Capitolul 2

Probleme de optimizare convexă

2.1 Preliminarii

Fie problema de optimizare,

$$\begin{aligned} \min_{x\in\mathbb{R}^n}\ & f(x) \\ \text{s.l.:}\ & g_i(x) \leq 0 \ \ \forall i = 1,\dots,m \\ & Ax = b. \end{aligned} \tag{2.1}$$

Dacă funcţia obiectiv f şi funcţiile ce definesc constrângerile de inegalitate g_i sunt convexe atunci problema se numeşte *problemă de optimizare convexă*. Convexitatea deţine un rol crucial în optimizare, deoarece problemele cu această proprietate prezintă trăsături teoretice foarte bune (e.g., punctele de optim locale sunt, de asemenea, puncte de optim globale); şi ceea ce este mai important, pot fi rezolvate numeric în mod eficient, ceea ce nu este valabil pentru problemele neconvexe. Vom studia mai departe noţiunile de *funcţie convexă* şi *mulţime convexă* pentru a sublinia proprietăţile remarcabile ale problemelor convexe.

2.1.1 Mulţimi convexe

Pentru o expunere clară şi concisă, introducem următoarele definiţii:

Definiţia 1. Multimea $S \subseteq \mathbb{R}^n$ se numeşte *convexă* dacă pentru oricare două puncte $x_1, x_2 \in S$ şi un scalar $\alpha \in [0,1]$ avem $\alpha x_1 + (1-\alpha)x_2 \in S$, i.e. segmentul generat de oricare două puncte din S este inclus în S (Fig. 2.1).

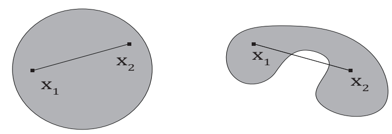

Figura 2.1: *Exemplu de mulţime convexă (stânga) şi mulţime neconvexă (dreapta).*

Exemplul 11. Fie mulţimea $\mathcal{L}_n = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} : \|x\| \leq t \right\}$, denumită şi *conul Lorentz* sau conul de îngheţată. Să se traseze graficul acestei mulţimi în Matlab pentru $n = 2$.

Rezolvare. Figura conului (Fig. 2.2) poate fi trasată de următoarea secvenţă de cod:

```
function create_cone
[y1,y2]=meshgrid(-1:0.01:1,-1:0.01:1);
y3=sqrt(y1.^2+y2.^2); dom=[y3>1];
z3=y3; z3(dom)=inf;
figure(1);
hold on;
surf(y1,y2,z3); set(gca,'FontSize',15);
hold off;
end
```

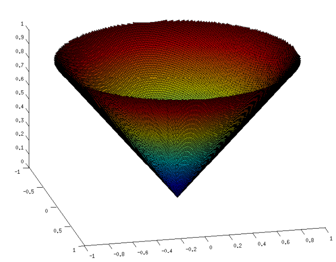

Figura 2.2: *Con de ordinul II (con Lorentz) în $\mathbb{R}^3$ pentru $n = 2$.*

Studiul mulţimilor convexe şi al proprietăţilor acestora facilitează analiza mulţimilor fezabile aferente problemelor de tipul (2.1) şi a şirurilor generate de algoritmi cu restricţii la aceste mulţimi. În continuare vom introduce şi analiza noţiunea de funcţie convexă.

2.1.2 Funcţii convexe

Definiţia 2. Funcţia $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte *convexă* dacă domeniul său efectiv dom$f$ este o mulţime convexă şi dacă următoarea inegalitate:

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2), \tag{2.2}$$

este satisfăcută pentru orice $x_1, x_2 \in \text{dom}f$ şi $\alpha \in [0,\ 1]$.

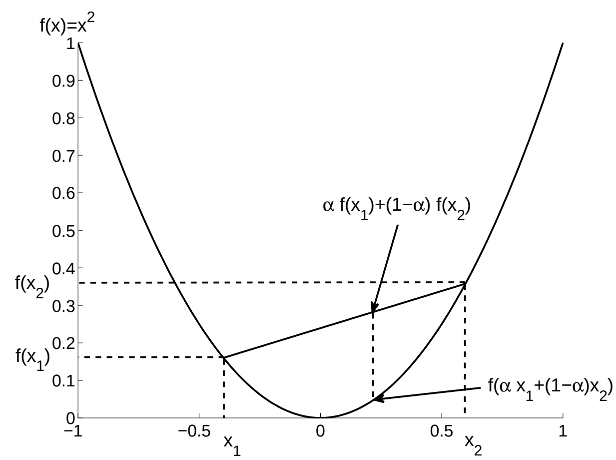

Figura 2.3: *Exemplu de funcţie convexă* $f(x) = x^2$.

Inegalitatea (2.2) se poate generaliza la un număr p de puncte, purtând numele de *inegalitatea lui Jensen*: f este o funcţie convexă dacă şi numai dacă:

$$f\left(\sum_{i=1}^{p} \alpha_i x_i\right) \leq \sum_{i=1}^{p} \alpha_i f(x_i)$$

pentru orice $x_i \in \text{dom}f$ şi $\alpha_i \in [0,\ 1]$, cu $\sum_i^p \alpha_i = 1$.

Interpretarea geometrică a convexităţii este foarte simplă (Fig. 2.3). Pentru o funcţie convexă, fie două puncte din domeniul său $x, y \in \mathrm{dom} f$, atunci valorile funcţiei evaluate în punctele din intervalul $[x, y]$ sunt mai mici sau egale decât cele aflate pe segmentul cu capetele $(x, f(x))$ şi $(y, f(y))$.

Remarca 1. O funcţie este convexă dacă şi numai dacă restricţia domeniului său la o dreaptă (care intersectează domeniul) este de asemenea, convexă. Cu alte cuvinte, f este convexă dacă şi numai dacă oricare ar fi $x \in \mathrm{dom} f$ şi o direcţie d, funcţia $g(\alpha) = f(x + \alpha d)$ este convexă pe domeniul $\{\alpha \in \mathbb{R} : x + \alpha d \in \mathrm{dom} f\}$. Această proprietate este foarte utilă în problemele ce implică teste de convexitate ale funcţiilor.

Pentru a sublinia legătura dintre noţiunile introduse anterior, facem următoarea observaţie: orice mulţime M definită ca mulţimea subnivel a unei anumite funcţii f, i.e. $M = \{x \in \mathbb{R}^n : f(x) \leq c\}$, unde c este o constantă, este convexă dacă şi numai dacă funcţia f este convexă (pentru demonstraţie vezi [1]).

Exemplul 12. Fie funcţia convexă $f : \mathbb{R}^2 \to \mathbb{R}$, definită de $f(x) = \frac{1}{2} x^T \begin{bmatrix} 2 & -1 \\ -1 & 0.5 \end{bmatrix} x + [-0.5 \ 1] x$. Să se traseze graficul funcţiei în Matlab.

Rezolvare. Graficul funcţiei (Fig. 2.4) este generat de următoarea secvenţă de cod (unde am folosit funcţia $gradient(z)$ pentru a descrie culoarea, graduală de-a lungul axei z, graficului rezultat):

```
function create_parab
[x,y] = meshgrid([-10:.2:10]);
z=x.^2+0.25*(y.^2)-x.*y-0.5*x+y;
surf(x,y,z,gradient(z))
end
```

2.2 Pachet CVX

În această subsecţiune expunem o scurtă prezentare a pachetului de optimizare CVX. CVX reprezintă un sistem de modelare a problemelor de optimizare convexă pe baza limbajului Matlab. CVX realizează transformarea comenzilor Matlab într-un limbaj de modelare, cu

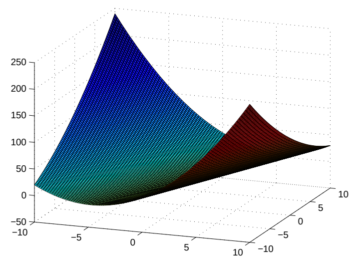

Figura 2.4: *Graficul funcţiei pătratice* $f(x) = x_1^2 + 1/4x_2^2 - x_1x_2 - 1/2x_1 + x_2$.

posibilitatea de declarare a constrângerilor şi funcţiei obiectiv prin intermediul sintaxei Matlab standard.
De exemplu, considerăm următorul model de optimizare convexă:

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$
$$\text{s.l.: } Cx = d, \|x\|_\infty \leq 1.$$

Următoarea secvenţă de cod generează şi rezolvă o instanţă aleatorie a acestui model:

```
m = 10; n = 8; p = 2;
A = randn(m,n); b = randn(m,1);
C = randn(p,n); d = randn(p,1);
cvx_begin
    variable x(n)
    minimize(norm(A*x - b, 2))
    subject to
        C*x == d
        norm(x, Inf) <= 1
cvx_end
```

2.2.1 Instalare

Un scurt ghid de instalare presupune următorii paşi:

se descarcă arhiva de la adresa : `http://cvxr.com/cvx/`, fişier `.zip` sau `.tar.gz`;

se dezarhivează într-un director dorit (diferit de directorul `toolbox` al Matlab-ului);

se porneşte Matlab-ul; se comută directorul curent în directorul unde am realizat dezarhivarea şi se execută în consolă comanda `cvx_setup`.

Remarca 2. În unele cazuri (de cele mai multe ori în Linux) este nevoie de crearea sau modificarea fişierului `startup.m` ce elimină necesitatea ca utilizatorul să introducă comanda `cvx_setup` la fiecare pornire a Matlab-ului.

2.2.2 Elemente de bază

Dacă instalarea este realizată cu succes, începem prin a preciza că orice program CVX se scrie în interiorul unei funcţii Matlab. Pentru a diferenţia conţinutul codului CVX de restul programului Matlab, programele CVX se delimitează cu comenzile `cvx_begin` şi `cvx_end`. Valorile variabilelor create în porţiunea de cod Matlab se pot folosi ca parametri în problemele de optimizare rezolvate cu CVX. La începutul oricărui program CVX se definesc variabilele de decizie şi dimensiunile acestora, e.g.:

```
m = 10; n = 8; A = randn(m,n); b = randn(m,1);
cvx_begin
    variable x(n)
    minimize( norm(A*x-b) )
cvx_end
```

În acest exemplu, variabila de decizie este vectorul x din $\mathbb{R}^n$, iar matricea A şi vectorul b reprezintă variabile Matlab ce sunt utilizate ca parametri în cadrul CVX. *Variabilele* CVX se declară folosind comanda `variable` şi specificarea dimensiunii variabilei, înainte de a se utiliza în constrângeri sau în expresia funcţiei obiectiv. Sintaxa declarării unei variabile poate fi una din următoarele forme:

```
variable x(10)    variable Y(20,10)   variable Z(5,5,5).
```

Sunt disponibile o varietate de opţiuni adiţionale pentru precizarea structurii matriceale, în cazul în care variabila de decizie este de tip matrice cu proprietăţi speciale (simetrică, Toeplitz) e.g.:

```
variable Y(50,50) symmetric
variable Z(100,100) hermitian toeplitz.
```

Pentru lista întreagă a opţiunilor se poate consulta adresa:

`http://cvxr.com/cvx/doc/basics.html.`

Declararea *funcţiei obiectiv* a problemei de optimizare necesită precizarea tipului de problemă (e.g. minimizare, maximizare) prin intermediul cuvintelor cheie `minimize` şi `maximize`, e.g.

```
minimize( norm( x, 1 ) )
maximize( geo_mean( x ) ).
```

Precizăm că este imperativ ca funcţia obiectiv să fie convexă când folosim `minimize` şi concavă când folosim `maximize`. În caz contrar, pachetul CVX va furniza un mesaj de eroare corespunzător. *Constrângerile* suportate de modelele CVX sunt cele de egalitate (liniare) impuse prin operatorul `==` şi de inegalitate impuse de operatorii `<=` şi `>=`. Pentru constrângeri de tip box (lanţ de inegalităţi) este disponibilă sintaxa `l<=x<=u`.

2.3 Probleme rezolvate de laborator

Exemplul 13. Să se determine bila de rază maximă (i.e. centrul şi raza ei) din spaţiul Euclidian bidimensional, ce poate fi înscrisă într-un poliedru descris de inegalităţi liniare (Fig. 2.5).

Rezolvare. În cazul general, fiind dat setul de inegalităţi liniare $a_i^T x \leq b_i$, formularea problemei în termeni de optimizare este dată de:

$$\max_{x\in\mathbb{R}^n, r\in\mathbb{R}} \; r$$
$$\text{s.l.: } a_i^T x + r\|a_i\|_2 \leq b_i \;\; \forall i = 1, \ldots, n,$$

unde r reprezintă raza bilei, iar x centrul acesteia. Alegând datele inegalităţilor liniare, coordonatele centrului şi raza bilei dorite sunt generate de următoarea secvenţă de program:

```
% Generare date de intrare
a1 = [ 2;  1]; a2 = [ 2; -1]; a3 = [-1;  2];
a4 = [-1; -2]; b = ones(4,1);
% Crearea si rezolvarea problemei de optimizare
cvx_begin
    variable r(1)
```

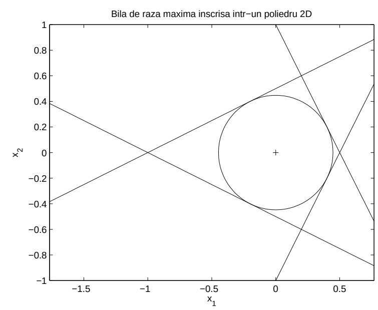

Figura 2.5: *Bila de rază maximă înscrisă într-un poliedru.*

```
    variable x_c(2)
    maximize ( r )
    a1'*x_c + r*norm(a1,2) <= b(1);
    a2'*x_c + r*norm(a2,2) <= b(2);
    a3'*x_c + r*norm(a3,2) <= b(3);
    a4'*x_c + r*norm(a4,2) <= b(4);
cvx_end

% Generare figura
x = linspace(-2,2);
theta = 0:pi/100:2*pi;
plot( x, -x*a1(1)./a1(2) + b(1)./a1(2),'b-');
hold on
plot( x, -x*a2(1)./a2(2) + b(2)./a2(2),'b-');
plot( x, -x*a3(1)./a3(2) + b(3)./a3(2),'b-');
plot( x, -x*a4(1)./a4(2) + b(4)./a4(2),'b-');
plot( x_c(1) + r*cos(theta), x_c(2) + r*sin(theta), 'r');
plot(x_c(1),x_c(2),'k+'); xlabel('x_1'); ylabel('x_2');
title('Bila de raza maxima inscrisa intr-un poliedru 2D');
axis([-1 1 -1 1]); axis equal;
```

Exemplul 14. Determinaţi filtrul FIR cu cel mai apropiat comportament de cel al unei funcţii de transfer date $H_d(\omega)$ (e.g. $H_d(\omega) = e^{-j\omega D}$). Să se realizeze proiectarea filtrului prin intermediul minimizării erorii absolute maxime (în norma infinit).

Rezolvare. Formularea problemei precedente în termeni de optimizare conduce la rezolvarea următoarei probleme:

$$\min_{H} \max_{\omega} |H(\omega) - H_d(\omega)|, \tag{2.3}$$

unde H este funcţia răspunsului în frecvenţă, iar variabila de decizie are rolul răspunsului la impuls. O aproximare convenabilă a problemei anterioare presupune extragerea unui set finit m de frecvenţe ω_i, cu $i = 1, \dots, m$, şi definirea unei matrice A:

$$A = \begin{bmatrix} 1 & e^{-j\omega_1} & e^{-2j\omega_1} & \dots & e^{-nj\omega_1} \\ 1 & e^{-j\omega_2} & e^{-2j\omega_2} & \dots & e^{-nj\omega_2} \\ & \dots & \dots & \dots & \dots \\ 1 & e^{-j\omega_m} & e^{-2j\omega_m} & \dots & e^{-nj\omega_m}. \end{bmatrix}$$

rezultând o aproximare a problemei (2.3), i.e.

$$\min_{x \in \mathbb{R}^n} \max_{1 \le i \le m} |A_i x - H_d(\omega_i)|, \tag{2.4}$$

unde A_i reprezintă linia i a matricei A.
Mai departe, prezentăm secvenţa de cod ce rezolvă problema de optimizare (2.4) şi determinăm aproximarea optimă a funcţiei de transfer $H_d(\omega) = e^{-j\omega D}$ (reprezentată în Fig. 2.6):

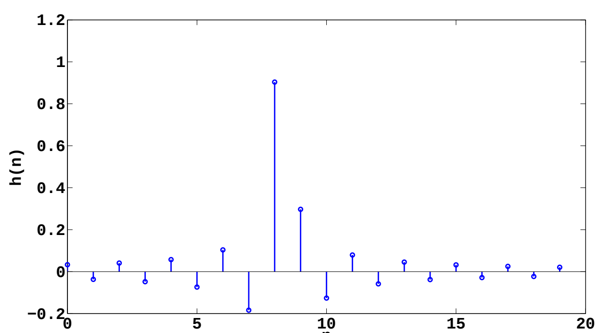

Figura 2.6: *Aproximarea optimă în sens Chebyshev a funcţiei de transfer* $H_d(\omega) = e^{-j\omega D}$.

```
n = 20; m = 15*n;
w = linspace(0,pi,m)'; % omega
% Construim un raspuns in frecventa dorit
```

```
D = 8.25;           % valoare intarziere
Hdes = exp(-j*D*w);  % raspuns in frecventa dorit

% Filtru Gaussian cu faza liniara (decomentati liniile
% de mai jos pentru proiectare)
% var = 0.05;
% Hdes = 1/(sqrt(2*pi*var))*exp(-(w-pi/2).^2/(2*var));
% Hdes = Hdes.*exp(-j*n/2*w);

% Rezolvam problema minimax de proiectare a filtrului
% A reprezinta matricea folosita la calculul
% raspunsului in frecventa
% A(w,:) = [1 exp(-j*w) exp(-j*2*w) ... exp(-j*n*w)]
A = exp( -j*kron(w,[0:n-1]) );

% formularea optimala a filtrului Chebyshev
cvx_begin
  variable h(n,1)
  minimize( max( abs( A*h - Hdes ) ) )
cvx_end

% verificam daca problema a fost rezolvata cu succes
disp(['Problem is ' cvx_status])
if ~strfind(cvx_status,'Solved')
  h = [];
end

% Generam figurile aferente filtrului
figure(1); stem([0:n-1],h); xlabel('n'); ylabel('h(n)');

% figura raspunsului in frecventa
H = [exp(-j*kron(w,[0:n-1]))]*h;
figure(2)
% magnitudine
subplot(2,1,1);
plot(w,20*log10(abs(H)),w,20*log10(abs(Hdes)),'--');
xlabel('w'); ylabel('mag H in dB'); axis([0 pi -30 10]);
legend('optimizat','dorit','Location','SouthEast')
% faza
```

```
subplot(2,1,2); plot(w,angle(H));
axis([0,pi,-pi,pi]); xlabel('w'), ylabel('faza H(w)');
```

2.4 Probleme rezolvate de seminar

2.4.1 Mulţimi şi funcţii convexe

Problema 1. Fie mulţimea S descrisă de:

$$\begin{aligned} S &= \{x \in \mathbb{R}^n : a_i^T x = b_i,\ c_j^T x \leq d_j, \quad i = 1, \dots, m, j = 1, \dots, p\} \\ &= \{x \in \mathbb{R}^n : Ax = b,\ Cx \leq d\}. \end{aligned}$$

Să se demonstreze că mulţimea este convexă (mulţimile definite de egalităţi şi inegalităţi liniare, cum este cea din enunţ, se numesc *poliedre*).

Rezolvare. Se arată că pentru orice $x_1, x_2 \in S$ şi $\alpha \in [0, 1]$ rezultă $\alpha x_1 + (1 - \alpha)x_2 \in S$. Dacă $x_1, x_2 \in S$ atunci avem:

$$\begin{cases} Ax_1 = b, & Cx_1 \leq d \\ Ax_2 = b, & Cx_2 \leq d. \end{cases}$$

Notând $x_\alpha = \alpha x_1 + (1 - \alpha)x_2$, putem finaliza demonstraţia prin următoarea observaţie:

$$\begin{aligned} Ax_\alpha &= \alpha Ax_1 + (1 - \alpha)Ax_2 = \alpha b + (1 - \alpha)b = b \\ Cx_\alpha &= \alpha Cx_1 + (1 - \alpha)Cx_2 \leq \alpha d + (1 - \alpha)d = d. \end{aligned}$$

În acest fel am demonstrat că mulţimea definită anterior este convexă.

Problema 2. Să se demonstreze că următoarele mulţimi se pot defini sub forma unor poliedre:

(i) $S_1 = \{x \in \mathbb{R}^n : x \geq 0, \|x\|_1 = 1\}$

(ii) $S_2 = \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1\}$

(iii) $S_3 = \{x \in \mathbb{R}^n : \|x\|_1 \leq 1\}$.

Rezolvare. : (i) În cazul mulţimii S_1 observăm următoarele echivalenţe:

$$S_1 = \left\{ x \in \mathbb{R}^n : x \geq 0,\ \sum_{i=1}^{n} x_i = 1 \right\} = \left\{ x \in \mathbb{R}^n : -x \leq 0,\ \sum_{i=1}^{n} x_i = 1 \right\}$$
$$= \{ x \in \mathbb{R}^n : -I_n x \leq 0,\ [1 \dots 1]x = 1 \},$$

unde I_n reprezintă matricea identitate de ordin n. Ultima formulare denotă forma poliedrală a mulţimii S_1.
(ii) De asemenea, în cazul mulţimii S_2, urmând acelaşi raţionament:

$$S_2 = \left\{ x \in \mathbb{R}^n : \max_{1 \leq i \leq n} |x_i| \leq 1 \right\} = \{ x \in \mathbb{R}^n : |x_i| \leq 1, \quad \forall i = 1, \dots, n \}$$
$$= \{ x \in \mathbb{R}^n : -1 \leq x_i \leq 1\ \forall i = 1, \dots, n \}$$
$$= \{ x \in \mathbb{R}^n : -x_i \leq 1, x_i \leq 1, \quad \forall i = 1, \dots, n \}$$
$$= \left\{ x \in \mathbb{R}^n : \begin{bmatrix} I_n \\ -I_n \end{bmatrix} x \leq \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right\},$$

obţinem o formă poliedrală.
(iii) Pentru a determina forma poliedrică a mulţimii S_3 definim o mulţime poliedrică auxiliară:

$$Q = \left\{ [x^T\ t^T]^T \in \mathbb{R}^{2n} : \sum_{i=1}^{n} t_i = 1, |x_i| \leq t_i \right\}.$$

Mai departe, definim proiecţia unei mulţimi poliedrice pe un subspaţiu de dimensiuni reduse şi folosim această noţiune pentru a studia mulţimea S_3.

Definiţia 3. Fie mulţimea $P = \{x \in \mathbb{R}^{n_1}, y \in \mathbb{R}^{n_2} : \ Ax + By \leq b\}$, unde $A \in \mathbb{R}^{m \times n_1}, B \in \mathbb{R}^{m \times n_2}, b \in \mathbb{R}^m$. *Proiecţia* mulţimii P pe subspaţiul variabilelor x este dată de:

$$P_x = \{x \in \mathbb{R}^{n_1} :\ \exists y \in \mathbb{R}^{n_2},\ [x^T\ y^T]^T \in P\}.$$

Proiecţia mulţimii Q pe subspaţiul variabilelor x conduce la următoarea serie de echivalenţe:

$$Q_x = \{x \in \mathbb{R}^n : \exists t \in \mathbb{R}^n \ \text{ a.î. } \ \sum_{i=1}^{n} t_i = 1, |x_i| \leq t_i\}$$
$$= \{x \in \mathbb{R}^n : \sum_{i=1}^{n} |x_i| \leq 1\} = \{x \in \mathbb{R}^n : \|x\|_1 \leq 1\} = S_3.$$

Ţinând cont că proiecţia oricărei mulţimi poliedrice este o mulţime poliedrică, concluzionăm că mulţimea S_3 este de tip poliedric.

Problema 3. Fie mulţimile:

(i) $\mathcal{L}^n = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} : \|x\| \leq t \right\}$ numit şi con Lorentz sau con de ordinul II;

(ii) $\mathcal{S}_+^n = \{X \in \mathcal{S}^n : X \succeq 0\}$ conul semidefinit.

Să se demonstreze că mulţimile precedente sunt conuri auto-duale.

Rezolvare: (i) Din definiţia conului dual avem:

$$\mathcal{L}_n^* = \{y \in \mathbb{R}^n : \langle x, y\rangle \geq 0 \ \forall x \in \mathcal{L}_n\}.$$

Problema se reduce la a demonstra că $\mathcal{L}_n^* = \mathcal{L}_n$, ceea ce este echivalent cu satisfacerea simultană a incluziunilor: $\mathcal{L}^{n*} \subseteq \mathcal{L}^n$ şi $\mathcal{L}^n \subseteq \mathcal{L}^{n*}$. Arătăm acum prima incluziune. Fie $y = \begin{bmatrix} y_1 \\ v \end{bmatrix} \in \mathcal{L}_n^*$, atunci pentru orice $x = \begin{bmatrix} x_1 \\ t \end{bmatrix} \in \mathcal{L}_n$ avem:

$$\langle x, y\rangle = y_1^T x_1 + vt \geq 0. \tag{2.5}$$

Ştiind că $x \in \mathcal{L}_n$ atunci $\|x\| \leq t$. Rămâne să demonstrăm că $\|y\| \leq v$. Din ipoteza că (2.5) are loc pentru orice vector $x \in \mathcal{L}_n$, atunci inegalitatea este satisfăcută, de asemenea, pentru un x ales. Alegând $x = \begin{bmatrix} x_1 \\ t \end{bmatrix} = \begin{bmatrix} -\frac{y_1}{\|y_1\|} \\ 1 \end{bmatrix}$ şi înlocuind în (2.5) obţinem:

$$-\frac{y_1^T y_1}{\|y_1\|} + v = -\|y_1\| + v \geq 0,$$

din care rezultă prima incluziune.

Pentru a doua incluziune, fie $y = \begin{bmatrix} y_1 \\ v \end{bmatrix} \in \mathcal{L}_n, x = \begin{bmatrix} x_1 \\ t \end{bmatrix} \in \mathcal{L}_n$. Din inegalitatea Cauchy-Schwartz $|\langle x, y\rangle| \leq \|x\|\|y\|$ rezultă:

$$y_1^T x_1 + vt > -\|y_1\|\|x_1\| + vt \geq -vt + vt = 0,$$

unde în a doua inegalitate am utilizat ipoteza că $x, y \in \mathcal{L}_n$. În concluzie, pentru orice $y \in \mathcal{L}_n$ avem că $y \in \mathcal{L}_n^*$.
(ii) În mod similar, arătăm prin dublă incluziune că $\mathcal{S}_+^n = \mathcal{S}_+^{n*}$. Fie $Y \in S_+^{n*}$, atunci $\mathrm{Tr}(YX) \geq 0$. Din următoarele relaţii:

$$x^T Y x = Tr(x^T Y x) = Tr(Y x x^T) = Tr(YX) \geq 0,$$

se deduce $Y \in S_+^n$. Pentru a doua incluziune, presupunem Y, X matrice pozitiv semidefinite. Remarcăm următoarea relaţie:

$$Tr(YX) = Tr(YV^T \Delta V) = Tr(VYV^T \Delta) \geq 0, \tag{2.6}$$

în care am folosit descompunerea valorilor proprii corespunzătoare matricei X şi proprietatea de permutare a funcţiei matriceale $Tr(\cdot)$. În concluzie, $Y \in S_+^{n*}$ datorită relaţiei (2.6), de unde reiese a doua incluziune.

Problema 4. Să se demonstreze că următoarele funcţii sunt convexe pe mulţimile specificate:
(i) $f(x) = -\log x, \quad \mathrm{dom} f = (0, \infty)$.
(ii) $f(x) = \frac{1}{2} x^T Q x + q^T x + r, \quad \mathrm{dom} f = \mathbb{R}^n, Q \succeq 0$.
(iii) $f(x, t) = \frac{x^T x}{t}, \quad \mathrm{dom} f = \mathbb{R}^n \times (0, \infty)$.
(iv) $f(x) = \|Ax - b\|, \ A \in \mathbb{R}^{m \times n}, \quad \mathrm{dom} f = \mathbb{R}^n$.
(v) $f(X) = -\log \det X, \quad \mathrm{dom} f = \mathcal{S}_{++}^n$.

Rezolvare: (i) Observăm că funcţia $f(x) = -\log x$ satisface condiţiile de ordinul II ale convexităţii:

$$f''(x) = \frac{1}{x^2} > 0 \ \forall x > 0.$$

(ii) Hessiana funcţiei $f(x) = \frac{1}{2} x^T Q x + q^T x + r$ este dată de $\nabla^2 f(x) = Q$. Observăm că funcţia este convexă deoarece matricea $Q \succeq 0$ (este pozitiv semidefinită).
(iii) În aceeaşi manieră arătăm că funcţia $f(x, t) = \frac{x^T x}{t}$ este convexă pe domeniul $\mathbb{R}^n \times (0, \infty]$. Din definiţia Hessianei rezultă:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{2}{t} I_n & -\frac{2x}{t} \\ -\frac{2x^T}{t^2} & \frac{2x^T x}{t^3} \end{bmatrix},$$

unde cu I_n am notat matricea identitate de ordin n. Pentru a determina dacă funcţia îndeplineşte condiţiile de ordinul II ale convexităţii observăm că:

$$\begin{aligned}
[u^T v^T]\nabla^2 f(x) \begin{bmatrix} u \\ u \end{bmatrix} &= [u^T v^T] \begin{bmatrix} \frac{2}{t}u - \frac{2xv}{t^2} \\ -\frac{2x^T u}{t^2} + \frac{2x^T xv}{t^3} \end{bmatrix} \\
&= \frac{2}{t}u^T u - \frac{2u^T xv}{t^2} - \frac{2x^T uv}{t^2} + \frac{2x^T xv^2}{t^3} \\
&= \frac{2}{t^3}\left(t^2 u^T u - 2tu^T xv + x^T xv^2\right) \\
&= \frac{2}{t^3}\|tu - xv\|^2.
\end{aligned}$$

Ţinând cont de domeniul de definiţie al funcţiei, remarcăm că pentru orice vector $\begin{bmatrix} u \\ v \end{bmatrix}$, termenul drept din ultima egalitate este pozitiv. De aici este evidentă proprietatea de pozitiv definire a matricei Hessiane corespunzătoare funcţiei.

(iv) Deoarece funcţia $\|\cdot\|$ este nediferenţiabilă în punctul 0, observăm că funcţia $f(x) = \|Ax - b\|$ nu este diferenţiabilă în punctele x ce satisfac $Ax = b$. Fie două puncte din mulţimea dom$f$ pentru care $f(x_1) = \|Ax_1 - b\|, f(x_2) = \|Ax_2 - b\|$. Pentru a arăta convexitatea funcţiei f, notăm $x_\alpha = \alpha x_1 + (1-\alpha)x_2$ şi deducem şirul de relaţii:

$$\begin{aligned}
f(x_\alpha) &= f(\alpha x_1 + (1-\alpha)x_2) \\
&= \|A(\alpha x_1 + (1-\alpha)x_2) - b\| \\
&\leq \|\alpha(Ax_1 - b)\| + \|(1-\alpha)(Ax_2 - b)\| \\
&= \alpha\|Ax_1 - b\| + (1-\alpha)\|Ax_2 - b\| \\
&= \alpha f(x_1) + (1-\alpha) f(x_2).
\end{aligned}$$

(v) Fie funcţia $f(X) = -\log\det X,\ X \in \mathcal{S}^n_{++}$. Arătăm convexitatea lui f prin intermediul reducerii domeniului acesteia la o dreaptă. Mai exact, folosim următoarea proprietate a funcţiilor convexe: f este convexă dacă funcţia scalară ce se obţine din restricţionarea la o dreaptă este de asemenea convexă (conform Remarca 1). Revenind la funcţia matriceală, considerăm matricele $X \in \mathcal{S}^n_{++}, D \in \mathcal{S}^n$ şi arătăm că funcţia scalară $g(t) = f(X + tD)$ este convexă, observând următoarele egalităţi:

$$\begin{aligned}
g(t) &= -\log\det(X + tD) \\
&= -\log\det(X^{1/2}X^{1/2} + tD) \\
&= -\log\det\left(X^{1/2}\left(I_n + tX^{-1/2}DX^{-1/2}\right)X^{1/2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= -\log\left(\det X^{1/2} \det\left(I_n + tX^{-1/2}DX^{-1/2}\right) \det X^{1/2}\right)\\
&= -\log\left(\det X \det\left(I_n + tX^{-1/2}DX^{-1/2}\right)\right)\\
&= -\log\det X - \log\det\left(I_n + tX^{-1/2}DX^{-1/2}\right).
\end{aligned}
$$

Pe de altă parte, ştiind că pentru orice matrice $A \in \mathcal{S}^n$, cu spectrul $\Lambda(A) = \{\lambda_1, \dots, \lambda_n\}$, transformarea $B = I_n + tA, t \in \mathbb{R}$, modifică spectrul astfel încât $\Lambda(B) = \{1+t\lambda_1, \dots, 1+t\lambda_n\}$ şi $\det B = \prod\limits_{i=1}^{n}(1+t\lambda_i)$. Pentru a aplica această proprietate în şirul de relaţii precedent, notăm $Z = X^{-1/2}DX^{-1/2}$ având spectrul $\Lambda(Z) = \{\mu_1, \dots, \mu_n\}$. Rescriind $g(t)$, rezultă:

$$
\begin{aligned}
g(t) &= -\log\det X - \log\prod_{i=1}^{n}(1+t\mu_i)\\
&= -\log\det X - \sum_{i=1}^{n}\log(1+t\mu_i).
\end{aligned}
$$

Mai departe, observăm că funcţia $h(t) = -\log(1+tu)$ din componenţa celei anterioare, este convexă, deoarece cea de-a două derivată satisface:

$$
h''(t) = \frac{u}{(1+tu)^2} \geq 0 \ \ \forall u \in \mathbb{R}.
$$

Ştiind că funcţia definită de o sumă de funcţii convexe este convexă, ajungem la concluzia că $g(t)$ este convexă în t. Deci, $f(X)$ este convexă.

Problema 5. Fie funcţia $f : \mathbb{R}^n \to \mathbb{R}$. Să se determine funcţia conjugată $f^*(y)$ pentru următoarele exemple:

(i) $f(x) = e^x$ (ii) $f(x) = x\log x$

(iii)$f(x) = \frac{1}{2}x^TQx, Q \succ 0$ (iv) $f(x) = \log\sum\limits_{i=1}^{n} e^{x_i}$.

Rezolvare. Definim funcţia conjugată corespunzătoare funcţiei f:

$$
f^*(y) = \max_{x \in \mathrm{dom} f} \langle y, x\rangle - f(x).
$$

În primele două cazuri monovariabile (i) şi (ii), observăm că funcţia obiectiv a problemei de maximizare este concavă pe domeniul funcţiei

f, de aceea condiţiile de optimalitate de ordinul I conduc la următoarele expresii:

$$(i)\ f^*(y) = y\log y - y, \qquad (ii)\ f^*(y) = e^{y-1}.$$

(iii) În acest caz, conjugata funcţiei f are forma $f^*(y) = \max_{x\in\mathbb{R}^n} y^T x - \frac{1}{2}x^T Q x$. Datorită proprietăţii de convexitate, observăm că soluţia problemei de optimizare este dată de $x^* = Q^{-1}y$. Înlocuind în expresia funcţiei conjugate rezultă $f^*(y) = y^T Q^{-1} y - \frac{1}{2} y^T Q^{-1} y = \frac{1}{2} y^T Q^{-1} y$.
(iv) Condiţiile necesare de optimalitate de ordinul I corespunzătoare problemei de maximizare $\max_{x\in\mathbb{R}^n}\langle y, x\rangle - \log \sum_{i=1}^n e^{x_i}$ se reduc la următoarele relaţii:

$$\frac{e^{x_i^*}}{\sum_{i=1}^n e^{x_i^*}} = y_i, \quad \forall i = 1,\dots,n. \tag{2.7}$$

Observăm că relaţia (2.7) este satisfăcută şi funcţia $f^*(y)$ ia valori finite numai în cazul în care argumentul funcţiei conjugate satisface $\sum_{i=1}^n y_i = 1, y \geq 0$. Presupunând că argumentul y satisface cele două condiţii, avem $x_i^* = \log \sum_{i=1}^n e^{x_i^*} + \log y_i$ pentru orice $i = 1,\dots,n$. Substituind în expresia funcţiei conjugate rezultă:

$$\begin{aligned}\langle y, x^*\rangle - \log\sum_{i=1}^n e^{x_i^*} &= \sum_{i=1}^n y_i \left(\log\sum_{i=1}^n e^{x_i^*} + \log y_i\right) - \log\sum_{i=1}^n e^{x_i^*}\\ &= \sum_{i=1}^n y_i\log y_i + \log\sum_{i=1}^n e^{x_i^*}\left(\sum_{i=1}^n y_i - 1\right) = \sum_{i=1}^n y_i \log y_i.\end{aligned}$$

În concluzie, funcţia conjugată $f^*(y)$ este definită de:

$$f^*(y) = \begin{cases} \sum_{i=1}^n y_i\log y_i, & \text{dacă } \sum_{i=1}^n y_i = 1, y\geq 0,\\ \infty, & \text{altfel.}\end{cases}$$

Problema 6. Să se demonstreze că următoarea problemă de optimizare este convexă:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} f(x) = {x_1}^2 + {x_2}^2 \\ &\text{s.l.: } g_1(x) = \frac{x_1}{1+{x_2}^2} \leq 0, g_2(x) = e^{x_1+x_2} - 1 \leq 0 \\ &\qquad h(x) = (x_1 - x_2 - 1)^2 = 0. \end{aligned}$$

Rezolvare. Pentru a arăta convexitatea problemei din enunţ:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} f(x) \\ &\text{s.l.: } g_1(x) \leq 0, g_2(x) \leq 0,\ h(x) = 0. \end{aligned}$$

este suficient să demonstrăm convexitatea funcţiilor f, g_1 şi g_2, şi liniaritatea funcţiei h. Observăm că Hessiana funcţiei f are forma explicită $\nabla^2 f(x) = I_2 \succ 0$, unde I_2 este matricea identitate de ordinul II. Deci, funcţia f este convexă deoarece satisface condiţia de convexitate de ordinul II. În cazul funcţiei g_1, distingem faptul că inegalitatea $\frac{x_1}{1+{x_2}^2} \leq 0$ este satisfăcută doar în cazul în care $x_1 \leq 0$. În concluzie, constrângerea $g_1(x) \leq 0$ este echivalentă cu o constrângere liniară (şi deci convexă) $x_1 \leq 0$. Pentru inegalitate $g_2(x) \leq 0$, observăm că este echivalentă cu $x_1 + x_2 \leq 0$, i.e. este şi aceasta echivalentă cu o constrângere liniară. Similar, pentru egalitatea definită de funcţia h, găsim următoarea echivalenţă între mulţimi:

$$\{x \in \mathbb{R}^2 : h(x) = 0\} = \{x \in \mathbb{R}^2 : x_1 - x_2 = 1\},$$

rezultând o funcţie liniară în x. În final, putem rescrie problema sub forma unui QP convex:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} {x_1}^2 + {x_2}^2 \\ &\text{s.l.: } \ x_1 \leq 0,\ \ x_1 + x_2 \leq 0,\ \ x_1 - x_2 = 1. \end{aligned}$$

Problema 7. Să se demonstreze că următoarea problemă de optimizare este convexă:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} f(x) \quad (= -\log(a^T x - b)) \\ &\text{s.l.: } g_1(x) = e^{x^T x} - e \leq 0,\ g_2(x) = (c^T x - d)^2 - 1 \leq 0 \\ &\qquad h(x) = (x_1 + 2x_2)^4 = 0. \end{aligned}$$

Rezolvare. Pentru a demonstra că o problemă de optimizare este convexă, trebuie să demonstrăm că funcţia obiectiv este convexă şi mulţimea fezabilă definită de constrângeri este convexă. Pentru funcţia $f(x) = -\log(a^T x - b)$ deducem expresia gradientului şi a Hessianei:

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

unde avem desigur că $(a^T x - b)^2 > 0$. Dacă notăm $y = a^T x$, observăm că:

$$x^T a a^T x = y^T y = \|y\|_2^2 \geq 0,$$

deci matricea aa^T este pozitiv semidefinită. Drept rezultat, $\nabla^2 f(x) \succeq 0$ peste întregul dom $f = \{x \in \mathbb{R}^n : a^T x - b > 0\}$, i.e. satisface condiţiile de convexitate de ordin II şi deci funcţia obiectiv este convexă. Pentru a arăta că mulţimea constrângerilor este convexă, este suficient să arătăm că funcţiile din constrângerile de egalitate sunt liniare, iar cele din constrângerile de inegalitate sunt convexe. Observăm că constrângerea $h(x) = 0$ este echivalentă cu egalitatea $x_1 + 2x_2 = 0$, a cărei funcţie este liniară. Pentru $g_1(x) \leq 0$, observăm că este echivalentă cu $x^T x - 1 \leq 0$. Funcţia $x^T x - 1$ este o funcţie pătratică, diferenţiabilă de două ori, cu Hessiana $2I_2 \succ 0$, deci satisface condiţiile de convexitate de ordin II şi este implicit convexă. Constrângerea $g_2(x) \leq 0$ este echivalentă cu:

$$-1 \leq (c^T x - d) \leq 1 \Rightarrow \begin{cases} c^T x - d \leq 1 \\ -c^T x + d \leq 1 \end{cases} \Rightarrow \begin{bmatrix} c^T \\ -c^T \end{bmatrix} x + \begin{bmatrix} -d \\ d \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

deci aceasta se reduce la o constrângere de inegalitate unde funcţia este liniară şi implicit convexă.

Problema 8. Să se determine problema convexă de programare semidefinită ce aproximează următoarea problemă neconvexă:

$$\begin{aligned} &\max_{x \in \mathbb{R}^n} x^T A^T A x \\ &\text{ s.l.: } \|x\|_2 \leq 1, \end{aligned}$$

unde $A \in \mathbb{R}^{m \times n}$ şi $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$.

Rezolvare. Reamintim că pentru orice $Q \in \mathbb{R}^{n\times n}$ şi $x \in \mathbb{R}^n$, funcţia $\mathrm{Tr}(Q)$ satisface relaţia: $\mathrm{Tr}(x^T Q x) = \mathrm{Tr}(Qxx^T)$. Pe baza acestei relaţii, problema precedentă se scrie sub următoarea formă echivalentă:

$$\begin{aligned}\max_{X\in\mathbb{R}^{n\times n}} \; & Tr\left(A^T A X\right)\\ \text{s.l.:} \; & \mathrm{rang}(X) = 1,\; Tr(X) = 1.\end{aligned}$$

Se obţine relaxarea convexă prin renunţarea la constrângerea de egalitate neliniară $\mathrm{rang}(X) = 1$. În concluzie, avem următoarea aproximare convexă a problemei de optimizare originală:

$$\begin{aligned}\max_{X\in\mathbb{R}^{n\times n}} \; & Tr\left(A^T A X\right)\\ \text{s.l.:} \; & Tr(X) = 1.\end{aligned}$$

2.5 Probleme propuse

Problema 1. Să se determine care dintre următoarele funcţii sunt convexe, concave sau niciuna dintre cele două variante. Să se argumenteze rezultatele obţinute.

(i) $f(x_1, x_2) = 2x_1^2 - 4x_1x_2 - 8x_1 + 3x_2$,

(ii) $f(x_1, x_2) = x_1 e^{-(x_1+3x_2)}$,

(iii) $f(x_1, x_2) = -x_1^2 - 3x_2^2 + 4x_1x_2 + 10x_1 - 10x_2$,

(iv) $f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1^2 + x_2^2 + 2x_3^2 - 5x_1x_3$,

(v) $f(x_1, x_2, x_3) = -2x_1^2 - 3x_2^2 - 2x_3^2 + 8x_1x_2 + 3x_1x_3 + 4x_2x_3$.

Problema 2. Să se determine submulţimea din $\{x \in \mathbb{R} : x > 0\}$ pe care funcţia $f(x) = e^{ax^b}$ este convexă. Parametrii a, b satisfac $a > 0, b \geq 1$.

Problema 3. Să se determine în domeniul axei reale în care funcţia $f(x) = x^2(x^2 - 1)$ este convexă.

Problema 4. Fie funcţiile convexe $f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$. Să se demonstreze că următoarele compuneri ale acestora sunt convexe:

(i) $g(x) = \sum_{i=1}^{m} \alpha_i f_i(x),\ \alpha_i > 0,$

(ii) $h(x) = \max\{f_1(x), \ldots, f_m(x)\}.$

Problema 5. Fie mulţimea $K = \{x \in \mathbb{R}^n : x \geq 0\}$. Să se demonstreze că mulţimea K este un con şi să se determine conul dual corespunzător.

Problema 6. Să se demonstreze că funcţia $f(x) = \sum_{i=1}^{r} \alpha_i x_{[i]}$ este convexă în x, unde $\alpha_1 \geq \cdots \geq \alpha_r \geq 0$, iar $x_{[n]}$ reprezintă a n-a cea mai mare componentă din vectorul x.

Problema 7. Determinaţi funcţiile conjugate corespunzătoare funcţiilor:

(i) $f(x) = \max_{1 \leq i \leq n} x_i$ cu domeniul $\mathbb{R}^n$;

(ii) $f(x) = x^p$ cu domeniul $\mathbb{R}_{++}$, unde $p > 1$. Comentaţi cazul în care $p < 0$.

(iii) $f(x) = \left(-\prod_{i=1}^{n} x_i\right)^{1/n}$ cu domeniul $\mathbb{R}^n_{++}$;

(iv) $f(x,t) = -\log(t^2 - x^T x)$ cu domeniul $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 \leq t\}$.

Problema 8. Să se demonstreze că funcţia:

$$f(x) = \cfrac{1}{x_1 - \cfrac{1}{x_2 - \cfrac{1}{x_3 - \cfrac{1}{x_4}}}},$$

definită pe domeniul în care toţi numitorii sunt pozitivi, este convexă şi strict descrescătoare. (Cazul când $n = 4$ nu prezintă nicio particularitate, caracteristicile funcţiei se menţin şi pentru cazul general n oarecare.)

Problema 9. Presupunem că funcţia f este convexă, iar $\lambda_1 > 0$, $\lambda_i \leq 0$ pentru $i = 2, \ldots, n$ şi $\sum_{i=1}^{n} \lambda_i = 1$. Fie punctele $x_1, \ldots, x_n \in dom f$. Să se demonstreze că are loc inegalitatea:

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \geq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n).$$

Problema 10. Să se arate că o funcţie $f : \mathbb{R} \to \mathbb{R}$ este convexă dacă şi numai dacă domeniul $dom f$ este convex şi

$$\det \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ f(x) & f(y) & f(z) \end{bmatrix} \geq 0,$$

pentru orice $x, y, z \in dom f$ şi $x < y < z$.
Problema 11. Să se arate că maximul unei funcţii convexe peste poliedrul $\mathcal{P} = \operatorname{conv}\{v_1, \ldots, v_k\}$ este atins într-unul dintre vârfurile poliedrului, i.e.,

$$\sup_{x \in \mathcal{P}} f(x) = \max_{i=1,\ldots,k} f(v_i).$$

Indiciu: Se presupune că afirmaţia este falsă şi se utilizează inegalitatea lui Jensen.

Problema 12. Fie o funcţie convexă f şi o funcţie g definită după cum urmează:

$$g(x) = \min_{\alpha > 0} \frac{f(\alpha x)}{\alpha}.$$

(i) Să se demonstreze că funcţia g este omogenă, i.e. $g(tx) = tg(x)$ pentru orice $t \geq 0$.

(ii) Să se demonstreze că dacă o funcţie h este omogenă şi $h(x) \leq f(x)$ pentru orice x, atunci avem $h(x) \leq g(x)$ pentru orice x.

(iii) Să se demonstreze că funcţia g definită anterior este convexă.

Problema 13. Fie funcţia omogenă $f : \mathbb{R}^n \to \mathbb{R}$, i.e. $f(tx) = tf(x)$ pentru orice $t \geq 0$ şi $x \in \mathbb{R}^n$. O funcţie omogenă se numeşte subaditivă dacă satisface, în plus, următoarea relaţie:

$$f(x) + f(y) \geq f(x + y).$$

Să se demonstreze că, pentru funcţiile omogene, subaditivitatea este echivalentă cu convexitatea.

Problema 14. Fie o mulţime S nevidă, mărginită şi convexă în $\mathbb{R}^n$ şi o funcţie f definită de:

$$f(x) = \max_{y \in S} \; y^T x.$$

Funcţia f se numeşte *funcţia suport* a mulţimii S.

(i) Să se arate că funcţia f este omogenă şi convexă.

(i) Să se determine explicit funcţia suport pentru mulţimile S definite de

$$S = \{x \in \mathbb{R}^n : \|Ax\|_p \leq 1\} \quad p = 1, 2, \infty,$$

unde matricea A este inversabilă.

Problema 15. Fie funcţia convexă $f : \mathbb{R}^n \to \mathbb{R}$. Definim proprietatea de convexitate tare, enumerând condiţiile de ordin $0, I$ şi II (ce depind de gradul de diferenţiabilitate al funcţiei), într-un mod similar cu cele corespunzătoare cazului convex. Pentru $\mu > 0$, funcţia f se numeşte $\mu-$*tare convexă* în raport cu norma-p dacă:

(Condiţii de ordinul 0) funcţia f satisface următoarea inegalitate pentru orice $x, y \in \mathbb{R}^n$:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) - \frac{\mu\alpha(1-\alpha)}{2}\|x - y\|_p^2.$$

(Condiţii de ordinul I) funcţia f este diferenţiabilă şi satisface următoarea inegalitate pentru orice $x, y \in \mathbb{R}^n$:

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|x - y\|_p^2.$$

(Condiţii de ordinul II) funcţia f este diferenţiabilă de două ori şi satisface următoarea inegalitate pentru orice $x \in \mathbb{R}^n$:

$$x^T \nabla^2 f(x) x \geq \mu \|x\|_p^2,$$

unde $\|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}$, $1 \leq p \leq \infty$.

Să se demonstreze că:

(i) Pentru norma 2, condiţiile de ordin I implică condiţiile de ordin 0, i.e. dacă funcţia f satisface condiţiile de ordin I atunci ea satisface şi condiţiile de ordin 0.

(ii) Pentru norma 2, orice funcţie diferenţiabilă o dată şi $\mu-$tare convexă satisface relaţia:

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \mu \|x - y\|_2^2.$$

(iii) Funcţia $g(x) = \frac{1}{2}\|x\|_2^2$ este 1-tare convexă în raport cu norma $\|\cdot\|_2$.

(iv) Funcţia $g(x) = \frac{1}{2}\|x\|_p^2$ este $(p-1)$-tare convexă în raport cu norma $\|\cdot\|_p$, pentru $p > 1$.

(v) Funcţia $g(x) = \log n - \sum\limits_{i=1}^{n} x_i \log x_i$ este 1-tare convexă în raport cu norma $\|\cdot\|_1$ pe domeniul $\{x \in \mathbb{R}^n_+ : \|x\|_1 \leq 1\}$.

Problema 16. Fie funcţia f definită de:

$$f(y,t) = \max_{x\in\mathbb{R}^n} \langle y, x\rangle - \frac{t}{2}\|x\|^2.$$

Determinaţi domeniul funcţiei f şi arătaţi că expresia acesteia se rescrie în următoarea formă:

$$f(y,t) = \begin{cases} 0, & \text{dacă } y = 0, t = 0 \\ \frac{\|y\|^2}{2t}, & \text{dacă } t > 0. \end{cases}.$$

Problema 17. Să se demonstreze că pentru o matrice simetrică şi pozitiv definită $Q \in \mathbb{R}^{n\times n}$ şi doi vectori $u, v \in \mathbb{R}^n$ are loc următoarea inegalitate:

$$|u^T v| \leq u^T Q u v^T Q^{-1} v.$$

Problema 18. Fie matricele E, H şi F de dimensiuni compatibile şi $F^T F \preceq I$. Să demonstreze că pentru orice $\alpha > 0$, avem

$$EFH + H^T F^T E^T \preceq \alpha E E^T + \frac{1}{\alpha} H^T H.$$

Indiciu: Se foloseşte inegalitatea $(\alpha E^T - FH)^T(\alpha E^T - FH) \succeq 0$.

Problema 19. Fie matricea simetrică M de forma:

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

în care matricea C este inversabilă. Se numeşte *complementul Schur* al lui M, matricea $S = A - BC^{-1}B^T$. Un rezultat des folosit în teoria matriceală precizează că următoarele relaţii sunt echivalente:

(i) $M \succeq 0$ (M este pozitiv semidefinită);

(ii) $A \succeq 0,\ (I - AA^{\dagger})B = 0,\ C - B^T A^{\dagger} B \succeq 0$;

(iii) $C \succeq 0,\ (I - CC^{\dagger})B = 0,\ A - B^T C^{\dagger} B \succeq 0$,

unde cu $A^{\dagger}$ am notat pseudo-inversa matricei A. Mai mult, dacă $M \succ 0$ obţinem un caz particular al echivalenţelor:

(i)' $M \succ 0$ (M este pozitiv definită);

(ii)' $A \succ 0,\ C - B^T A^{-1} B \succeq 0$;

(iii)' $C \succ 0,\ A - B^T C^{-1} B \succeq 0$.

Evidenţiem, mai departe, un exemplu de aplicaţie al acestui rezultat. Fie funcţia $f : \mathbb{R}^n \times \mathcal{S}^n_{++} \to \mathbb{R}$ definită de:

$$f(x, Y) = x^T Y^{-1} x.$$

Să se demonstreze că f este convexă folosind proprietăţile complementului Schur.