a } X_k \text{ i } \alpha_k = \frac{{G_k}'G_k}{{G_k}'QG_k}$$

$$G_0 = QX_0 - B = \begin{bmatrix} 19/3 & -8\sqrt{2}/3 \\ -8\sqrt{2}/3 & 11/3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 19 \\ -8\sqrt{2} \end{bmatrix} = \begin{bmatrix} -\frac{8\sqrt{2}+57}{3} \\ \frac{11+24\sqrt{2}}{3} \end{bmatrix}$$

$${G_0}'G_0 = 742,94082 \quad {G_0}'QG_0 = 6680,2367 \quad \alpha_0 = 0,1112147$$

$$X_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - 0,1112147 \begin{bmatrix} -\frac{8\sqrt{2}+57}{3} \\ \frac{11+24\sqrt{2}}{3} \end{bmatrix} = \begin{bmatrix} 2,5324972 \\ -0,6660379 \end{bmatrix}$$

FUNCIÓ D'ERROR A X_0 I A X_1

Per tal de calcular $\mathcal{E}(X)=\frac{1}{2}(X-X^*)'Q(X-X^*)$ necessitem coneixer X^* el qual és solució de $QX^*=B$. Resolent el sistema:

$$\begin{bmatrix} 19/3 & -8\sqrt{2}/3 \\ -8\sqrt{2}/3 & 11/3 \end{bmatrix} X^* = \begin{bmatrix} 19 \\ -8\sqrt{2} \end{bmatrix} \quad \text{obtenim } X^* = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

Fent els càlculs per a X_0 i per al punt trobat X_1 trobem que $\mathcal{E}(X_0)=\frac{1}{2}(X_0-X^*)'Q(X_0-X^*)$=41,647042 i que $\mathcal{E}(X_1)=\frac{1}{2}(X_1-X^*)'Q(X_1-X^*)$=0,33111455 .

FITA SUPERIOR A LA TAXA DE CONVERGÈNCIA

Sabem que:

$$\frac{\mathcal{E}(X_{k+1})}{\mathcal{E}(X_k)} \leq \left(\frac{A-a}{A+a}\right)^2 \quad \text{on } A \text{ i } a \text{ són el més gran i el més petit valor propi de } Q.$$

Hem de calcular els valors propis de Q i ho farem resolent el polinomi característic $\det(Q-\lambda\mathbb{1})=0$ (de grau 2 en aquest cas).

$$(\frac{19}{3}-\lambda)(\frac{11}{3}-\lambda)-\frac{128}{9}=0 \implies \lambda^2-10\lambda+9=0 \implies \begin{cases} \lambda_1 = A = 9 \\ \lambda_2 = a = 1 \end{cases}$$

COMPROVACIÓ DE LA TAXA DE CONVERGÈNCIA LOCAL

Efectivament:

$$\frac{\mathcal{E}(X_1)}{\mathcal{E}(X_0)} = \frac{0,33111455}{41,647042} = 0,00795 \leq \left(\frac{A-a}{A+a}\right)^2 = \left(\frac{9-1}{9+1}\right)^2 = 0,64$$

COMPROVACIÓ DE L'ORTOGONALITAT DE G_1 I (X_1-X_0)

$$G_1 = QX_1 - B = \begin{bmatrix} 19/3 & -8\sqrt{2}/3 \\ -8\sqrt{2}/3 & 11/3 \end{bmatrix} \begin{bmatrix} 2,5324972 \\ -0,6660379 \end{bmatrix} - \begin{bmatrix} 19 \\ -8\sqrt{2} \end{bmatrix} = \begin{bmatrix} -0,4490648 \\ -0,6790754 \end{bmatrix}$$

$${G_1}'(X_1-X_0) = \begin{bmatrix} -0,4490648 & -0,6790754 \end{bmatrix} \begin{bmatrix} 2,5324972 \\ -1,6660379 \end{bmatrix} = 0 \quad \text{tal com diu la teoria}$$

Solució del problema 31.

Càlcul de la direcció de decens::

Es calcula en primer lloc la direcció de moviment del mètode del gradient:

$$f(x) = x_1^2 + x_2^2 - (x_2 - x_3)^3 + x_1x_2x_3 \quad ; \quad x^0 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

$$\nabla f(x)' = \begin{bmatrix} 2x_1 + x_2x_3 \\ 2x_2 - 3(x_2 - x_3)^2 + x_1x_3 \\ 3(x_2 - x_3)^2 + x_1x_2 \end{bmatrix} \quad ; \quad \nabla f(x^0) = \begin{bmatrix} 3 \\ -3 \\ -1 \end{bmatrix}$$

Direcció de moviment del mètode del gradient: $\boxed{d^0 = -\nabla f(x^0) = \begin{bmatrix} -3 \\ 3 \\ 1 \end{bmatrix}}$

Exploració lineal::

A continuació es realitza exploració lineal per Fibonacci amb $N = 4$ i interval d'incertessa $[0,\ 0.5]$, de londitud inicial $d_1 = 0.5$. La funció univaluada sobre la que s'aplicarà Fibonacci és:

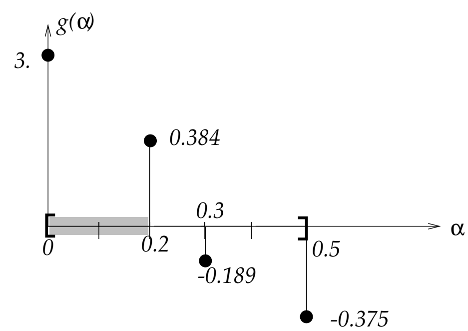
$$g(\alpha) = f(x^0 + \alpha d^0) = f(-3\alpha + 1, 3\alpha - 1, \alpha - 1)$$
$$g(\alpha) = -17\alpha^3 + 33\alpha^2 - 19\alpha + 3$$

El seu valor sobre els extrems de l'interval d'incertessa és $g(0) = 3$, $g(0.5) = -0.375$. Els valors dels números de Fibonacci que s'usaran són:

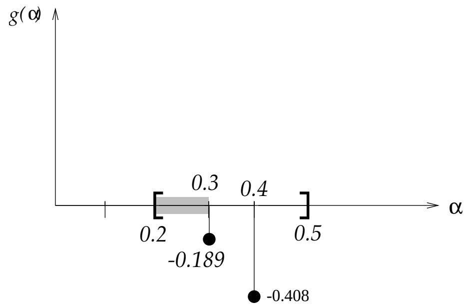
$$F_0 = F_1 = 1 \quad ; \quad F_2 = 2 \quad ; \quad F_3 = 3 \quad ; \quad F_4 = 5$$

i les iteracions del procés són:

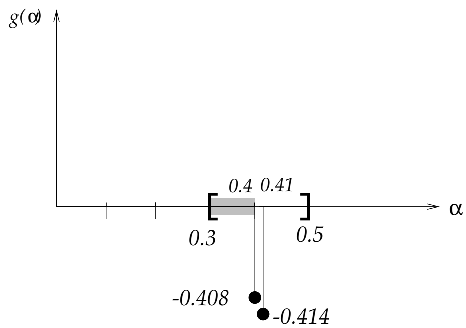
- **Primera iteració:**
  - $\ast$ $\alpha_1 = 0 + \frac{F_3}{F_4}d_1 = \frac{3}{5}0.5 = 0.3 \quad ; \quad g(0.3) = -0.189$
  - $\ast$ $\alpha_2 = 0.5 - \frac{F_3}{F_4}d_1 = 0.5 - 0.3 = 0.2 \quad ; \quad g(0.2) = 0.384$
  - $\ast$ $g(0.2) > g(0.5)$: es descarta $[0, 0.2[$. El nou interval d'incertessa és $[0.2, 0.5]$.



- **Segona iteració:**
  - $\ast$ $\alpha_3 = \alpha_2 + \frac{F_2}{F_4}d_1 = 0.4 \quad ; \quad g(0.4) = -0.408$
  - $\ast$ $g(0.4) < g(0.5)$: es descarta $[0.2, 0.3[$. El nou interval d'incertessa és $[0.3, 0.5]$



- **Tercera iteració:**
  - $*$ $\alpha_4 = \alpha_1 + \frac{F_1}{F_4} d_1 + 0.001 = 0.41 \quad ; \quad g(0.41) = -0.414$
  - $*$ $g(0.4) > g(0.41)$: es descarta $[0.3, 0.4[$. L'interval final d'incertessa és $[0.4, 0.5]$.



Es pot prendre com a longitud de pas el punt mig $\boxed{\alpha^0 = 0.45}$.

Actualització de les variables:

$$x^1 = x^0 + \alpha^0 d^0 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} + 0.45 \begin{bmatrix} -3 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.35 \\ 0.35 \\ -0.35 \end{bmatrix}$$

Test d'optimalitat:

$$\|\nabla f(x^1)\|_2 = \| \left[-0.89250 \quad -0.23250 \quad 2.30750 \right] \|_2 \approx 2.485 > \epsilon = 0.1 \Rightarrow x^1 \neq x^*$$

Solució del problema 33.

1.- *Condició d'òptim:* $\|\nabla(x^0)\|_2 = 2\sqrt{2} \not< \epsilon = 0.5$

2.- *Direcció de moviment:* $d^0 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$

3.- *Exploració lineal per Fibonacci:* $\alpha^0 = 1/6$

4.- *Actualització de variables:* $x^1 = x^0 + \alpha^0 d^0 = \begin{bmatrix} 4/3 \\ 5/3 \end{bmatrix}$

Solució del problema 35.

a) $x^0 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$; $\nabla f(x) = (\, 2x_1 + 2x_2 - 2 \quad 3x_2 + 2x_1 - 2\,)'$; $\nabla f(x^0) = (\, -4 \quad -4\,)'$

1.- Condició d'aturada: $\|\nabla f(x^0)\|_2| = \sqrt{16+16} = \sqrt{32} = 5.6568 > \epsilon = 0.5$

2.- Direcció de descens : $d^0 = -\nabla f(x^0) = (\, 4 \quad 4\,)'$

3.- Exploració lineal : $g(\alpha) = f(x^0 + \alpha d^0) = f(-1+4\alpha, 4\alpha) = 72\alpha^2 - 32\alpha + 3$ $g(\alpha)' = 144\alpha - 32$; $g(\alpha^*)' = 144\alpha^* - 32 = 0$; $\alpha^* = 32/144 \approx 0.2$

4.- Actualització : $x^1 = x^0 + \alpha^* d^0 = (\, -1/9 \quad 8/9\,)'$

b) $N = 4$: $F_0 = 1, F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5$, $g(\alpha) = 72\alpha^2 - 32\alpha + 3$, $\alpha^0 = 0$ Interval d'incertessa inicial : $g(0) = 3$, $g(1) = 43 > g(0) \Rightarrow \mathrm{I}^0 = [0, 1]$

1.- Primera iteració : $d_2 = F_3/F_4 d_1 = 3/5 = 0.6$; $\alpha_1 = 0 + 0.6 = 0.6$, $\alpha_2 = 1 - 0.6 = 0.4$

$g(\alpha_1) = 9.72 \quad > \quad g(\alpha_2) = 1.72 \quad \Rightarrow$ es descarta $[0.6, 1]$: $\mathrm{I}^2 = [0, 0.6]$

2.- Segona iteració : $d_3 = F_2/F_4 d_1 = 2/5 = 0.4$; $\alpha_3 = 0.6 - 0.4 = 0.2$

$g(\alpha_3) = -0.52 \quad < \quad g(\alpha_2) = 1.72 \quad \Rightarrow$ es descarta $[0.4, 0.6]$: $\mathrm{I}^3 = [0, 0.4]$

3.- Tercera iteració : $d_4 = F_1/F_4 d_1 = 1/5 = 0.2$; $\alpha_4 = 0 + 0.2 + 0.01 = 0.21$

$g(\alpha_4) = -0.54 \quad < \quad g(\alpha_3) = -0.52 \quad \Rightarrow$ es descarta $[0, 0.2]$: $\mathrm{I}^4 = [0.2, 0.4]$

Interval final d'incertessa : $\boxed{\mathrm{I}^4 = [0.2, 0.4]}$. Es pren com a longitud de pas òptima el millor extrem de I^4: $\boxed{\alpha^* = 0.2}$

Solució del problema 36.

a) $x^1 = [\, -2.8571 \quad 4.8571 \quad -0.5928\,]$

b) Si es calcula la definició de la matriu Hessiana (constant, doncs $f(x)$ és quadràtica), s'obté que és indefinida. Així doncs, el problema **(PNL)** no té solució, doncs no n'hi ha cap punt de $\mathbb{R}^3$ que satisfaci les condicions necessàries de segon ordre ($\nabla f(x) = [0]$ i $\nabla^2 f(x)$ semidef $+$).

c) Atés que el problema no té solució, no té sentit calcular la taxa de convergència. Si es calculés s'obtindria $\beta = 7.77 > 1$, que no té sentit.

3.7 Mètode de Newton.

Solució del problema 38.

Sense fer cap càlcul podem adonar-nos de que, aplicant el mètode de Newton a les funcions $f_1(x)$ i $f_2(x)$ del problema 25, en una iteració trobarem el punt òptim. Aquestes funcions poden ser escrites de forma matricial com

$$f(x) = \frac{1}{2} x^T Q x - b^T x$$

El gradient i la hessiana són directament:

$$\nabla f(x)^T = Qx - b \qquad\qquad \nabla^2 f(x) = Q$$

Sabem que el punt òptim x^* és aquell que anul·la el gradient, és a dir $x^* = Q^{-1}b$ (tal i com es va poder veure al problema 25, aquest punt òptim és igual per als dos problemes i correspon al punt $x^* = (0 \;\; 1)^T$).

Si apliquem el mètode de Newton, començant a iterar des d'un punt x_0 qualsevol, tenim que el primer punt x_1 és (usant els valors de gradient i hessiana abans trobats):

$$\begin{aligned} x_1 &= x_0 - \nabla^2 f(x_0)^{-1}\nabla f(x_0)^T \\ &= x_0 - Q^{-1}(Qx_0 - b) = x_0 - Q^{-1}Qx_0 - Q^{-1}b = x_0 - x_0 - Q^{-1}b \\ &= Q^{-1}b \end{aligned}$$

amb el qual tenim que $x_1 = Q^{-1}b = x^* = (0 \;\; 1)^T$. Hem comprovat, doncs, sense fer cap càlcul, que si apliquéssim el mètode de Newton a $f_1(x)$ i $f_2(x)$ (o a qualsevol altra funció quadràtica) obtindríem el punt solució en una iteració.

Solució del problema 39.

VECTOR GRADIENT, MATRIU HESSIANA I LA SEVA TRIANGULARITZACIÓ.

$$\nabla f(X_0)' = \begin{bmatrix} 3x_1^2 + x_2^2 + 6x_3 + 3x_5 \\ 2x_2 + 2x_1x_2 + 2x_4 + 2x_2x_5 \\ 11x_3 + 6x_1 + 6x_5 \\ 3x_4^2 + 2x_2 + x_5 \\ 5x_5 + 3x_1 + x_2^2 + 6x_3 + x_4 \end{bmatrix} \quad ; \quad \nabla f(X_0)' = \begin{bmatrix} -21/4 \\ 1 \\ -8 \\ 3/4 \\ -4 \end{bmatrix}$$

$$\nabla^2 f(X) = \begin{bmatrix} 6x_1 & 2x_2 & 6 & 0 & 3 \\ 2x_2 & 2x_1 + 2x_5 + 2 & 0 & 2 & 2x_2 \\ 6 & 0 & 11 & 0 & 6 \\ 0 & 2 & 0 & 6x_4 & 1 \\ 3 & 2x_2 & 6 & 1 & 5 \end{bmatrix} \quad ; \quad \nabla^2 f(X_0) = \begin{bmatrix} 3 & 0 & 6 & 0 & 3 \\ 0 & 3 & 0 & 2 & 0 \\ 6 & 0 & 11 & 0 & 6 \\ 0 & 2 & 0 & 3 & 1 \\ 3 & 0 & 6 & 1 & 5 \end{bmatrix}$$

$$\nabla^2 f(X_0) = TDT' = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & \mathbf{0} & \\ 2 & 0 & 1 & & \\ 0 & 2/3 & 0 & 1 & \\ 1 & 0 & 0 & 3/5 & 1 \end{bmatrix} \begin{bmatrix} 3 & & & & \\ & 3 & & \mathbf{0} & \\ & & -1 & & \\ & \mathbf{0} & & 5/3 & \\ & & & & 7/5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ & 1 & 0 & 2/3 & 0 \\ & & 1 & 0 & 0 \\ & \mathbf{0} & & 1 & 3/5 \\ & & & & 1 \end{bmatrix}$$

CORRECCIÓ UNIFORME δ DE LUENBERGER A LA MATRIU DIAGONAL D

PER TAL QUE EL MÍNIM VALOR DIAGONAL RESULTANT SIGUI $\epsilon = 0.1 > 0$.
La correció necessària és $\delta = 1.1$:

$$\tilde{D} = D+\delta I = \begin{bmatrix} 3 & & & & \\ & 3 & & \mathbf{0} & \\ & & -1 & & \\ & \mathbf{0} & & 5/3 & \\ & & & & 7/5 \end{bmatrix} + 1.1 \begin{bmatrix} 1 & & & & \\ & 1 & & \mathbf{0} & \\ & & 1 & & \\ & \mathbf{0} & & 1 & \\ & & & & 1 \end{bmatrix} = \begin{bmatrix} 4.1 & & & & \\ & 4.1 & & \mathbf{0} & \\ & & 0.1 & & \\ & \mathbf{0} & & 2.7\hat{6} & \\ & & & & 2.5 \end{bmatrix}$$

RESOLUCIÓ DE $T\tilde{D}T'P = -\nabla f(X_0)$.

$$\begin{bmatrix} 1 & & & & \\ 0 & 1 & & \mathbf{0} & \\ 2 & 0 & 1 & & \\ 0 & 2/3 & 0 & 1 & \\ 1 & 0 & 0 & 3/5 & 1 \end{bmatrix} \begin{bmatrix} 4.1 & & & & \\ & 4.1 & & \mathbf{0} & \\ & & 0.1 & & \\ & \mathbf{0} & & 2.7\hat{6} & \\ & & & & 2.5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ & 1 & 0 & 2/3 & 0 \\ & & 1 & 0 & 0 \\ & \mathbf{0} & & 1 & 3/5 \\ & & & & 1 \end{bmatrix} P = \begin{bmatrix} 21/4 \\ -1 \\ 8 \\ -3/4 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & & & & \\ 0 & 1 & & \mathbf{0} & \\ 2 & 0 & 1 & & \\ 0 & 2/3 & 0 & 1 & \\ 1 & 0 & 0 & 3/5 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 21/4 \\ -1 \\ 8 \\ -3/4 \\ 4 \end{bmatrix} \Rightarrow \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 21/4 \\ -1 \\ -5/2 \\ -1/12 \\ -6/5 \end{bmatrix}$$

$$\begin{bmatrix} 4.1 & 0 & 8.2 & 0 & 4.1 \\ & 4.1 & 0 & 2.7\hat{3} & 0 \\ & & 0.1 & 0 & 0 \\ & \mathbf{0} & & 2.7\hat{6} & 1.66 \\ & & & & 2.5 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} 21/4 \\ -1 \\ -5/2 \\ -1/12 \\ -6/5 \end{bmatrix} \Rightarrow \boxed{\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} 10.11951 \\ -0.31384 \\ -5.8 \\ 0.11337 \\ -0.48 \end{bmatrix}}$$

COMPROVACIÓ DE QUE P ÉS DE DESCENS: $\nabla f(X_0)P < 0$.

$$\begin{bmatrix} -21/4 & 1 & -8 & 3/4 & -4 \end{bmatrix} \begin{bmatrix} 10.11951 \\ -0.31384 \\ -5.8 \\ 0.11337 \\ -0.48 \end{bmatrix} = -5.01375 < 0$$

EXPLORACIÓ LINEAL PER AJUST QUADRÀTIC EN LA DIRECCIÓ P OBTINGUDA.

$$q(X_0 + \lambda P) = a\lambda^2 + b\lambda + c \quad , \quad q'(X_0 + \lambda P) = 2a\lambda + b \quad , \quad q''(X_0 + \lambda P) = 2a$$

$$\underline{\lambda=0}: \quad \begin{cases} q(X_0)=f(X_0)=\boxed{c=-8.25} \\ q'(X_0)=\nabla f(X_0)P=\boxed{b=-5.01375} \end{cases}$$

$$\underline{\lambda=1}: \quad \begin{cases} X_0+P=\begin{bmatrix}1/2\\0\\-1\\1/2\\0\end{bmatrix}+\begin{bmatrix}10.11951\\-0.31384\\-5.8\\0.11337\\-0.48\end{bmatrix}=\begin{bmatrix}10.61951\\-0.31384\\-6.8\\0.61337\\-0.48\end{bmatrix}\Rightarrow f(X_0+P)=1024.2122 \\ q(X_0+P)=a+b+c=f(X_0+P)=1024.2122\Rightarrow \boxed{a=1037.4759} \end{cases}$$

$$\min q(X_0+\lambda P)\Rightarrow \boxed{\lambda_q=-\frac{b}{2a}=-\frac{-5.01375}{2\times 1037.4759}=0.00241632} \quad (\ddot{q}(X_0+\lambda_q P)= 2074.95>0: \text{mínim})$$

EXPLORACIÓ LINEAL PER AJUST QUADRÀTIC EN LA DIRECCIÓ $\hat{P}=[2\quad 0\quad -1\quad 0\quad 0]'$.

$$q(X_0+\lambda\hat{P})=\hat{a}\lambda^2+\hat{b}\lambda+\hat{c} \quad , \quad q'(X_0+\lambda\hat{P})=2\hat{a}\lambda+\hat{b} \quad , \quad q''(X_0+\lambda\hat{P})=2\hat{a}$$

$$\nabla f(X_0)\hat{P}=[-21/4\quad 1\quad -8\quad 3/4\quad -4]\begin{bmatrix}2\\0\\-1\\0\\0\end{bmatrix}=-2.5$$

$$\underline{\lambda=0}: \quad \begin{cases} q(X_0)=f(X_0)=\boxed{\hat{c}=-8.25} \\ q'(X_0)=\nabla f(X_0)\hat{P}=\boxed{\hat{b}=-2.5} \end{cases}$$

$$\underline{\lambda=1}: \quad \begin{cases} X_0+\hat{P}=\begin{bmatrix}1/2\\0\\-1\\1/2\\0\end{bmatrix}+\begin{bmatrix}2\\0\\-1\\0\\0\end{bmatrix}=\begin{bmatrix}5/2\\0\\-2\\1/2\\0\end{bmatrix}\Rightarrow f(X_0+\hat{P})=7.75 \\ q(X_0+\hat{P})=\hat{a}+\hat{b}+\hat{c}=f(X_0+\hat{P})=7.75\Rightarrow \boxed{\hat{a}=7.75} \end{cases}$$

$$\min q(X_0+\lambda\hat{P})\Rightarrow \boxed{\hat{\lambda}_q=-\frac{\hat{b}}{2\hat{a}}=-\frac{-2.5}{2\times 7.5}=\frac{1}{6}} \quad (\ddot{q}(X_0+\hat{\lambda}_q\hat{P})=15>0: \text{mínim})$$

Solució del problema 43.

OBTENCIÓ DIRECCIÓ DE DESCENS AMB DIRECCIONS DE CORBATURA NEGATIVA

Sent el gradient zero i l'hessiana indefinida, cal utilitzar com a direcció de descens una direcció de corbatura negativa D_{CN}. Trobarem D_{CN} utilitzant el procediment que es deriva de la triangularització de Gill-Murray de l'hessiana: $\tilde{S}\tilde{D}\tilde{S}'D_{CN}=E_m$, sent E la matriu diagonal d'elements e_{jj} (correccions a les diagonals). L'obtenció de D_{CN} és per resolució del sistema $\tilde{S}'D_{CN}=E_m$, sent E_m la columna m de la matriu unitat i m l'índex de la més gran correcció e_{mm} a E.

Obtenció de E per diferència entre $\tilde{S}\tilde{D}\tilde{S}'$ i $\nabla^2 f(X_0)$ i determinació de e_{mm}.

$$\tilde{S}\tilde{D}\tilde{S}' = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ 0 & 0 & 1 & \\ 2/3 & 1/6 & 6 & 1 \end{bmatrix} \begin{bmatrix} 6 & & & \\ & 12 & & \\ & & 1/6 & \\ & & & 9 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 2/3 \\ & 1 & 0 & 1/6 \\ & & 1 & 6 \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 6 & -6 & 0 & 4 \\ -6 & 18 & 0 & -2 \\ 0 & 0 & 1/6 & 1 \\ 4 & -2 & 1 & 18 \end{bmatrix}$$

$$\nabla f(X)' = \begin{bmatrix} 3x_1^2 - 3x_2^2 + 4x_1x_4 \\ -6x_1x_2 + x_3x_4 + 6 \\ x_2x_4 \\ 2x_1^2 + x_2x_3 \end{bmatrix} \quad ; \quad \nabla^2 f(X) = \begin{bmatrix} 6x_1 + 4x_4 & -6x_2 & 0 & 4x_1 \\ -6x_2 & -6x_1 & x_4 & x_3 \\ 0 & x_4 & 0 & x_2 \\ 4x_1 & x_3 & x_2 & 0 \end{bmatrix}$$

$$\nabla f(X_0)' = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad ; \quad \nabla^2 f(X_0) = \begin{bmatrix} 6 & -6 & 0 & 4 \\ -6 & -6 & 0 & -2 \\ 0 & 0 & 0 & 1 \\ 4 & -2 & 1 & 0 \end{bmatrix}$$

$$E = \tilde{S}\tilde{D}\tilde{S}' - \nabla^2 f(X_0) = \begin{bmatrix} 0 & & & \\ & 24 & & \\ & & 1/6 & \\ & & & 18 \end{bmatrix}$$

d'on $m = 2$, ja que $e_{mm} = e_{22} = 24$, i el sistema a resoldre és:

$$\tilde{S}' D_{CN} = \begin{bmatrix} 1 & -1 & 0 & 2/3 \\ & 1 & 0 & 1/6 \\ & & 1 & 6 \\ & & & 1 \end{bmatrix} \begin{bmatrix} d_{CN_1} \\ d_{CN_2} \\ d_{CN_3} \\ d_{CN_4} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \Rightarrow D_{CN} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

comprovació:

$$D'_{CN} \nabla^2 f(X_0) D_{CN} = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 6 & -6 & 0 & 4 \\ -6 & -6 & 0 & -2 \\ 0 & 0 & 0 & 1 \\ 4 & -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = -12 < 0$$

Nota: també $-D_{CN}$ és una direcció de corbatura negativa a partir de X_0 on $\nabla f(X_0)' = \mathbf{0}$.

EXPLORACIÓ LINEAL PER AJUST CÚBIC AMB $u(\lambda) = a\lambda^3 + b\lambda^2 + c\lambda + d$.

$$u(\lambda) = a\lambda^3 + b\lambda^2 + c\lambda + d \quad , \quad u'(\lambda) = 3a\lambda^2 + 2b\lambda + c \quad , \quad u''(\lambda) = 6a\lambda + 2b$$

a) En la direcció D_{CN}:

$$\underline{\lambda = 0}: \begin{cases} u(0) = f(X_0) = \boxed{d = 4} \\ u'(0) = \nabla f(X_0) D_{CN} = \boxed{c = 0} \end{cases}$$

$$\underline{\lambda = 1}: \begin{cases} X_0 + D_{CN} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ -2 \\ 0 \end{bmatrix} \Rightarrow f(X_0 + D_{CN}) = -4 \\ \nabla f(X_0 + D_{CN}) = [0 \quad -18 \quad 0 \quad 4] \quad ; \quad \nabla f(X_0 + D_{CN}) D_{CN} = -18 \\ \left.\begin{array}{l} u(1) = a + b + c + d = f(X_0 + D_{CN}) = -4 \\ u'(1) = 3a + 2b + c = \nabla f(X_0 + D_{CN}) D_{CN} = -18 \end{array}\right\} \left.\begin{array}{l} a + b = -8 \\ 3a + 2b = -18 \end{array}\right\} \Rightarrow \boxed{\begin{array}{l} a = -2 \\ b = -6 \end{array}} \end{cases}$$

Tenim doncs $u(\lambda) = -2\lambda^3 - 6\lambda^2 + 4$. Igualant a zero $u'(\lambda_u) = 3a\lambda_u^2 + 2b\lambda_u + c = 0$ arribem a $\lambda_u(-6\lambda_u - 12) = 0$ que té dos solucions: $\lambda_u = 0$ i $\lambda_u = -2$. Examinant amb $u''(\lambda) = 6a\lambda + 2b$ el caràcter de màxim o mínim d'aquests punts veiem que:

$\underline{\lambda_u = 0}$ dóna $u''(0) = -12 < 0$ i correspon a un **màxim**.

$\underline{\lambda_u = -2}$ dóna $u''(-2) = 24 - 12 = 12 > 0$ i correspon a un **mínim**.

La passa a aplicar en la direcció D_{CN} és doncs $\boxed{\lambda_u = -2}$.

Fenòmens atípics observats:

i) La passa $\lambda_u = -2$ és negativa, malgrat que no ens ha d'extranyar perquè tant és de descens la direcció D_{CN} com $-D_{CN}$ a partir de X_0 amb $\nabla f(X_0) = \mathbf{0}$ (vegeu figura 1).

ii) Anomenant $X_1 = X_0 + \lambda_u D_{CN}$ tenim que:

$$X_1 = X_0 + \lambda_u D_{CN} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -2 \\ 0 \end{bmatrix} \quad ; \quad \begin{cases} f(X_1) = -4 = \\ = f(X_0 + D_{CN}) = -4 \end{cases}$$

(vegeu figura 1). L'ajust sembla indicar que és millor explorar cap a valors positius de $\lambda > 1$, que no pas quedar-se a $\lambda_u = -2$, ja que a $\lambda = 1$ la derivada és negativa (-18), i la funció baixa de valor augmentant λ.

iii) Si haguessim aplicat el criteri de considerar com a mínim de l'ajust cúbic la solució amb radicand positiu de $u'(\lambda) = -6\lambda^2 - 12\lambda = 0$, $\lambda = \dfrac{12 \pm \sqrt{(-12)^2}}{-12}$ també hauriem obtingut que $\lambda = -2$ correspon a un nínim, i que $\lambda = 0$ correspon a un màxim, però cal anar amb compte de no canviar els signes a $u'(\lambda)$, (quan $c = 0$ sembla natural considerar $6\lambda^2 + 12\lambda = 0$), perquè això provocaria que aquest criteri no fos vàlid.

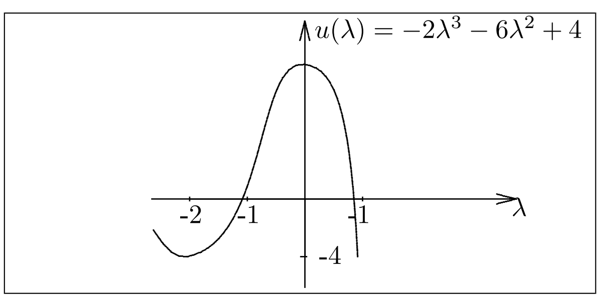


Figura 1

b) En la direcció $\tilde{D}_{CN} = -D_{CN} = [\,-1 \quad -1 \quad 0 \quad 0\,]'$:
Igualment al cas **a)** seguim tenint que $\boxed{c = 0\,,\; d = 4}$.

$$\underline{\lambda = 1}: \begin{cases} X_0 + \tilde{D}_{CN} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2 \\ 0 \end{bmatrix} \Rightarrow f(X_0 + \tilde{D}_{CN}) = 0 \\ \nabla f(X_0 + \tilde{D}_{CN}) = [\,0 \quad 6 \quad 0 \quad 0\,] \quad ; \quad \nabla f(X_0 + \tilde{D}_{CN})\tilde{D}_{CN} = -6 \\ \left.\begin{array}{l} u(1) = a + b + c + d = f(X_0 + \tilde{D}_{CN}) = 0 \\ u'(1) = 3a + 2b + c = \nabla f(X_0 + \tilde{D}_{CN})\tilde{D}_{CN} = -6 \end{array}\right\} \left.\begin{array}{l} a + b = -4 \\ 3a + 2b = -6 \end{array}\right\} \Rightarrow \boxed{\begin{array}{l} a = 2 \\ b = -6 \end{array}} \end{cases}$$

Tenim doncs $u(\lambda) = -2\lambda^3 - 6\lambda^2 + 4$. Igualant a zero $u'(\lambda_u) = 6\lambda_u^2 - 12\lambda_u = 0$ arribem a $\lambda_u(6\lambda_u - 12) = 0$ que té dos solucions: $\lambda_u = 0$ i $\lambda_u = 2$. Examinant amb $u''(\lambda) = 12\lambda - 12$ el caràcter de màxim o mínim d'aquests punts veiem que:

$\underline{\lambda_u = 0}$: dóna $u''(0) = -12 < 0$ i correspon a un **màxim**.

$\underline{\lambda_u = 2}$: dóna $u''(2) = 24 - 12 = 12 > 0$ i correspon a un **mínim** (vegeu figura 2).

La passa a aplicar en la direcció $\tilde{D}_{CN}$ és doncs $\boxed{\lambda_u = 2}$, equivalent a la passa amb D_{CN}. Això però és un fenòmen particular d'aquest problema, perquè en general, partint de X_0 amb $\nabla f(X_0) = \mathbf{0}$, les passes en una direcció i en la contrària, s'assemblen però poden no coincidir, ja que la variació de la funció objectiu en una direcció i en la contraria no son sempre iguals.

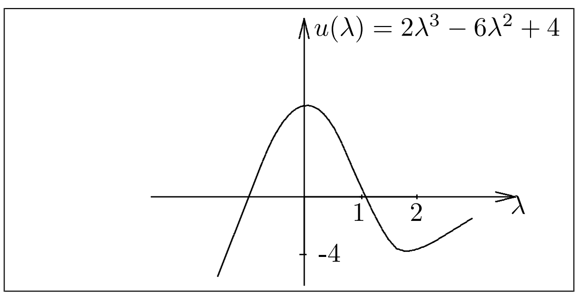


Figura 2

AJUST CÚBIC EMPRANT LA DIRECCIÓ $P = [\,7/3 \quad 1 \quad 0 \quad 0\,]'$.
Igualment al cas **a)** anterior seguim tenint que $\boxed{c = 0\,,\; d = 4}$.

$$\underline{\lambda = 1}: \begin{cases} X_0 + P = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 7/3 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 10/3 \\ 2 \\ -2 \\ 0 \end{bmatrix} \Rightarrow f(X_0 + P) = 244/27 \\ \nabla f(X_0 + P) = [\,64/3 \quad -34 \quad 0 \quad 164/9\,] \quad ; \quad \nabla f(X_0 + P)P = 142/9 \\ \left.\begin{array}{l} u(1) = a + b + c + d = f(X_0 + P) = 244/27 \\ u'(1) = 3a + 2b + c = \nabla f(X_0 + P)P = 142/9 \end{array}\right\} \left.\begin{array}{l} a + b = 136/27 \\ 3a + 2b = 142/9 \end{array}\right\} \Rightarrow \boxed{\begin{array}{l} a = \frac{154}{27} \\ b = -\frac{2}{3} \end{array}} \end{cases}$$

Tenim doncs $u(\lambda) = \frac{154}{27}\lambda^3 - \frac{2}{3}\lambda^2 + 4$. Igualant a zero $u'(\lambda_u) = \frac{154}{9}\lambda_u^2 - \frac{4}{3}\lambda_u = 0$ arribem a

$\lambda_u(\frac{154}{9}\lambda_u - \frac{4}{3}) = 0$ que té dos solucions: $\lambda_u = 0$ i $\lambda_u = \frac{6}{77}$. Examinant amb $u''(\lambda) = \frac{308}{9}\lambda - \frac{4}{3}$ el caràcter de màxim o mínim d'aquests punts veiem que:

$\underline{\lambda_u = 0}$: dóna $u''(0) = -\frac{4}{3} < 0$ i correspon a un **màxim**.

$\underline{\lambda_u = \frac{6}{77}}$: dóna $u''(\frac{6}{77}) = \frac{4}{3} > 0$ i que correspon a un **mínim**.

Anomenant $X_p = X_0 + \lambda_u P$ tenim que:

$$X_p = X_0 + \lambda_u P = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \frac{6}{77}\begin{bmatrix} 7/3 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2/11 \\ 83/77 \\ -2 \\ 0 \end{bmatrix} \quad ; \quad f(X_p) = \frac{34624}{5929} < f(X_0 + P) = \frac{142}{9}$$

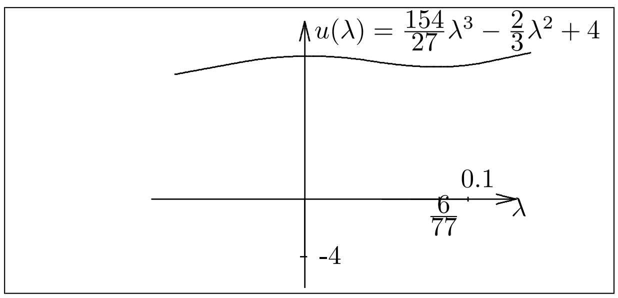


Figura 3

4 Solucions dels problemes de PNL amb constriccions.

4.1 Condicions d'òptim per a problemes amb constriccions.

Solució del problema 47.

a) Una corba diferenciable sobre una superficie $\mathcal{S}$ s'expressa com una parametrització de les variables x_i en funció del paràmetre t. L'enunciat ens dona l'expressió paramètrica corresponent a la primera variable: hem de trobat l'expressió paramètrica de x_2 i substituir per $t = 0$:

$$\left.\begin{aligned} x_1(t) = t \\ h(x(t)) = x_1^2(t) - x_2(t) = 0 \end{aligned}\right\} t^2 - x_2(t) = 0\,,\; x_2(t) = t^2 \Rightarrow x(t) = \begin{bmatrix} t \\ t^2 \end{bmatrix}\;,\; \boxed{x(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

b)

$$\underline{\left.\frac{\mathrm{d}}{\mathrm{dt}} f(x(t))\right|_{t=0}} = \nabla f(x(0)) \left.\frac{\mathrm{d}}{\mathrm{dt}} x(t)\right|_{t=0} = [\,1 \quad 0\,] \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \underline{1}$$

$$\nabla f(x) = [\,(x_2+1)e^{x_1(x_2+1)} \quad x_1 e^{x_1(x_2+1)}\,] \ , \ \nabla f(x(0)) = [\,1 \quad 0\,]$$

$$\frac{\mathrm{d}}{\mathrm{dt}} x(t) = \begin{bmatrix} \frac{\mathrm{d}}{\mathrm{dt}} x_1(t) \\ \frac{\mathrm{d}}{\mathrm{dt}} x_2(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 2t \end{bmatrix} \ , \ \left.\frac{\mathrm{d}}{\mathrm{dt}} x(t)\right|_{t=0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

c) $x(0)$ no pot ser un punt estacionari del problema indicat, doncs una condició necessària és que $\left.\frac{\mathrm{d}}{\mathrm{dt}} f(x(t))\right|_{t=0} = 0$ per a totes les corbes diferenciables de $\mathcal{S}$ que passin per $x(0)$, i aquesta condició no es satisfà per la corba trobada a l'apartat a). Si fessim un moviment a partir de $x(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ al llarg de la corba $x(t) = \begin{bmatrix} t \\ t^2 \end{bmatrix}$ amb $t < 0$ $f(x)$ milloraria.

Solució del problema 48.

$$\begin{array}{l} \min\ f(x_1,x_2) \\ \text{s.a.}: g_1(x_1,x_2) \le b_1 \\ \qquad g_2(x_1,x_2) \ge b_2\,; \\ \qquad x_1 \ge 0 \\ \qquad x_2 \ge 0 \end{array} \left\{ \begin{array}{ll} i) & \nabla f(x_1,x_2) + \mu_1 \nabla g_1(x_1,x_2) + \mu_2 \nabla g_(x_1,x_2) + \mu_3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mu_4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \\ ii) & \mu_1(b_1 - g_1(x_1,x_2)) = 0 \quad , \quad \mu_2(b_2 - g_2(x_1,x_2)) = 0 \\ & \mu_3 x_1 = 0 \quad , \quad \mu_4 x_2 = 0 \\ iii) & \mu_1 \ge 0 \quad , \quad \mu_2 \le 0 \quad , \quad \mu_3 \le 0 \quad , \quad \mu_4 \le 0 \end{array} \right.$$

Solució del problema 49.

a) $\nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} >$ def $+ \Rightarrow$ la funció objectiu és convexa i la regió factible és un conjunt convex, doncs està definit a partir de constriccions lineals. Aixídoncs es trata d'un problema convex, satisfent-se llavors que qualsevol solució factible que cumpleixi les condicions necessàries de primer ordre és un mínim global.

b) Passem el problema **(PNL)** a la forma estàndar canviant el signe de la primera constricció:

$$\textbf{(PNL)} \left\{ \begin{array}{llccccc} \min & \text{f(x)} = & (x_1-1)^2 & + & (x_2-2)^2 & & \\ & & & & & & \\ \text{subj.a}: & & x_1 & - & x_2 & \le & 1 \\ & & x_1 & + & x_2 & \le & 2 \\ & & -x_1 & & & \le & 0 \\ & & & - & x_2 & \le & 0 \end{array} \right.$$

Condicions de Khun i Tucker:

$$i) \left\{ \begin{array}{lcccccccccc} 2x_1 - 2 & + & \mu_1 & + & \mu_2 & - & \mu_3 & & & = & 0 \\ 2x_2 - 4 & - & \mu_1 & + & \mu_2 & & & - & \mu_4 & = & 0 \end{array} \right.$$

$$ii) \left\{ \begin{array}{lcl} \mu_1(1 - x_1 + x_2) = 0 & ; & \mu_3 x_1 = 0 \\ \mu_2(2 - x_1 - x_2) = 0 & ; & \mu_4 x_2 = 0 \end{array} \right.$$

iii) $\mu_i \geq 0 \quad , \quad i = 1, \ldots, 4$

c) $x_1, x_2 > 0 \Rightarrow \mu_3 = \mu_4 = 0$. Substituïnt a les condicions de Kuhn i Tucker trobades a l'apartat anterior s'obté:

$$
\begin{aligned}
2x_1 + \mu_1 + \mu_2 &= 2 && (1)\\
2x_2 - \mu_1 + \mu_2 &= 4 && (2)\\
\mu_1(1 - x_1 + x_2) &= 0 && (3)\\
\mu_2(2 - x_1 - x_2) &= 0 && (4)\\
\mu_1 \geq 0,\ \mu_2 &\geq 0 && (5)
\end{aligned}
$$

Provant diversos valors de μ_1 y μ_2 s'obté:

a) $\mu_1 = \mu_2 = 0 \ \Rightarrow \ x_1 = 1,\ x_2 = 2 \ \Rightarrow \underline{x_1 + x_2 \nleq 2 \ \Rightarrow \text{viola la segunda restricción}}$
b) $\mu_1 \neq 0, \mu_2 \neq 0 \ \Rightarrow \ \mu_2 = 1,\ \underline{\mu_1 = -2 \ngeq 0 \ \Rightarrow \text{viola } (5)}$
c) $\mu_1 \neq 0,\ \mu_2 = 0 \ \Rightarrow \ x_1 = 2,\ x_2 = 1 \ \Rightarrow \underline{x_1 + x_2 \nleq 2 \ \Rightarrow \text{viola la segunda restricción}}$
d) $\mu_1 = 0,\ \mu_2 \neq 0 \ \Rightarrow \ \underline{x_1 = 1/2,\ x_2 = 3/2 \ \mu_1 = 0,\ \mu_2 = 1 \ > \ \text{óptimo}}$

Solució del problema 50.

Calcularem prèviament el gradient de la funció objectiu i la matriu Jacobiana de les constriccions:

$$
\begin{aligned}
f(x) =& (x_1x_2 - 3)^2 - 4x_1x_2 + 2(x_1x_3 - 1)^3 - \frac{15}{2}x_2 + 5x_3\\
\nabla f(x)' =& \begin{bmatrix} 2(x_1x_2 - 3)x_2 - 4x_2 + 6(x_1x_3 - 1)^2x_3 \\ 2(x_1x_2 - 3)x_1 - 4x_1 - 15/2 \\ 6(x_1x_3 - 1)^2x_1 + 5 \end{bmatrix}\\
g_1(x) =& 6x_1^2 - 3x_2x_3 - x_3^2 - 5 \quad ; \quad \nabla g_1(x)' = \begin{bmatrix} 12x_1 \\ -3x_3 \\ -3x_2 - 2x_3 \end{bmatrix}\\
g_2(x) =& x_1^2 + \frac{1}{2}x_2 + x_3^2 - 4 \quad ; \quad \nabla g_2(x)' = \begin{bmatrix} 2x_1 \\ 1/2 \\ 2x_3 \end{bmatrix}
\end{aligned}
$$

a) **Condicions de Kuhn i Tucker:** Sigui x factible **(PNL)** ($g(x) \leq 0$). Llavors si x és mínim local de **(PNL)** existeix el vector $\mu \in \mathbb{R}^2$ que satisfà les següents condicions:

i) $\nabla f(x) + \sum_{j=1}^{2} \mu_j \nabla g_j(x) = 0$. En el nostre cas:

$$
\left.
\begin{aligned}
2(x_1x_2 - 3)x_2 - 4x_2 + 6(x_1x_3 - 1)^2x_3 + 12x_1\mu_1 + 2x_1\mu_2 =& 0\\
2(x_1x_2 - 3)x_3 - 4x_1 - 15/2 - 3x_3\mu_1 + (1/2)\mu_2 =& 0\\
6(x_1x_3 - 1)^2x_1 + 5 - 3x_1\mu_1 - 2x_3\mu_1 + 2x_3\mu_2 =& 0
\end{aligned}
\right\} \qquad (1)
$$

ii) $\mu_j g_j(x) = 0 \; , \; j = 1, \ldots, 2$:

$$\left.\begin{aligned} \mu_1(6x_1^2 - 3x_2x_3 - x_3^2 - 5) =& 0 \\ \mu_2(x_1^2 + \frac{1}{2}x_2 + x_3^2 - 4) =& 0 \end{aligned}\right\}$$

iii) $\mu_j \geq 0 \; , \; j = 1, 2$

b) Hem de comprovar les condicions de l'apartat anterior sobre el punt $x = [\,0 \quad -2 \quad 1\,]'$:

$$g(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \end{bmatrix} = \begin{bmatrix} 0 \\ -4 \end{bmatrix} \leq 0 \Rightarrow x \text{ factible } \textbf{(PNL)}$$

ii) La segona constricció és inactiva ($g_2(x) = -4 \neq 0$): això implica que el multiplicador μ_2 s'ha d'anular: $\underline{\mu_2 = 0}$.

i) Substituïnt els valors de $x = [\,0 \quad -2 \quad 1\,]'$ i $\mu_2 = 0$ al sistema d'equacions (1) s'obté:

$$\left.\begin{aligned} 26 =& 0 \\ -15/2 + 6\mu_1 =& 0 \\ 5 + 4\mu_1 =& 0 \end{aligned}\right\} \text{sistema incompatible} \Rightarrow \boxed{x = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \text{ no és mínim de } \textbf{(PNL)}}$$

Solució del problema 52.

b) El punt $x^0 = [\,0 \quad 1\,]'$ no és un mínim local de **(PNL)** perquè viola la condició de signe dels multiplicadors de Lagrange ($\lambda_1 = -3/5$).

c) El conjunt factible és convex, doncs la primera constricció és de menor o igual i la funció que la defineix és convexa ($\nabla^2 g_1(x) = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$). Les constriccions de signe $x \geq 0$ també defineixen conjunts convexos, i la intersecció de conjunts convexos és un conjunt convex. La matriu Hessiana de la funció objectiu és:

$$\nabla^2 f(x) = \begin{bmatrix} 2x_2 + 2 & 2(x_1 - 1) - 2 \\ 2(x_1 - 1) - 2 & 2 \end{bmatrix}$$

Hem d'estudiar la definició d'aquesta matriu:

* $\Delta_1 = 2x_2 + 2 \geq 0 \Rightarrow x_2 \geq -1$. Això sempre es satisfà per a tota solució factible **(PNL)**.
* $\Delta_2 = -4x_1^2 + 16x_1 + 4x_2 - 12 \geq 0$: aquesta condició es pot satisfer en funció dels valors de x_1 i x_2. Si provem, però, el la solució factible de l'anterior apartat $x^0 = [\,0 \quad 1\,]'$ s'obté $\Delta_2 = -8 \not\geq 0$. Així doncs, $f(x)$ no és convexa ($\nabla^2 f(x)$ no és semidefinida positiva) sobre la regió factible, i, llavors, **(PNL)** no és un problema de programació convexa.

Solució del problema 53.

a) Calulem les primeres derivades de la funció objectiu i de les constriccions:

$$\nabla f(x)' = \begin{bmatrix} (x_2 + 1)e^{x_1(x_2+1)} \\ x_1 e^{x_1(x_1+1)} \end{bmatrix} \quad ; \quad \nabla g_1(x) = \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix} \quad ; \quad \nabla g_2(x) = \begin{bmatrix} 6x_1^2 - 4x_1 \\ -1 \end{bmatrix}$$

Condicions de Kuhn i Tucker: n'hi ha dos multiplicadors de Lagrange μ_1 i μ_2 associats al problema **(PNL)**

i) $\nabla f(x) + \sum_{j=1}^{2} \mu_j \nabla g_j(x) = 0$

$$(x_2+1)e^{x_1(x_2+1)} + 2\mu_1 x_1 + \mu_2 6x_1^2 - 4\mu_2 x_1 = 0 \quad (1)$$
$$x_1 e^{x_1(x_2+1)} - \mu_1 - \mu_2 = 0 \quad (2)$$

ii) $\mu_j g_j(x) = 0\ ,\ j = 1, \ldots, 2$:

$$\mu_1(x_1^2 - x_2) = 0 \quad (3)$$
$$\mu_2(2x_1^3 - 2x_1^2 - x_2 + 1) = 0 \quad (4)$$

iii) $\mu_j \geq 0\ ,\ j = 1, 2$

$$\mu_1 \geq 0 \quad (5)$$
$$\mu_2 \geq 0 \quad (6)$$

Condicions de factibilitat:

$$x_1^2 - x_2 \leq 0 \quad (7)$$
$$2x_1^3 - 2x_1^2 - x_2 + 1 \leq 0 \quad (8)$$

Per a resoldre **(PNL)** a partir de les equacions (1)-(8) caldria explorar els quatre casos següents:

1) Dues constriccions actives: hem de trobar un punt solució del sistema (1)-(8) sencer.
2) Cap constricció activa: de (3) i (4) s'obté $\mu_1 = \mu_2 = 0$ i el sistema es redueix a calcular $\nabla f(x) = 0$, és a dir, un punt estacionari de $f(x)$ a l'interior de la regió factible de **(PNL)**.
3) $g_1(x) = 0$, $g_2(x) < 0$: llavors $\mu_2 = 0$ i s'ha de resoldre (1), (2), (5), (7) i (8), el·liminant μ_2 de (1) i (2).
4) $g_1(x) < 0$, $g_2(x) = 0$: de (3) tenim $\mu_1 = 0$ i s'ha de resoldre (1), (2), (6) i (8), el·liminant μ_1 de (1) i (2).

La major dificultat consisteix en que els sistemes que s'obtenen en els quatre casos anteriors són sistemes d'equacions no lineals que, en general, no es poden resoldre de forma directa.

b) A la sortida de GINO observem que el valor d'un dels multiplicadors de Lagrange, μ_2, és negatiu, violant la condició (6): el punt $\tilde{x} = \begin{bmatrix} -0.455418 \\ 0.396285 \end{bmatrix}$ no és mínim local de **(PNL)**.

c) Comprovem primer que el punt $\tilde{x}$ és regular. Calculem la jacobiana de les constriccions actives J = {2}:

$$\nabla g_2(\tilde{x}) = [\,3.0661 \quad -1\,] \qquad \text{rang}(\nabla g_2(\tilde{x})) = 1 \Rightarrow \text{rang complet}$$

El subespai tangent sobre $\tilde{x}$ regular és:

$$\text{M} = \{y \mid \nabla g_j(\tilde{x})y = 0\ ,\ j \in \text{J}\} = \{y \mid \ [\,3.0661 \quad -1\,] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0\} = \{y \mid y_2 = 3.0661 y_1\}$$

La dimensió del subespai M és dim(M)=1, llavors qualsevol vector de M serveix com a base d'aquest subespai:

$$\boxed{Z = \begin{bmatrix} 1 \\ 3.0661 \end{bmatrix}}$$

Solució del problema 55.

Per tal de comprovar si x^* és òptim de **(PNL)** hem de comprovar les condicions suficients de segon ordre (que inclouen, recordeu, a les de primer ordre). El primer pas és trobar els multiplicadors de Lagrange μ_1 i μ_2 associats a x^*, si es que aquests existeixen. Per tal de calcular aquests multiplicadors usarem les condicions necessàries de primer ordre (condicions de Khun i Tucker). Calculem el gradient i la matriu Jacobiana sobre el punt x^*:

$$\nabla f(x)' = \begin{bmatrix} (x_2+1)e^{x_1(x_2+1)} \\ x_1 e^{x_1(x_2+1)} \end{bmatrix} \quad , \quad \nabla f(x^*)' = \begin{bmatrix} 0.6691 \\ -0.2675 \end{bmatrix}$$

$$\nabla g(x) = \begin{bmatrix} 2x_1 & -1 \\ 6x_1^2 - 4x_1 & -1 \end{bmatrix} \quad , \quad \nabla g(x^*) = \begin{bmatrix} -1 & -1 \\ 3.5 & -1 \end{bmatrix}$$

S'observa que la matriu Jacobiana és de rang complet, sent doncs x^* regular. Les dues constriccions són actives sobre x^*. Si plantejem la primera condició de Khun i Tucker tenim:

$$\nabla f(x^*) + [\,\mu_1^* \quad \mu_2^*\,]\,\nabla g(x^*) = 0$$

$$\left.\begin{array}{r} 0.6691 - \mu_1^* + 3.5\mu_2^* = 0 \\ -0.2676 - \mu_1^* - \mu_2^* = 0 \end{array}\right\} \boxed{\mu_1^* = -0.05947 \;,\; \mu_2^* = -0.20816}$$

Les constriccions del problema **(PNL)** estan plantejades com de ≥ 0. Això implica que la condició de signe sobre els multiplicador (tercera condició de Kuhn i Tucker) és $\mu^* \leq 0$, que és satisfeta pel vector trobat. Així doncs, el parell $x^* = [\,-1/2 \quad 1/4\,]'$, $\mu^* = [\,-0.05947 \quad -0.20816\,]$ satisfan les condicions necessàries de primer ordre. Comprovem ara les de segon ordre, calculant la definició de la matriu Hessiana de la funció Lagrangiana sobre el subespai tangent M. Calculem primer $\nabla_x^2 \mathcal{L}(x^*, \mu^*)$:

$$\nabla^2 f(x^*) = \begin{bmatrix} 0.8363 & 0.2007 \\ 0.2007 & 0.1338 \end{bmatrix} , \nabla^2 g_1(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} , \nabla^2 g_2(x^*) = \begin{bmatrix} -10 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\nabla_x^2 \mathcal{L}(x^*, \mu^*) = \nabla^2 f(x^*) + \sum_{j=1}^{2} \mu_j \nabla^2 g_j(x^*) =$$

$$= \begin{bmatrix} 0.8363 & 0.2007 \\ 0.2007 & 0.1338 \end{bmatrix} - 0.05947 \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} - 0.20816 \begin{bmatrix} -10 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2.799 & 0.2007 \\ 0.2007 & 0.1338 \end{bmatrix} \left.\begin{array}{l} \Delta_1 = 2.799 > 0 \\ \Delta_2 = 0.334 > 0 \end{array}\right\} \text{def } +$$

Donat que $\nabla_x^2 \mathcal{L}(x^*, \mu^*)$ és def + sobre $\mathbb{R}^2$, també ho serà sobre qualsevol subespai de $\mathbb{R}^2$. En particular, serà def + sobre el subespai tangent M. Així doncs, el punt x^* és un mínim local estricte de **(PNL)**.

Solució del problema 56.

S'ha de comprovar la definició de l'Hessiana de la funció Lagrangiana sobre x^*, λ^*, $\nabla_x^2 \mathcal{L}(x^*, \lambda^*)$

sobre el subespai M:

$$\nabla_x^2 \mathcal{L}(x^*,\lambda^*) = \nabla^2 f(x^*) + \lambda^{*\prime}\nabla^h(x^*) = \begin{bmatrix} 0 & 1/2 & -1/4 \\ 1/2 & 0 & -1/4 \\ -1/4 & -1/4 & 1/4 \end{bmatrix} + (-1/4)\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \end{bmatrix} =$$

$$= \begin{bmatrix} -1/2 & 1/2 & -1/4 \\ 1/2 & -1/2 & -1/4 \\ -1/4 & -1/4 & 1/4 \end{bmatrix}$$

La matriu $\nabla_x^2 \mathcal{L}(x^*,\lambda^*)$ és indefinida. Hem de comprovar la definició de $Z'\nabla_x^2 \mathcal{L}(x^*,\lambda^*)Z$ on Z és una base del subespai M $= \{y \mid \nabla h(x^*)y = 0\}$. S'han de trobar dos vectors z_1 i z_2 pertanyents a M i linealment independents. La condició de pertinença a M és:

$$\nabla h(x^*) = [\,2 \quad 2 \quad -1\,] \quad ; \quad \nabla h(x^*)y = 0 \Rightarrow \underline{2y_1 + 2y_2 - y_3 = 0} \qquad (1)$$

Trobem ara dos vectors vectors linealment independent que satisfacin (1):

$$\left.\begin{array}{l} y_3 = 0 \quad ; \quad y_1 = 1 \Rightarrow y_2 = -1 \quad ; \quad z_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\ y_3 = 1 \quad ; \quad y_1 = 0 \Rightarrow y_2 = 1/2 \quad ; \quad z_2 = \begin{bmatrix} 0 \\ 1/2 \\ 1 \end{bmatrix} \end{array}\right\} Z = \begin{bmatrix} 1 & 0 \\ -1 & 1/2 \\ 0 & 1 \end{bmatrix}$$

Calulem $Z'\nabla_x^2 \mathcal{L}(x^*,\lambda^*)Z$ i la seva definició:

$$Z'\nabla_x^2 \mathcal{L}(x^*,\lambda^*)Z = \begin{bmatrix} -2 & 1/2 \\ 1/2 & -1/8 \end{bmatrix} \left.\begin{array}{l} \Delta_1 = -2 < 0 \\ \Delta_2 = 0 \end{array}\right\} \text{semidef}-$$

Donat que $Z'\nabla_x^2 \mathcal{L}(x^*,\lambda^*)Z$ és semidef - , el punt estacionari x^* pot ser màxim local, tot i que no es pot assegurar.

Solució del problema 59.

El problema a solucionar és

$$\begin{array}{rl} \min & f(x) = x_1^2 + x_2^2 \\ \text{subj. a} & h(x) = x_1 + x_2 - a = 0 \end{array}$$

i) Si ho solucionem introduint el canvi de variable $x_2 = a - x_1$ a la funció objectiu, obtenim el següent problema d'una variable sense restriccions:

$$\min u(x) = x_1^2 + (a - x_1)^2 = 2x_1^2 - 2ax_1 + a^2$$

Buscant els punts que anul·len $u'(x)$ obtenim:

$$u'(x) = 4x_1 - 2a \quad 4x_1 - 2a = 0 \Rightarrow x_1 = a/2$$

Com que $u''(x) = 4 > 0$ tenim que el punt $x_1 = a/2$ és un mínim del nostre problema sense restriccions. El costat de l'altre quadrat és $x_2 = a - a/2 = a/2$. Els costats que donen, doncs, l'àrea mínima són $x_1 = x_2 = a/2$ (l'àrea mínim és $a^2/2$).

ii) Solucionem-ho ara aplicant el mètode dels multiplicadors de Lagrange. Definim la funció lagrangiana:

$$L(x, \lambda) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - a)$$

Un punt per ser mínim ha de satisfer en primer lloc:

$$\left.\begin{aligned} \nabla_{x_1} L(x,\lambda) &= 2x_1 + \lambda = 0 \\ \nabla_{x_2} L(x,\lambda) &= 2x_2 + \lambda = 0 \\ \nabla_{\lambda} L(x,\lambda) &= x_1 + x_2 - a = 0 \end{aligned}\right\} \Rightarrow \lambda = -a, \quad x_1 = x_2 = a/2$$

El punt anterior és, doncs, l'únic candidat a ser un mínim. Per estar segurs que ho és, cal que observem la condició suficient de segon ordre:

$$y^T[\nabla^2_{xx} L(x,\lambda)]y > 0 \quad \forall y : \nabla h(x) y = 0$$

La matriu hessiana $\nabla^2_{xx} L(x,\lambda)$ és:

$$\nabla^2_{xx} L(x,\lambda) = \nabla_x \left(2x_1 + \lambda \quad 2x_1 + \lambda \right) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Com que aquesta matriu és definida positiva, tenim que per a tot $y \neq 0$ es verifica que $y^T[\nabla^2_{xx} L(x,\lambda)]y > 0$. Per tant es satisfà la condició suficient de segon ordre, amb el qual garantim que $x_1 = x_2 = a/2$ és un mínim local (i en aquest cas, a més global, podeu raonar per què?) del nostre problema amb restriccions. Observem com la solució coincideix amb la calculada considerant el problema sense restriccions.

Solució del problema 60.

i) Considerant el radi r i l'alçària h com a variables del problema, aquest pot ser formulat com:

$$\begin{aligned} &\max \quad \pi r^2 h \quad \equiv -\min f(r,h) = -\pi r^2 h \\ &\text{subj. a} \\ &\qquad h(r,h) = 2\pi r h + 2\pi r^2 - 6\pi = 0 \\ &\qquad r \geq 0 \quad h \geq 0 \end{aligned}$$

Treballarem, per comoditat, amb el problema equivalent de minimització. Les restriccions de no-negativitat segur que són inactives a l'òptim, ja que si fossin actives tindríem que $r = 0$ o $h = 0$, amb el qual el volum seria igual a 0 (el qual, evidentment, no correspon a un volum màxim).

ii) La funció lagrangiana és

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

Les condicions necessàries (de primer ordre) que ha de satisfer un mínim són:

$$\begin{aligned} \nabla_r L(r,h,\lambda) &= -2\pi r h + \lambda 2\pi h + \lambda 4\pi r = 0 \\ \nabla_h L(r,h,\lambda) &= -\pi r^2 + \lambda 2\pi r = 0 \\ \nabla_\lambda L(r,h,\lambda) &= 2\pi r h + 2\pi r^2 - 6\pi = 0 \end{aligned}$$

Simplificant la segona condició anterior tenim que $r(2\lambda - r) = 0$, el qual implica que $r = 0$ o $r = 2\lambda$. Abans hem justificat que r no pot ser 0. Per tant ens quedem només amb que $r = 2\lambda$. Substituint $\lambda = r/2$ a la primer condició, s'obté que $r(2r - h) = 0$, el qual implica que $r = 0$ o $h = 2r$. Descartem, com abans, que $r = 0$, i ens quedem només amb $h = 2r$. Substituint $h = 2r$ a la tercera condició, i operant, s'arriba finalment a que $r^2 = 1$, el qual implica que $r = 1$ o $r = -1$. Descartem que $r = -1$ (un radi negatiu no té sentit) i ens quedem amb $r = 1$. Com que hem vist abans que $h = 2r$ i que $\lambda = r/2$, tenim que el punt $r = 1,\ h = 2,\ \lambda = 1/2$ satisfà les condicions de primer ordre, i és, doncs, un candidat a ser òptim.

Per estar segurs que és un punt òptim (mínim en aquest cas), cal veure si es satisfan les condicions suficients de segon ordre

$$y^T[\nabla^2_{xx}L(x,\lambda)]y > 0 \quad \forall y : \nabla h(x)y = 0$$

(x en aquest cas fa referència a (r, h)). La matriu hessiana, avaluada al punt $r = 1,\ h = 2,\ \lambda = 1/2$, és:

$$\nabla^2_{xx}L(r=1,h=2,\lambda=1/2) = \begin{pmatrix} \nabla^2_{rr}L(r,h,\lambda) & \nabla^2_{rh}L(r,h,\lambda) \\ \nabla^2_{hr}L(r,h,\lambda) & \nabla^2_{hh}L(r,h,\lambda) \end{pmatrix} =$$

$$= \begin{pmatrix} -2\pi h + 4\pi\lambda & -2\pi r + 2\lambda\pi \\ -2\pi r + 2\lambda\pi & 0 \end{pmatrix}_{(1,2,1/2)} = \begin{pmatrix} -2\pi & -\pi \\ -\pi & 0 \end{pmatrix}$$

Ara cal determinar els vectors y tals que $\nabla h(x)y = 0$. Tenim que

$$\nabla h(x) = \nabla_{(r,h)}[2\pi r h + 2\pi r^2 - 6\pi] = \begin{pmatrix} 8\pi & 2\pi \end{pmatrix}$$

Per tants els vectors y han de satisfer:

$$\begin{pmatrix} 8\pi & 2\pi \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 8\pi y_1 + 2\pi y_2 = 0 \Rightarrow y_2 = -4y_1$$

Finalment observem el signe de $y^T[\nabla^2_{xx}L(x,\lambda)]y$:

$$\begin{pmatrix} y_1 & -4y_1 \end{pmatrix}\begin{pmatrix} -2\pi & -\pi \\ -\pi & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ -4y_1 \end{pmatrix} = 6\pi y_1^2 > 0 \ \forall\ y \neq 0$$

Per tant es garanteix la condició suficient de segon ordre, el qual ens assegura que $r = 1, h = 2$ proporciona el volum màxim del cilindre (que és de $\pi r^2 h = 2\pi$).

iii) Per solucionar el problema anterior amb Lingo, podríem entrar el problema següent:

```
data:
     pi= 3.141592;
enddata
max= pi*r^2*h;
2*pi*r*h+2*pi*r^2=6*pi;
r>=0;
h>=0;
```

La solució proporcionada per Lingo és la següent:

```
Rows=      4 Vars=      2 No. integer vars=      0
```

```
Nonlinear rows=      2 Nonlinear vars=      2 Nonlinear constraints=    1
Nonzeros=      7 Constraint nonz=     4 Density=0.583

Optimal solution found at step:         8
Objective value:                 6.283184

                      Variable           Value        Reduced Cost
                            PI        3.141592        0.0000000E+00
                             R        1.000000        0.0000000E+00
                             H        2.000000        0.0000000E+00

                           Row    Slack or Surplus      Dual Price
                             1        6.283184            1.000000
                             2      0.0000000E+00       0.5000000
                             3        1.000000          0.1953993E-06
                             4        2.000000          0.0000000E+00
```

Podem comprovar com, efectivament, els valors de r i h obtinguts corresponen amb els que hem trobat, com el valor del volum màxim és de $2\pi = 6.283184$, i també com el multiplicador λ de la restricció d'igualtat (la 2 del llistat de Lingo) és de 0.5. També podem comprovar com els multiplicadors μ de les restriccions $r \geq 0$ i $h \geq 0$ (restriccions 3 i 4 al llistat) són 0 (una és de $0.1953993 \cdot 10^{-06}$, que és aproximadament 0), el qual concorda amb el que ens diuen les condicions d'optimalitat per a problemas amb restriccions de desigualtat. En definitiva, el punt proporcionat per Lingo verifica les condicions d'optimalitat del problema, tal i com ha de ser.

Solució del problema 61.

El problema de programació lineal pot ser escrit com:

$$\begin{aligned} \min \quad & f(x) = c^T x \\ \text{subj. a} & \\ & h(x) = Ax - b = 0 \\ & g(x) = -x \leq 0 \end{aligned}$$

Associant els multiplicadors λ a les restriccions $h(x) = 0$ i μ a les restriccions $g(x) \leq 0$, la funció lagrangiana que considerem és:

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

Per poder aplicar les condicions necessàries, s'ha de garantir que el punt x és regular. Això vol dir que $\nabla h(x)$ i $\nabla g(x)$ tenen les files linealment independents. Però com que $h(x) = Ax - b$ i $g(x) = -x$ directament tenim que $\nabla h(x) = A$, les files de la qual són linealment independents perquè l'enunciat del problema ens diu que A és de rang complet, i $\nabla g(x) = -$, que clarament també té les files linealment independents.

També cal usar un conjunt $\mathcal{A}(x) = \{j : g_j(x) = 0\}$ d'índexs de restriccions actives. En aquest cas el conjunt $\mathcal{A}(x)$ correspondrà a les variables j que verifiquen que $x_j = 0$. Com que sabem, per l'enunciat, que la solució bàsica és no degenerada, les variables bàsiques satisfan $x_B > 0$.

Per tant el conjunt $\mathcal{A}(x)$ correspon, de fet, al conjunt de variables no bàsiques.

Recordem que les condicions necessàries (de Kuhn-Tucker) que ha de satisfer un punt per ser òptim, són:

$$\begin{array}{rl} i) & \nabla_x L(x,\lambda,\mu) = 0 \\ ii) & \mu_i \geq 0 \;\forall i, \quad \mu_i = 0 \;\forall i \notin \mathcal{A}(x) \\ iii) & y^T[\nabla^2_{xx} L(x,\lambda,\mu)]y \geq 0 \quad \forall y : \nabla h(x) y = 0,\; \nabla g_j(x) y = 0 \;\forall i \in \mathcal{A}(x) \end{array}$$

La primera condició ens imposa que

$$c^T + \lambda^T A - \mu^T \Leftrightarrow \mu^T = c^T + \lambda^T A$$

La segona condició ens obliga a que $\mu_i = 0$ per aquells índexs i que no estan a $\mathcal{A}(x)$. Com que a $\mathcal{A}(x)$ només hi ha els índexs de les variables no bàsiques (com hem vist abans), el que realment estem dient és que $\mu_i = 0$ per a tot i associat a una variable bàsica. Si particionem el vector μ en μ_B i μ_N (associats respectivament a les variables bàsiques i no bàsiques), podem escriure la segona condició com:

$$\mu \geq 0, \quad \mu_B = 0$$

La darrera condició sempre es satisfarà, ja que $\nabla^2_{xx} L(x,\lambda,\mu) = \nabla_x [c^T + \lambda^T A - \mu^T] = 0$, i per tant sempre verificarem que, per a tot y

$$y^T[\nabla^2_{xx} L(x,\lambda,\mu)]y = 0 \geq 0$$

Ara només hem de jugar una mica amb el que ens diuen la primera i segona condició. Particionant la matriu de restriccions $A = [B\ N]$ en una part bàsica i una de no bàsica, considerant el particionament de μ abans introduït, i usant que $\mu \geq 0$ i $\mu_B = 0$, podem escriure la primera condició en dues parts:

$$\begin{array}{l} \mu_B^T = c_B^T + \lambda^T B = 0 \Rightarrow \lambda^T = -c_B^T B^{-1} \\ \mu_N^T = c_N^T + \lambda^T N \geq 0 \Rightarrow \mu_N^T = c_N^T - c_B^T B^{-1} N \geq 0 \end{array}$$

Adonem-nos que precisament la condició $\lambda^T = -c_B^T B^{-1}$ ens proporciona el valor negat de les variables duals, mentre que $\mu_N^T = c_N^T - c_B^T B^{-1} N \geq 0$ no és més que la condició de no-negativitat que han de satisfer els costos reduïts de les variables no bàsiques (els μ_N són, de fet, els costos reduïts de les variables x_N). Aquestes condicions coincideixen amb les que ha de garantir un solució bàsica òptima.

Solució del problema 62.

a) La funció objectiu és $f(x,y) = e^x y^4 - 100y^2(1+x)$. El seu gradient i Hessiana es calculen com:

$$\nabla f(x,y) = (\, e^x y^4 - 100y^2 \quad 4e^x y^3 - 200y(1+x)\,)$$

$$\nabla^2 f(x,y) = \begin{pmatrix} e^x y^4 & 4e^x y^3 - 200y \\ 4e^x y^3 - 200y & 12 e^x y^2 - 200(1+x) \end{pmatrix}$$

Pot comprovar-se com a tota la regió factible $\Omega = \{0 \leq x \leq 1,\ 0 \leq y\}$ la Hessiana no és definida positiva. Per exemple, al punt $x = 0$, $y = 1$, tenim

$$\nabla^2 f(0,1) = \begin{pmatrix} e^0 1^4 & 4e^0 1^3 - 200 \cdot 1 \\ 4e^0 1^3 - 200 \cdot 1 & 12e^0 1^2 - 200(1+0) \end{pmatrix} = \begin{pmatrix} 1 & -196 \\ -196 & -188 \end{pmatrix}.$$

Aquesta matriu no es definida positiva (ja que $|\nabla^2 f(0,1)| = -38604 < 0$), de forma que la funció objectiu no és convexa a tota la regió factible. Per tant no podem garantir que l'òptim que obtindríem amb un determinat algorisme d'optimització fos l'òptim global.

b) La direcció $d = (d_x, d_y)^T$ serà de descens al punt $(x, y) = (0, 1)$ si es verifica que

$$\nabla f(0,1) d < 0.$$

Només cal trobar $\nabla f(0,1)$, usant l'expressió obtinguda a l'apartat a)

$$\nabla f(0,1) = \begin{pmatrix} e^0 1^4 - 100 \cdot 1^2 & 4e^0 1^3 - 200 \cdot 1(1+0) \end{pmatrix} = \begin{pmatrix} -99 & -196 \end{pmatrix},$$

i comprovar que verifica la condició de descens:

$$\nabla f(x=0, y=1) d = \begin{pmatrix} -99 & -196 \end{pmatrix} \begin{pmatrix} 1 \\ -0.5 \end{pmatrix} = -1 < 0.$$

Adonem-nos, però, que la direcció proporcionada només és de descens en un entorn petit del punt $(0, 1)$. Així, per exemple, tenim que $f(0, 1) = -99$, que si ens belluguem una mica (usant una longitud de pas de $\alpha = 0.01$) en la direcció de d millorem la funció:

$$f((0,1) + 0.01(1, -0.5)) = f(0.01, 0.995) = -99.0025 < -99,$$

però que si ens belluguem una mica més (usant $\alpha = 0.02$) augmentem el seu valor:

$$f((0,1) + 0.02(1, -0.5)) = f(0.02, 0.99) = -98.9901 > -99,$$

c) Per trobar els candidats a ser òptim usarem les condicions d'optimalitat de primer ordre. En primer lloc definim la Lagrangiana del nostre problema:

$$L(x, y, \mu_1, \mu_2, \mu_3) = e^x y^4 - 100y^2(1+x) + \mu_1(x-1) + \mu_2(-x) + \mu_3(-y).$$

Per l'enunciat del problema sabem que a l'òptim la 2a i 3a restriccions són inactives, garantint aleshores que $\mu_2^* = \mu_3^* = 0$. Per tant escriurem la Lagrangiana de forma més simple:

$$L(x, y, \mu_1) = e^x y^4 - 100y^2(1+x) + \mu_1(x-1).$$

La primera condició necessària que ha de satisfer un punt óptim és:

$$\frac{\partial L(x,y,\mu)}{\partial x} = e^x y^4 - 100y^2 + \mu_1 := 0$$
$$\frac{\partial L(x,y,\mu)}{\partial y} = 4e^x y^3 - 200y(1+x) := 0.$$

La segona condició necessària ens diu que $\mu_1 \geq 0$, i que $\mu_1 = 0$ si la restricció $x - 1 \leq 0$ és inactiva a l'òptim. A priori no sabem si serà o no activa aquesta restricció. Considerem, doncs, les dues possibilitats:

i) La restricció és inactiva ($x-1<0$ i $\mu_1=0$).
El sistema que obtenim en aquest cas és:

$$\begin{aligned} e^x y^4 - 100y^2 &= 0 \\ 4e^x y^3 - 200y(1+x) &= 0. \end{aligned}$$

Pot comprovar-se com les solucions d'aquest sistema d'equacions són $(x, y=0)$ per una banda, i $(x=1, y=10/\sqrt{e})$ per una altra. Tanmateix, cap d'aquestes dues solucions són vàlides, ja que sabem per l'enunciat que $y>0$ (tercera restricció és inactiva), i hem suposat en aquest cas que $x<1$.

ii) La restricció és activa ($x=1$ i $\mu_1 \geq 0$).
El sistema que obtenim en aquest cas és:

$$\begin{aligned} ey^4 - 100y^2 + \mu_1 &= 0 \\ 4ey^3 - 400y &= 0. \end{aligned}$$

Pot comprovar-se com les solucions d'aquest sistema d'equacions són $(y=0, \mu_1=0)$, i $(y=10/\sqrt{e}), \mu_1=0)$. La primera d'elles, però, s'ha de descartar, ja que sabem per l'enunciat que $y>0$ (la tercera restricció és inactiva). Per tant, l'únic candidat a ser òptim que trobem ve donat per la segona solució, de forma que

$$(x^*, y^*, \mu_1^*) = (\,1 \quad 10/\sqrt{e} \quad 0\,)\,.$$

Tot i que l'enunciat no ho demana, pot comprovar-se com aquest punt satisfà les condicions de segon ordre (suficients i necessàries). Només cal veure que $\nabla^2_{(x,y),(x,y)} L(x,y,\mu_1)$ és definida positiva al punt $x^*=1$, $y^*=10/\sqrt{e}$:

$$\nabla^2_{(x,y),(x,y)} L = \begin{pmatrix} e^x y^4 & 4e^x y^3 - 200y \\ 4e^x y^3 - 200y & 12e^x y^2 - 200(1+x) \end{pmatrix}_{x=1, y=10/\sqrt{e}} = \begin{pmatrix} 3678.79 & 1213.06 \\ 1213.06 & 800.0 \end{pmatrix}.$$

També cal fer notar que el fet d'haver trobat que $\mu_1^*=0$ implica que la restricció $x-1 \leq 0$, tot i ser activa ($x^*=1$), no afecta a la determinació del punt òptim. Si eliminéssim aquesta restricció continuaríem obtenint la solució $x^*=1$, $y^*=10/\sqrt{e}$.

4.2 Optimització amb constriccions lineals

Solució del problema 63.

i) Al primer apartat hem de verificar que el punt $x^* = (3\ 0\ 2\ 5)^T$ i els multiplicadors $\lambda^* = (-1\ 0)^T$ i $\mu^* = (0\ 1\ 0\ 0)^T$ donats verifiquen les condicions d'optimalitat. En primer lloc, fàcilment observem que el punt x^* proporcionat és factible, ja que totes les components són no negatives, i satisfà les dues restriccions lineals ($3+0+2=5$, $2\cdot 3+0-5=1$). També podem comprovar com es satisfan les condicions necessàries sobre μ: $\mu^* \geq 0$, i per a les variables que no tenen el seu límit actiu (és a dir, aquelles que $x_i^*>0$, que són x_1, x_3 i x_4) es té que la μ_i^* associada és igual a 0 (en aquest cas $\mu_1=\mu_3=\mu_4=0$). També podem observar com es satisfà una de les condicions suficients: $\mu_i^*>0$ per a tota variable que té el límit actiu (en aquest cas només és x_2, i comprovem com $\mu_2^* = 1 > 0$).

Per comprovar la resta de condicions usarem la funció lagrangiana:

$$L(x,\lambda,\mu) = x_1^2+x_2^2+x_3^2-x_1x_3-3x_1+2x_2+\lambda_1(x_1+x_2+x_3-5)+\lambda_2(2x_1+x_2-x_4-1)-\sum_{i=1}^{4}\mu_i x_i$$

S'ha de garantir que $\nabla_x L(x^*,\lambda^*,\mu^*) = 0$. En aquest cas tenim:

$$\begin{aligned}
\nabla_{x_1} L(x^*,\lambda^*,\mu^*) &= 2x_1^* - x_3^* - 3 + \lambda_1^* + 2\lambda_2^* - \mu_1^* = 2\cdot 3 - 2 - 3 - 1 + 2\cdot 0 - 0 = 0\\
\nabla_{x_2} L(x^*,\lambda^*,\mu^*) &= 2x_2^* + 2 + \lambda_1^* + \lambda_2^* - \mu_2^* = 2\cdot 0 + 2 - 1 + 0 - 1 = 0\\
\nabla_{x_3} L(x^*,\lambda^*,\mu^*) &= 2x_3^* - x_1^* + \lambda_1^* - \mu_3^* = 2\cdot 2 - 3 - 1 - 0 = 0\\
\nabla_{x_4} L(x^*,\lambda^*,\mu^*) &= -\lambda_2^* - \mu_4^* = -0 - 0 = 0
\end{aligned}$$

Per tant, els valors de x^*, λ^* i μ^* proporcionats satisfan les equacions anteriors.

Ara cal comprovar les condicions de segon ordre. En primer lloc trobem $\nabla_{xx}L(x,\lambda,\mu)$:

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

Aquesta matriu és semidefinida positiva (els valors propis són 0, 1, 2 i 3). Per tant, es satisfà que $\forall y\ y^T\nabla_{xx}L(x,\lambda,\mu)y \geq 0$. Això ens indica que es satisfà la condició necessària de segon ordre, però no la suficient. Per garantir la condició suficient (la última de les condicions suficients que ens queda, amb el qual podrem assegurar que x^* és un mínim local) s'ha de verificar que

$$y^T\nabla_{xx}L(x,\lambda,\mu)y > 0 \quad \forall y : \nabla h(x^*)y = 0, \quad \nabla g_j(x^*)y = 0\ \forall j : x_j^* = 0$$

on $h(x)$ i $g_j(x)$ fan referència a les restriccions d'igualtat i desigualtat respectivament. En primer lloc busquem les y afectades per les condicions anteriors:

$$\left.\begin{array}{l}
\nabla h(x^*)y = 0 \iff \begin{pmatrix} 1 & 1 & 1 & 0\\ 2 & 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_3\\ y_4 \end{pmatrix} = 0\\[2ex]
\nabla g_2(x^*)y = 0 \iff \begin{pmatrix} 0 & -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_3\\ y_4 \end{pmatrix} = 0
\end{array}\right\} \Rightarrow y = y_1\begin{pmatrix} 1\\ 0\\ -1\\ 2 \end{pmatrix}$$

I ara ja podem calcular:

$$y^T\nabla_{xx}L(x,\lambda,\mu)y = 6y_1^2 > 0\ \forall y \neq 0$$

Per tant, es satisfà la darrera condició suficient que ens quedava, amb el qual podem assegurar que el punt x^* donat és un mínim local del nostre problema (en aquest cas, a més, és mínim global).

ii) Partint del punt $x^0 = (0\ 2\ 3\ 1)^T$, realitzarem ara dues iteracions del mètode del gradient reduït. Abans, però, cal saber quin és el gradient de la funció objectiu:

$$\nabla f(x) = (\, 2x_1 - x_3 - 3 \quad 2x_2 + 2 \quad 2x_3 - x_1 \quad 0\,)$$

Ens cal també determinar quines variables seran considerades dependents (o bàsiques) i independents (o no bàsiques). La única restricció és que una variable bàsica no pot valer 0. Per comoditat, donat que ens serà molt fàcil calcular B^{-1}, considerarem com a bàsiques x_3 i x_4 (x_1 i x_2 seran no bàsiques). En aquest cas tenim que

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad\qquad N = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$$

Ara estem en disposició de realitzar les dues iteracions requerides.

1a. iteracio)

Al punt x^0 la funció objectiu val $f(x^0) = 17$. Per la seva banda, el gradient de $f(x)$ al punt actual i el valor de $\rho = \nabla_{x_N} f(x^0) - \nabla_{x_B} f(x^0) B^{-1} N$ són:

$$\nabla f(x^0) = (\, -6 \quad 6 \quad 6 \quad 0\,) \qquad \rho = (\, -6 \quad 6\,) - (\, 6 \quad 0\,)\, B^{-1} N = (\, -12 \quad 0\,)$$

La direcció de moviment de les variables no bàsiques d_{x_N} es troba a partir del vector ρ anterior de forma:

$$d_{x_{N_i}} = \begin{cases} -\rho_i & \text{si } \rho_i < 0 \text{ o } x_{N_i} > 0 \\ 0 & \text{altrament} \end{cases}$$

Per la seva banda, la direcció de moviment de les variables bàsiques es calcula com $d_{x_B} = -B^{-1} N d_{x_N}$, garantint d'aquesta forma que el nou punt verificarà les restriccions d'igualtat. Realitzant els càlculs al punt actual, tenim que:

$$\left.\begin{array}{l} d_{x_N} = (\, 12 \quad 0\,)^T \\ d_{x_B} = -B^{-1} N\, (\, 12 \quad 0\,)^T = (\, -12 \quad 24\,)^T \end{array}\right\} \Rightarrow d_x = \begin{pmatrix} 12 \\ 0 \\ -12 \\ 24 \end{pmatrix}$$

Ara hem de calcular la longitud de pas α. En primer lloc trobem els valors α_1 i α_2 corresponents a les longituds de pas màximes per garantir que les variables bàsiques i no bàsiques continuen essent no negatives. En aquest cas tenim:

$$\alpha_1 = \max\{\alpha : \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} -12 \\ 24 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}\} \Rightarrow \alpha_1 = 1/4$$
$$\alpha_2 = \max\{\alpha : \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} 12 \\ 0 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \geq 0\} \Rightarrow \alpha_2 = +\infty$$

Ara realitzem una cerca lineal partint del punt x^0 i usant la direcció d_x abans calculada, sabent que la longitud de pas ha de ser menor que el $\min\{\alpha_1, \alpha_2\}$, que en aquest cas és $\alpha_1 = 1/4$. Per tant ara hem de calcular:

$$\alpha_3 = \arg\min\{g(\alpha) = f(x^0 + \alpha dx) = f(\begin{pmatrix} 12\alpha \\ 2 \\ 3 - 12\alpha \\ 1 + 24\alpha \end{pmatrix}) = 432\alpha^2 - 144\alpha + 17, 0 \leq \alpha \leq 1/4\}$$

Igualant a 0 la derivada de $g(\alpha)$, obtenim que $144(6\alpha - 1) = 0$, amb el qual el mínim de la funció anterior és $\alpha = 1/6$. Com que aquest valor és menor que $1/4$, tenim directament que és la solució de la cerca lineal. Per tant $\alpha_3 = 1/6$.

Ara podem calcular el nou punt

$$x^1 = x^0 + \alpha_3 d_x = \begin{pmatrix} 0 \\ 2 \\ 3 \\ 1 \end{pmatrix} + 1/6 \begin{pmatrix} 12 \\ 0 \\ -12 \\ 24 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 5 \end{pmatrix}$$

que té un valor de funció objectiu de $f(x^1) = 5$ (hem passat d'un punt x^0 que tenia un cost de 17 a un on la funció objectiu val 5). Com que el valor $\alpha_3 \neq \alpha_1$ cap variable bàsica ha esdevingut 0, amb el qual no cal canviar la partició de variables bàsiques i no bàsiques.

2a. iteracio)

A la segona iteració cal fer els mateixos passos que a la primera, però usant ara el nou punt x^1 (no entrarem en tants detalls, doncs). El gradient i direcció de descens al punt actual són:

$$\nabla f(x^1) = (\,0 \quad 6 \quad 0 \quad 0\,) \qquad \rho = (\,0 \quad 6\,) \qquad d_{x_N} = \begin{pmatrix} 0 \\ -6 \end{pmatrix} \qquad d_{x_B} = \begin{pmatrix} 6 \\ -6 \end{pmatrix}$$

Els valors α_i en aquesta segona iteració corresponen a:

$$\begin{aligned}
\alpha_1 &= \max\{\alpha : \begin{pmatrix} 1 \\ 5 \end{pmatrix} + \alpha \begin{pmatrix} 6 \\ -6 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}\} \Rightarrow \alpha_1 = 5/6 \\
\alpha_2 &= \max\{\alpha : \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ -6 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \geq 0\} \Rightarrow \alpha_2 = 1/3 \\
\alpha_3 &= \arg\min\{g(\alpha) = f(x^1 + \alpha dx) = 72\alpha^2 - 36\alpha + 5, 0 \leq \alpha \leq \min\{5/6, 1/3\} = 1/3\}
\end{aligned}$$

El mínim de $g(\alpha)$ correspon a $1/4$, que és menor que $1/3$. Per tant $\alpha_3 = 1/4$, amb el qual el nou punt x^2 és:

$$x^2 = x^1 + \alpha_3 d_x = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 5 \end{pmatrix} + 1/4 \begin{pmatrix} 0 \\ -6 \\ 6 \\ -6 \end{pmatrix} = \begin{pmatrix} 2 \\ 0.5 \\ 2.5 \\ 3.5 \end{pmatrix}$$

i $f(x^2) = 1/2 < 5 = f(x^1)$, amb el qual hem millorat el valor de la funció objectiu. Com que $\alpha_3 \neq \alpha_1$ no cal modificar la partició actual de variables.

iii) En aquest tercer apartat iterarem a partir del punt $x^0 = (1\ 1\ 3\ 2)^T$, que té un cost associat de $f(x^0) = 7$. Tal i com ens diu l'enunciat, considerarem la partició següent: x_1 i x_2 són bàsiques, i les altres dues variables seran les no bàsiques. Per tant

$$B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \Rightarrow B^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} \qquad\qquad N = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Tenint en compte les matrius anteriors, el gradient i direcció de descens al punt x^0 considerat són:

$$\nabla f(x^0) = (\,-4 \quad 4 \quad 5 \quad 0\,) \qquad \rho = (\,-7 \quad -8\,) \qquad d_{x_N} = \begin{pmatrix} 7 \\ 8 \end{pmatrix} \qquad d_{x_B} = \begin{pmatrix} 15 \\ -22 \end{pmatrix}$$

Els valors α_i en aquest punt són:

$$\alpha_1 = \max\{\alpha : \begin{pmatrix}1\\1\end{pmatrix} + \alpha\begin{pmatrix}15\\-22\end{pmatrix} \geq \begin{pmatrix}0\\0\end{pmatrix}\} \Rightarrow \alpha_1 = 1/22 = 0.0454$$
$$\alpha_2 = \max\{\alpha : \begin{pmatrix}3\\2\end{pmatrix} + \alpha\begin{pmatrix}7\\8\end{pmatrix} \geq \begin{pmatrix}0\\0\end{pmatrix} \geq 0\} \Rightarrow \alpha_2 = +\infty$$
$$\alpha_3 = \arg\min\{g(\alpha) = f(x^0 + \alpha dx) = 653\alpha^2 - 113\alpha + 7, 0 \leq \alpha \leq \min\{1/22, +\infty\} = 1/22\}$$

El mínim de $g(\alpha)$ correspon a 0.0865. Aquest valor, però, és més gran que 1/22= 0.0454. Per tant hem de considerar que $\alpha_3 = \alpha_1 = 1/22$ (això degut a que $g(\alpha)$ és convexa —$g''(\alpha) > 0$—, ja que si fos concava prendríem $\alpha_3 = 0$). El nou punt x^1 serà doncs:

$$x^1 = x^0 + \alpha_3 d_x = \begin{pmatrix}1\\1\\3\\2\end{pmatrix} + 1/22\begin{pmatrix}15\\-22\\7\\8\end{pmatrix} = \begin{pmatrix}1.6818\\0\\3.3182\\2.3636\end{pmatrix}$$

i $f(x^1) = 3.2128 < 7 = f(x^0)$, amb el qual hem millorat el valor de la funció objectiu. Observem com, en aquest cas, donat que $\alpha_3 = \alpha_1$, una de les variables bàsiques s'ha fet 0, situació que s'ha d'evitar al mètode del gradient reduït. Per tant, ens cal buscar una variable no bàsica amb valor positiu i intercanviar-la per aquesta variable bàsica anul·lada. Per exemple, podem fer entrar a la base la variable x_3, amb el qual les noves matrius B i N serien:

$$B = \begin{pmatrix}1 & 1\\2 & 0\end{pmatrix} \qquad N = \begin{pmatrix}1 & 0\\1 & -1\end{pmatrix}$$

i haurien de ser usades a la següent iteració per calcular el nou punt x^2.

Solució del problema 64.

a) Denotant per x_1 l'arc que va del node A al C, per x_2 l'arc que va del node A al B, i per x_3 i x_4 els arcs que van del node B al C, el problema a solucionar pot ser formulat com

$$\begin{array}{lll}
\min & \displaystyle\sum_{i=1}^{4} x_i^2 & \\
\text{subj. a} & x_1 + x_2 = M & \text{[equació de balanç al node A]} \\
 & x_2 = x_3 + x_4 & \text{[equació de balanç al node B]} \\
 & x_1 + x_3 + x_4 = M & \text{[equació de balanç al node C]} \\
 & 0 \leq x_i \leq M \quad i = 1, ..., 4 &
\end{array}$$

b) Podem eliminar la restricció associada al balanç en un node (per exemple, el C), ja que és combinació lineal de les altres dues. És fàcil veure que sumant les equacions dels nodes A i B obtenim la del node C:

$$\left.\begin{array}{r} x_1 + x_2 = M \\ x_3 + x_4 = x_2 \end{array}\right\} \Rightarrow x_1 + x_3 + x_4 = M$$

Pel que fa als límits de les variables, podem eliminar el límit superior $x_i \leq M$ ja que no poden circular per la xarxa més de M unitats del fluid. Si eliminem els límits inferiors $x_i \geq 0$, però, podríem tenir situacions on apareguessin fluxos negatius que satisfan les restriccions d'igualtat del problema. Per exemple, el punt $x_1 = -M$, $x_2 = 2M$ i $x_3 = x_4 = M$ és factible si eliminem $x_i \geq 0$. Tanmateix, en aquest problema concret aquestes situacions són econòmicament indesitjables ja que tenen un cost molt elevat. És millor fer circular fluxos que siguin fraccions positives de M, les quals satisfaran les restriccions d'igualtat amb un cost menor.

El problema que tenim eliminant els límits i l'equació de balanç del node C és doncs:

$$\begin{array}{rl} \min & \sum_{i=1}^{4} x_i^2 \\ \text{subj. a} & x_1 + x_2 = M \\ & x_2 = x_3 + x_4 \end{array}$$

c) El problema obtingut a l'apartat anterior el solucionarem aplicant les condicions d'optimalitat (de Lagrange) per a problemes amb restriccions d'igualtat. Definint la Lagrangiana

$$L(x,\lambda) = \sum_{i=1}^{4} x_i^2 + \lambda_1(x_1 + x_2 - M) + \lambda_2(x_2 - x_3 - x_4),$$

les condicions de primer ordre s'obtenen derivant i igualant a 0 el gradient de la Lagrangiana:

$$\left.\begin{array}{l} \dfrac{\partial L(x,\lambda)}{\partial x_1} = 2x_1 + \lambda_1 := 0 \\ \dfrac{\partial L(x,\lambda)}{\partial x_2} = 2x_2 + \lambda_1 + \lambda_2 := 0 \\ \dfrac{\partial L(x,\lambda)}{\partial x_3} = 2x_3 - \lambda_2 := 0 \\ \dfrac{\partial L(x,\lambda)}{\partial x_4} = 2x_4 - \lambda_2 := 0 \\ \dfrac{\partial L(x,\lambda)}{\partial \lambda_1} = x_1 + x_2 - M := 0 \\ \dfrac{\partial L(x,\lambda)}{\partial \lambda_2} = x_2 - x_3 - x_4 := 0 \end{array}\right\} \Rightarrow \quad \begin{array}{l} x_1^* = 3/5M \\ x_2^* = 2/5M \\ x_3^* = x_4^* = 1/5M. \end{array}$$

Per estar segurs de que es tracta d'un punt mínim hem de comprovar les condicions de 2n ordre: $y^T[\nabla^2_{xx}L(x^*,\lambda^*)]y \geq 0$ per a tota $y \neq 0$ tal que $Ay = 0$ (on A representa la matriu de restriccions lineals del problema). Calculem $\nabla^2_{xx}L(x^*,\lambda^*)$:

$$\nabla^2_{xx}L(x^*,\lambda^*) = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}.$$

Donat que $\nabla^2_{xx}L(x^*,\lambda^*)$ és definida positiva per a tot vector y (no només per aquells tals que $Ay = 0$), es satisfan les condicions suficients de segon ordre, i per tant el punt $x^* =$

$(\,3/5M \quad 2/5M \quad 1/5M \quad 1/5M\,)^T$ és el mínim del problema.

d) Per garantir que és un mínim global hem de veure si ens trobem davant d'un problema convex, és a dir, que té la funció objectiu i una regió factible convexes:

i) La Hessiana de la funció objectiu ve donada per

$$\nabla^2_{xx} f(x) = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix},$$

que és una matriu definida positiva, garantint que $f(x)$ és estrictament convexa.

ii) Les dues equacions del problema són lineals, i sabem que aquestes sempre defineixen conjunts factibles convexos

Per tant, podem garantir que el punt obtingut a l'apartat c) és l'òptim global del nostre problema.

e) El problema a solucionar és:

$$\begin{array}{ll} \min & \sum_{i=1}^{4} x_i^2 \\ \text{subj. a} & x_1 + x_2 = M \\ & x_2 - x_3 - x_4 = 0 \\ & x_i \geq 0 \quad i = 1, ..., 4 \end{array} \qquad \text{on} \quad A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 \end{pmatrix} \qquad b = \begin{pmatrix} 10 \\ 0 \end{pmatrix}.$$

Per l'enunciat sabem que la base inicial està formada per x_1 i x_2, de forma que

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad B^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \qquad\qquad N = \begin{pmatrix} 0 & 0 \\ -1 & -1 \end{pmatrix}.$$

El gradient de la funció objectiu és

$$\nabla f(x) = (\,2x_1 \quad 2x_2 \quad 2x_3 \quad 2x_4\,)$$

1a Iteració)

El punt inicial d'iteració i gradient en aquest punt són:

$$x^0 = (\,10 \quad 0 \quad 0 \quad 0\,)^T \qquad\qquad \nabla f(x^0) = (\,20 \quad 0 \quad 0 \quad 0\,).$$

Usant la partició anterior de $A = [B|N]$ i una partició equivalent per a $\nabla f(x)$, calculem el gradient reduït al punt x^0:

$$\rho = \nabla_{x_N} f(x^0) - \nabla_{x_B} f(x^0) B^{-1} N = (\,-20 \quad -20\,).$$

Per tant la direcció de moviment obtinguda és

$$\left.\begin{array}{l} d_{x_N} = \begin{pmatrix} 20 \\ 20 \end{pmatrix} \\ d_{x_B} = -B^{-1}Nd_{x_N} = \begin{pmatrix} -40 \\ 40 \end{pmatrix} \end{array}\right\} \Rightarrow \quad d_x = \begin{pmatrix} -40 \\ 40 \\ 20 \\ 20 \end{pmatrix}.$$

Tot seguit calculem la passa α:

$$\begin{aligned} \alpha_1 &= \max\{\alpha \geq 0 : \begin{pmatrix} 10 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -40 \\ 40 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}\} \Rightarrow \alpha_1 = 1/4 \\ \alpha_2 &= \max\{\alpha \geq 0 : \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 20 \\ 20 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \geq 0\} \Rightarrow \alpha_2 = +\infty \\ \alpha_3 &= \arg\min\{g(\alpha) = f(x^0 + \alpha dx) = 4000\alpha^2 - 800\alpha + 100, 0 \leq \alpha \leq 1/4\}. \end{aligned}$$

El mínim de $g(\alpha)$ ve donat per $\alpha = 1/10$. Com que aquest valor ja és menor que $1/4$, tenim que $\alpha_3 = 1/10$.

El nou punt x^1 és ara:

$$x^1 = x^0 + \alpha_3 d_x = \begin{pmatrix} 10 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 1/10 \begin{pmatrix} -40 \\ 40 \\ 20 \\ 20 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 2 \\ 2 \end{pmatrix}.$$

Al nou punt hem millorat el valor de la funció objectiu:

$$f(x^0) = 10^2 = 100 > f(x^1) = 6^2 + 4^2 + 22^2 = 60.$$

2a Iteració)

La partició de variables bàsiques i no bàsiques no ha estat modificada. El gradient al nou punt és

$$\nabla f(x^1) = (\,12 \quad 8 \quad 4 \quad 4\,),$$

i el gradient reduït ve donat per

$$\rho = \nabla_{x_N} f(x^1) - \nabla_{x_B} f(x^1) B^{-1} N = (\,0 \quad 0\,).$$

Per tant $d_{x_N} = 0$ i $d_{x_B} = 0$. És a dir, ja no podem bellugar-nos millorant la nostra funció objectiu, de forma que el punt actual $x^1 = (\,6 \quad 4 \quad 2 \quad 2\,)$ ja és l'òptim del nostre problema. Adonem-nos que aquest resultat concorda amb l'obtingut de forma analítica a l'apartat b):

$$x^* = \begin{pmatrix} 3/5M \\ 2/5M \\ 1/5M \\ 1/5M \end{pmatrix} \quad \overset{[M = 10]}{\Longrightarrow} \quad \begin{pmatrix} 6 \\ 4 \\ 2 \\ 2 \end{pmatrix}.$$

Solució del problema 73.

Hem de començar realitzant la selecció de les variables dependents i independents sobre x. Es calcula prèviament la matriu Jacobiana $\nabla h(x)$:

$$h(x) = \begin{bmatrix} \frac{x_1^2}{2} + x_2 - x_3 + 2x_4x_5 - x_5 + 2 \\ \frac{3x_1^2}{2} + x_2x_3 - 2x_2 + \frac{3x_4^2}{2} + x_5 - \frac{7}{2} \end{bmatrix}$$
$$\nabla h(x) = \begin{bmatrix} x_1 & 1 & -1 & 2x_5 & 2x_4 - 1 \\ 3x_1 & x_3 - 2 & x_2 & 3x_4 & 1 \end{bmatrix}$$

que, avaluada sobre $x = [\,1 \quad 0 \quad 3 \quad 1 \quad 1/2\,]'$ proporciona:

$$\nabla h(x) = \begin{bmatrix} 1 & 1 & -1 & 1 & 1 \\ 3 & 1 & 0 & 3 & 1 \end{bmatrix}$$

La variable x_2 no hauria de ser independent, doncs es troba a fita inferior. Una possible selecció de variables dependents seria $y = [\,x_1 \quad x_3\,]'$ ja que ambdues es troben entre fites i la matriu:

$$\nabla_y h(x) = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix}$$

és no singular. Les variables independients serien $z = [\,x_2 \quad x_4 \quad x_5\,]'$, y la matriu $\nabla_z h(x)$:

$$\nabla_z h(x) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 1 \end{bmatrix}$$

Si calculem l'expressió del gradiente:

$$\nabla f(x) = \begin{bmatrix} \frac{x_3^2}{2} - x_5 + 3 & -6 & \frac{2x_1x_3}{2} - 4 & -4 & -x_1 \end{bmatrix}$$

i l'avaluem sobre $x = [\,1 \quad 0 \quad 3 \quad 1 \quad 1/2\,]'$ s'obté:

$$\nabla f(x) = [\,7 - 6 - 1 - 4 - 1\,] \quad ; \quad \nabla_y f(x) = [\,7 \quad -1\,] \quad ; \quad \nabla_z f(x) = [\,-6 \quad -4 \quad -1\,]$$

Determinem el gradient reduït sobre el punt x. A la pràctica, el càlcul del gradient reduït es duu a terme en dues passes. A la primera es calcula el producte $\lambda' = \nabla_y f(y,z)[\nabla_y h(y,z)]^{-1}$ mitjançant la resolució del sistema d'equacions $\nabla_y h(y,z)'\lambda = \nabla_y f(y,z)'$. La raó d'aquest procediment és que és més estable i eficient treballar amb la factorització LU de la matriu $\nabla_y h(y,z)$ que amb la seva inversa. La factorització LU de $\nabla_y h(x)$ és:

$$\nabla_y h(x) = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}$$

Resolent el sistema s'obté el valor de λ:

$$\begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix} \lambda = \begin{bmatrix} 1 & \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = - \begin{bmatrix} 7 \\ -1 \end{bmatrix} \Rightarrow \begin{cases} \lambda_1 = -1 \\ \lambda_2 = -2 \end{cases}$$

A continuació es procedeix al càlcul de $r = \nabla_z f(y,z)' - \nabla_z h(y,z)'\lambda$:

$$r = \begin{bmatrix} -6 \\ -4 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ -11 \\ -4 \end{bmatrix}$$

Tenint en compte que $z = [\,0 \quad 1 \quad 1/2\,]$ i les fites $l_z = 0$ i $u_z = [\,4 \quad 4 \quad 4\,]$, s'obté $\Delta z = -r$:

$$\Delta z = \begin{bmatrix} 9 \\ 11 \\ 4 \end{bmatrix}$$

Un cop determinada Δz (i després de comprobar que és diferent de zero), es procedeix al càlcul de Δy. En la pràctica, aquest càlcul es fa a través de la resolució del sistema d'equacions $\nabla_y h(y,z)\Delta y = -\nabla_z h(y,z)\Delta z$:

$$\begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix} \Delta y = \begin{bmatrix} 1 & \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ & 3 \end{bmatrix} \begin{bmatrix} \Delta y_1 \\ \Delta y_2 \end{bmatrix} = - \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 9 \\ 11 \\ 4 \end{bmatrix} = \begin{bmatrix} -24 \\ -46 \end{bmatrix} \Rightarrow \begin{cases} \Delta y_1 = -46/3 \\ \Delta y_2 = 26/3 \end{cases}$$

$$\Delta y = \begin{bmatrix} -46/3 \\ 26/3 \end{bmatrix}$$

Calculem ara la longitud de pas màxima $\bar{\alpha}$. Tenint en compte que $l = 0$ i $u = [\,4 \quad 4 \quad 4 \quad 4 \quad 4\,]'$ tenim que:

$$y = \begin{bmatrix} 1 \\ 3 \end{bmatrix}', \ \Delta y = \begin{bmatrix} -46/3 \\ 26/3 \end{bmatrix} \quad ; \quad \left.\begin{array}{l} \bar{\alpha}_{y_1} = \dfrac{1}{46/3} \\ \bar{\alpha}_{y_2} = \dfrac{(4-3)}{26/3} \end{array}\right\} \Rightarrow \bar{\alpha}_y = \min\{\frac{3}{46}, \frac{3}{26}\} = \frac{3}{46}$$

$$z = \begin{bmatrix} 0 \\ 1 \\ 1/2 \end{bmatrix}', \ \Delta z = \begin{bmatrix} 9 \\ 11 \\ 4 \end{bmatrix} \quad ; \quad \left.\begin{array}{l} \bar{\alpha}_{z_1} = (4-0)/9 = 4/9 \\ \bar{\alpha}_{z_2} = (4-1)/11 = 3/11 \\ \bar{\alpha}_{z_3} = (4-1/2)/4 = 7/8 \end{array}\right\} \Rightarrow \bar{\alpha}_z = \min\{\frac{4}{9}, \frac{3}{11}, \frac{7}{8}\} = \frac{3}{11}$$

$$\bar{\alpha} = \min\{\frac{3}{46}, 3/11\} = \frac{3}{46}$$

Prenem ara una longitud de pas arbitrària $\alpha^* = \bar{\alpha}/2 = 3/92$. El punt iterat serà:

$$\tilde{x} = x + \alpha^* \Delta x = \begin{bmatrix} 1 \\ 0 \\ 3 \\ 1 \\ 1/2 \end{bmatrix} + \frac{3}{92} \begin{bmatrix} -46/3 \\ 9 \\ 26/3 \\ 11 \\ 4 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 27/92 \\ 151/46 \\ 125/92 \\ 29/46 \end{bmatrix} ; \ \tilde{y} = \begin{bmatrix} 1/2 \\ 151/46 \end{bmatrix}, \ \tilde{z} = \begin{bmatrix} 27/92 \\ 125/92 \\ 29/46 \end{bmatrix}$$

Com era d'esperar, el punt $\tilde{x}$ no és factible, ja que $h(\tilde{x}) = \begin{bmatrix} 1.98488 \\ 3.26364 \end{bmatrix} \neq 0$. Apliquem una passa del procés iteratiu de recuperació de factibilitat a partir de $\tilde{x}$. El càlcul de $\Delta \tilde{y} =$

$-[\nabla_y h(y,z)]^{-1} h(\tilde{y},\tilde{z})$ es realitza resolent el sistema $\nabla_y h(y,z)\Delta\tilde{y} = -h(\tilde{y},\tilde{z})$:

$$\begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix} \Delta\tilde{y} = \begin{bmatrix} 1 & \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ & 3 \end{bmatrix} \begin{bmatrix} \Delta\tilde{y}_1 \\ \Delta\tilde{y}_2 \end{bmatrix} = -h(\tilde{y},\tilde{z}) = \begin{bmatrix} -1.98488 \\ -3.26364 \end{bmatrix} \Rightarrow \begin{cases} \Delta\tilde{y}_1 = -1.08788 \\ \Delta\tilde{y}_2 = 0.897 \end{cases}$$

Tenint en compte el valor de les variables $\tilde{y}' = [\,1/2 \quad 151/46\,]'$ i les fites $u_y = [\,4 \quad 4\,]$, $l_y = 0$, la longitud de pas màxima al llarg de $\Delta\tilde{y}$ és:

$$\left.\begin{aligned} \bar{\alpha}_{\tilde{y}_1} &= \frac{1/2}{1.08788} = 0.4596 \\ \bar{\alpha}_{\tilde{y}_2} &= \frac{4 - 151/46}{0.897} = 0.7997 \end{aligned}\right\} \Rightarrow \bar{\alpha}_{\tilde{y}} = 0.4596 \quad ; \quad \alpha_{\tilde{y}} = \min\{1, 0.4596\} = 0.4596$$

és a dir:

$$\tilde{y} \leftarrow \tilde{y} + \alpha_{\tilde{y}}\Delta\tilde{y} \approx \begin{bmatrix} 0. \\ 3.6948 \end{bmatrix} \quad ; \quad h(\tilde{y},\tilde{z}) = \begin{bmatrix} -0.3186 \\ 0.3969 \end{bmatrix}$$

El nou punt $\tilde{x}$ està més a prop de la hipersuperfície $\mathcal{S}$ de les constriccions $h(x) = 0$ que el punt original. Com que $\alpha_{\tilde{y}} = \bar{\alpha}_{\tilde{y}_1}$, hem d'intercanviar $\tilde{y}_1$ amb una variable independent $\tilde{z}_q$ apropiada. El valor de $\nabla h(x)$ sobre el nou punt $\tilde{x} = [\tilde{y}', \tilde{z}']$ és:

$$\nabla h(\tilde{x}) = \begin{bmatrix} 1/2 & 1 & -1 & 1.2608 \\ 3/2 & 1.2826 & 0.2934 & 4.0760 \end{bmatrix}$$

La variable $\tilde{z}_1 \equiv x_2$ està entre fites i, junt amb $\tilde{y}_2$, proporciona un nou $\nabla_y h(y,z)$ no singular:

$$\nabla_y h(y,z) = \begin{bmatrix} 1 & -1 \\ 1.2826 & 0.2934 \end{bmatrix} \quad , \quad \det(\nabla_y h(y,z)) = 1.57608$$

així doncs, es por intercanviar $\tilde{y}_1 \equiv x_1$ per $\tilde{z}_1 \equiv x_2$.

Solució del problema 65.

DETERMINACIÓ DE LES VARIABLES BÀSIQUES, SUPERBÀSIQUES I NO BÀSIQUES

Les variables bàsiques $\underline{0} \leq X_B \leq \overline{X}_B$, a ser possible estrictament entre fites, han de tenir imprescindiblement columnes d'A linealment independents. Les variables superbàsiques $\underline{0} < X_S < \overline{X}_S$ han d'estar estrictament entre fites. Les variables no bàsiques X_N han de tenir les seves components o a la seva fita superior: $x_{Nj} = \overline{x}_j$, $j \in \overline{I}$ o a la seva fita inferior: $x_{Nj} = 0$, $j \in \underline{I}$, sent $\overline{I}$ i $\underline{I}$ els conjunt d'índexos de variables no bàsiques a fita superior i a fita inferior respectivament.

Per triar les variables bàsiques tenim que:

- x_2 està a la fita inferior i x_5 i x_7 estan a la superior (per tant, en principi, les excloem del conjunt de les bàsiques)
- x_1, x_3 i x_4 tenen les columnes linealment dependents (no poden ser simultàniament bàsiques)
- x_1, x_3 i x_6 poden ser les variables bàsiques X_B perquè tenen les columnes linealment independents:

$$B = \begin{bmatrix} 3 & 0 & 0 \\ -6 & 1 & 4 \\ 3 & 0 & -3 \end{bmatrix} = \begin{bmatrix} 3 & & \\ -6 & 1 & \\ 3 & 0 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 4 \\ & & 1 \end{bmatrix}$$

- donat que hem trobat variables bàsiques estrictament entre fites, totes les variables que estan a una fita seran no bàsiques: $X_N = \{x_2, x_5, x_7\}$, i la resta de variables (estrictament entre fites) seran les superbàsiques $X_S = \{x_4, x_8\}$. (Altres combinacions de bàsiques i superbàsiques són també possibles).

MATRIUS S, N I $\widehat{A}$ I CÀLCUL (NO NECESSARI) DE Z

S és doncs $S = \begin{bmatrix} -3 & 3 \\ 8 & -3 \\ -3 & 0 \end{bmatrix}$ i $N = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 6 & 0 \\ -1 & 2 & 1 \end{bmatrix}$ d'on la matriu de constriccions actives

$$\widehat{A} \text{ és:} \quad \widehat{A} = \left[\begin{array}{c|c|c} B & S & N \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & \widetilde{\mathit{1}} \end{array}\right] = \begin{bmatrix} 3 & 0 & 0 & -3 & 3 & 1 & 3 & 1 \\ -6 & 1 & 4 & 8 & -3 & 1 & 6 & 0 \\ 3 & 0 & -3 & -3 & 0 & -1 & 2 & 1 \\ & & & & & 1 & & \\ & & \mathbf{0} & & & & -1 & \\ & & & & & & & -1 \end{bmatrix}$$

La matriu Z, la qual **no és necessari** tenir emmagatzemada de forma explícita, fóra: $Z =$

$$\begin{bmatrix} -B^{-1}S \\ \mathbb{1} \\ \mathbb{0} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 1 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ ja que } BV_1 = S_1 \text{ dóna } V_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \text{ i } BV_2 = S_2,\ V_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

GRADIENT I GRADIENT PROJECTAT

$$\nabla f(X)' = \begin{bmatrix} \frac{{x_3}^2}{2} - x_5 + 3 \\ -6 \\ x_1 x_3 - 4 \\ -4x_6 \\ -x_1 \\ -4x_4 + 2x_6 x_7 x_8 \\ {x_6}^2 x_8 \\ {x_6}^2 x_7 \end{bmatrix} \text{ i } \nabla f(X_k)' = \begin{bmatrix} -\frac{3}{2} \\ -6 \\ -2 \\ -4 \\ -2 \\ 8 \\ 2 \\ 5 \end{bmatrix}$$

Tenint en compte la ordenació i distribució de variables bàsiques $X_B = \{x_1, x_3, x_6\}$ superbàsiques $X_S = \{x_4, x_8\}$ i no bàsiques $X_N = \{x_2, x_5, x_7\}$, tenim que:

$$G_B = \begin{bmatrix} -\frac{3}{2} \\ -2 \\ 8 \end{bmatrix} \quad G_S = \begin{bmatrix} -4 \\ 5 \end{bmatrix} \quad G_N = \begin{bmatrix} -6 \\ -2 \\ 2 \end{bmatrix} \text{ i que la solució de } B'\Pi = G_B \text{ és:}$$

$$\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & -6 & 3 \\ & 1 & 0 \\ & & -3 \end{bmatrix} \cdot \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ -2 \\ 8 \end{bmatrix} \quad \rightarrow \quad \Pi = \begin{bmatrix} \frac{5}{6} \\ -2 \\ -\frac{16}{3} \end{bmatrix}$$

$$\text{i } Z'G = G_S - S'\Pi = \begin{bmatrix} -4 \\ 5 \end{bmatrix} - \begin{bmatrix} -3 & 8 & -3 \\ 3 & -3 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{5}{6} \\ -2 \\ -\frac{16}{3} \end{bmatrix} = \begin{bmatrix} -3/2 \\ -7/2 \end{bmatrix}$$

DIRECCIÓ DE MENYS EL GRADIENT PROJECTAT

Donat que $||Z'G||^2$=14,5 no està a prop de zero, haurem de fer una passa en la direcció P_S=P_z=$-Z'G$ (o en la que s'obtindria de resoldre $Z'H_kZP_z$=$-Z'G$) i calcular P_B=$-B^{-1}SP_S$, és a dir, resoldre $BP_B = -SP_S$ i fer P_N=$\underline{0}$.

$$\text{Calculant } SP_s = \begin{bmatrix} -3 & 3 \\ 8 & -3 \\ -3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3/2 \\ -9/2 \end{bmatrix}$$

$$\text{i resolent } \begin{bmatrix} 3 & & \\ -6 & 1 & \\ 3 & 0 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 4 \\ & & 1 \end{bmatrix} \cdot \begin{bmatrix} p_{B1} \\ p_{B2} \\ p_{B3} \end{bmatrix} = \begin{bmatrix} -6 \\ -3/2 \\ 9/2 \end{bmatrix} \text{ obtenim } P_B = \begin{bmatrix} -2 \\ 1/2 \\ -7/2 \end{bmatrix}$$

La derivada direccional $\nabla f(X_k)P$=G'_BP_B+G'_SP_S=$-29/2$<0 per tant queda comprovat que és una direcció de descens.

Solució del problema 66.

En la minimització de:

$$\begin{aligned} \min\ f(X) &= \frac{{x_1}^2 x_5}{10} + x_2x_3 - 3x_4{x_8}^2 + 2x_5x_7 - 2x_6 + 3x_8 \\ \text{subj.} & \begin{bmatrix} 3 & 1 & 0 & -3 & 3 & 0 & 1 & 3 \\ -6 & 1 & 1 & 8 & 6 & 4 & 0 & -3 \\ 3 & -1 & 0 & -3 & 2 & -3 & 1 & 0 \end{bmatrix} X = \begin{bmatrix} 31 \\ 92/3 \\ 11 \end{bmatrix} \\ & \underline{0} \leq X \leq \overline{X} \quad \text{amb} \quad \overline{X} = [\,5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5\,]' \end{aligned}$$

pel procediment de Murtagh-Saunders, s'ha trobat que en el punt X_0=$[\,3 \quad 0 \quad 3 \quad \frac{1}{3} \quad 5 \quad 4 \quad 5 \quad 1\,]'$ i partint de la subdivisió en variables bàsiques: X_B=$\{x_1, x_3, x_6\}$, superbàsiques: X_S=$\{x_4, x_8\}$ i no bàsiques: X_N=$\{x_2, x_5, x_7\}$, el gradient projectat val zero: $Z'\nabla f(X_0)'$=$\underline{0}$. Indiqueu si X_0 pot ser un òptim, i si no ho podés ser, quines variables haurien de canviar de valor i com calcularieu el seu canvi.

4.3 Optimització amb constriccions qualssevol

Solució del problema 76.

a) La funció objectiu és $f(x_1, x_2) = \frac{x_1^2}{10} - \ln(x_1 + x_2 + 2)$. Per comprovar si és o no convexa només cal veure si la seva hessiana és o no semidefinida positiva. El gradient de $f(x_1, x_2)$ és:

$$\nabla f(x_1, x_2) = \left(\ \frac{x_1}{5} - \frac{1}{x_1+x_2+2} \quad -\frac{1}{x_1+x_2+2}\ \right)$$

I la hessiana de $f(x_1, x_2)$ és:

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \frac{1}{5} + \frac{1}{(x_1+x_2+2)^2} & \frac{1}{(x_1+x_2+2)^2} \\ \\ \frac{1}{(x_1+x_2+2)^2} & \frac{1}{(x_1+x_2+2)^2} \end{pmatrix}$$

Per veure si $\nabla^2 f$ és semidefinida positiva, observem el signe del determinant dels dos menors principals Δ_1 i $\Delta_2 = \nabla^2 f$

$$\begin{aligned} \det(\Delta_1) &= \frac{1}{5} + \frac{1}{(x_1 + x_2 + 2)^2} > 0 \\ \det(\Delta_2) &= \left(\frac{1}{5} + \frac{1}{(x_1 + x_2 + 2)^2} \right) \frac{1}{(x_1 + x_2 + 2)^2} - \frac{1}{(x_1 + x_2 + 2)^2} \frac{1}{(x_1 + x_2 + 2)^2} \\ &= \frac{1}{5} \frac{1}{(x_1 + x_2 + 2)^2} > 0 \end{aligned}$$

Com que els determinants dels dos menors principals de la hessiana són positius, podem assegurar que la funció objectiu és convexa (en aquest cas, a més, podem garantir que és estrictament convexa, ja que els determinants són > 0).

b) Per poder definir la funció de penalització, primer reescrivim de forma escaient el problema:

$$\begin{aligned} \min_{x_1, x_2} \quad & \frac{x_1^2}{10} - \ln(x_1 + x_2 + 2) \\ \text{subj. a} \quad & \\ & 10 - x_1^2 - x_2^2 \geq 0 \\ & x_1 - 1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$

I ara la funció de penalització amb $\rho = 1$ serà:

$$P(x, 1) = \frac{x_1^2}{10} - \ln(x_1 + x_2 + 2) + \frac{1}{10 - x_1^2 - x_2^2} + \frac{1}{x_1 - 1} + \frac{1}{x_2}$$

c) La nova funció de penalització del SUMT amb la penalització logarísmica ve donada per l'expressió:

$$P(x, 1) = \frac{x_1^2}{10} - \ln(x_1 + x_2 + 2) - \ln(10 - x_1^2 - x_2^2) - \ln(x_1 - 1) - \ln x_2$$

Quan ens trobem en un punt x tal que $g(x) \geq 0$ (és a dir, x és un punt factible), la nova funció $-\ln g(x)$ evita que ens atansem a valors on $g(x) = 0$, ja que quan $g(x) \to 0^+$ aleshores $-\ln g(x) \to +\infty$, tot incrementant el valor de la funció $P(x, \rho)$ (per tant, penalitza el fet de que $g(x)$ s'atansi a 0).

Podem observar que $-\ln g(x) = \ln \frac{1}{g(x)}$, amb el qual veiem que la nova penalització no és més que la penalització inversament proporcional que ja teníem, però afectada per la funció ln. Això fa que quan $g(x) \to 0^+$ "suavitzem" el valor de la penalització (ja que $\ln \frac{1}{g(x)} << \frac{1}{g(x)}$). Hi ha però, una diferència destacable entre ambdues penalitzacions: $\frac{1}{g(x)}$ sempre penalitza (sempre pren valors positius dins la regió factible on $g(x) \geq 0$), mentre que $\ln \frac{1}{g(x)}$ pot prendre valors negatius (quan $\frac{1}{g(x)} < 1$) afavorint llavors el fet d'estar allunyats de $g(x) = 0$.

d) Donada la funció de penalització

$$P(x,1) = \frac{x_1^2}{10} - \ln(x_1 + x_2 + 2) - \ln(10 - x_1^2 - x_2^2) - \ln(x_1 - 1) - \ln x_2$$

el seu gradient ve donat per

$$\nabla P(x,1)\mathrm{T} = \begin{pmatrix} \frac{x_1}{5} - \frac{1}{x_1+x_2+2} - \frac{1}{x_1-1} + \frac{2x_1}{10-x_1^2-x_2^2} \\ -\frac{1}{x_1+x_2+2} - \frac{1}{x_2} + \frac{2x_2}{10-x_1^2-x_2^2} \end{pmatrix}$$

i la matriu hessiana és:

$$\nabla^2 P(x,1) = \begin{pmatrix} \frac{1}{5} + \frac{1}{(x_1+x_2+2)^2} + \frac{1}{(x_1-1)^2} + \frac{2}{10-x_1^2-x_2^2} + \frac{4x_1^2}{(10-x_1^2-x_2^2)^2} & \frac{1}{(x_1+x_2+2)^2} + \frac{4x_1x_2}{(10-x_1^2-x_2^2)^2} \\ \frac{1}{(x_1+x_2+2)^2} + \frac{4x_1x_2}{(10-x_1^2-x_2^2)^2} & \frac{1}{(x_1+x_2+2)^2} + \frac{1}{x_2^2} + \frac{2}{10-x_1^2-x_2^2} + \frac{4x_2^2}{(10-x_1^2-x_2^2)^2} \end{pmatrix}$$

Ara escrivim $\nabla^2 P(x,1)$ com la suma de 3 matrius $\nabla^2 P(x,1) = H_1 + H_2 + H_3$ on

$$H_1 = \begin{pmatrix} \frac{1}{5} + \frac{1}{(x_1+x_2+2)^2} & \frac{1}{(x_1+x_2+2)^2} \\ \frac{1}{(x_1+x_2+2)^2} & \frac{1}{(x_1+x_2+2)^2} \end{pmatrix} \qquad H_2 = \begin{pmatrix} \frac{1}{(x_1-1)^2} & 0 \\ 0 & \frac{1}{x_2^2} \end{pmatrix}$$

$$H_3 = \begin{pmatrix} \frac{2}{10-x_1^2-x_2^2} + \frac{4x_1^2}{(10-x_1^2-x_2^2)^2} & \frac{4x_1x_2}{(10-x_1^2-x_2^2)^2} \\ \frac{4x_1x_2}{(10-x_1^2-x_2^2)^2} & \frac{2}{10-x_1^2-x_2^2} + \frac{4x_2^2}{(10-x_1^2-x_2^2)^2} \end{pmatrix}$$

H_1 és la part de hessiana associada amb la funció objectiu, i hem vist a l'apartat a) que era definida positiva. H_2 és la part de hessiana associada amb la penalització dels límits de les variables, i clarament és una matriu definida positiva. H_3 està associada a la penalització de la restricció $x_1^2 + x_2^2 \leq 10$. Hem de veure que és una matriu definida positiva dins la regió factible, tal i com diu l'enunciat. Estudiarem el signe del determinant dels dos menors principals. Per comoditat definim $\alpha = 10 - x_1^2 - x_2^2$. Clarament $\alpha \geq 0$ dins la regió factible ($x_1^2 + x_2^2 \leq 10 \Rightarrow \alpha \geq 0$). Un cop definit α, observem el signe del determinant dels menors principals:

$$\det(\Delta_1) = \frac{2}{\alpha} + \frac{4x_1^2}{\alpha^2} > 0$$

$$\det(\Delta_2) = \left(\frac{2}{\alpha} + \frac{4x_1^2}{\alpha^2}\right)\left(\frac{2}{\alpha} + \frac{4x_2^2}{\alpha^2}\right) - \frac{16x_1^2x_2^2}{\alpha^4} = \frac{4}{\alpha} + \frac{8x_1^2}{\alpha^3} + \frac{8x_2^2}{\alpha^3} > 0$$

Com que H_1, H_2 i H_3 són definides positives dins la regió factible, podem concloure que $\nabla^2 P(x,1)$ és definida positiva dins la regió factible, i per tant $P(x,1)$ és una funció convexa dins aquesta regió.

Donada la convexitat de $P(x,1)$, podríem usar el mètode de Newton per a problemes sense restriccions, ja que tenim garantida una hessiana definida positiva dins la regió factible, el qual ens evita problemes de convergència associats amb el fet de trobar hessianes indefinides (aquest és un dels inconvenients més greus del mètode de Newton).

e) El punt inicial d'iteració és $x^0 = (2 \;\; 1)T$. El nou punt el trobarem com

$$x^1 = x^0 - (\nabla^2 P(x^0,1))^{-1} \nabla P(x^0,1)T$$

essent $P(x^0,1)$ la penalització logarísmica de l'apartat c) (el terme ρ val 1).

Només hem d'avaluar el gradient i hessiana de $P(x,1)$ al punt x^0. Usant l'expressió del gradient i hesssiana trobats a l'apartat anterior, directament obtenim:

$$\nabla P(x^0,1)T = \begin{pmatrix} 0 \\ -4/5 \end{pmatrix} \qquad \nabla^2 P(x^0,1) = \frac{1}{25}\begin{pmatrix} 57 & 9 \\ 9 & 40 \end{pmatrix} \Rightarrow (\nabla^2 P(x^0,1))^{-1} = \frac{25}{2199}\begin{pmatrix} 40 & -9 \\ -9 & 57 \end{pmatrix}$$

Finalment calculem el nou punt x^1:

$$x^1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{25}{2199}\begin{pmatrix} 40 & -9 \\ -9 & 57 \end{pmatrix}\begin{pmatrix} 0 \\ -4/5 \end{pmatrix} = \begin{pmatrix} 1.9182 \\ 1.5184 \end{pmatrix}$$

Podem comprovar que hem millorat el valor de la funció objectiu:

$$P(x^0,1) = \frac{2^2}{10} - \ln(2+1+2) - \ln(10-2^2-1^2) - \ln(2-1) - \ln 1 = -2.8189$$
$$P(x^1,1) = \frac{1.9^2}{10} - \ln(1.9+1.5+2) - \ln(10-1.9^2-1.5^2) - \ln(1.9-1) - \ln 1.5 = -3.0476$$