Capitolul 5

Probleme de optimizare constrânsă

5.1 Preliminarii

Acest capitol delimitează începutul unei importante părţi din teoria optimizării, şi anume analiza problemelor de optimizare cu restricţii. O problemă de optimizare constrânsă (cu restricţii) în formă generală este dată prin următoarea expresie:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.l.:} \quad & g(x) \leq 0, \ h(x) = 0, \end{aligned} \tag{5.1}$$

unde funcţia obiectiv $f : \mathbb{R}^n \to \mathbb{R}$ şi funcţiile vectoriale ce definesc constrângerile $g : \mathbb{R}^n \to \mathbb{R}^m, h : \mathbb{R}^n \to \mathbb{R}^p$, se presupun a fi de două ori diferenţiabile.
Introducem, mai departe, o serie de noţiuni caracteristice problemelor de optimizare constrânse. *Lagrangianul* corespunzător problemei (5.1) este:

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

unde $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^p$ reprezintă *multiplicatorii Lagrange* asociaţi constrângerilor de inegalitate, respectiv de egalitate. Condiţiile necesare de ordin I pentru probleme constrânse de tip (5.1) se formulează astfel: dacă x^* este un punct de minim local al problemei (5.1), atunci există

$\lambda^* \in \mathbb{R}^m$ şi $\mu^* \in \mathbb{R}^p$ astfel încât:

$$\begin{aligned}\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla h_i(x^*) &= 0\\ g(x^*) \leq 0, h(x^*) &= 0\\ \lambda_i^* g_i(x^*) &= 0\\ \lambda^* &\geq 0.\end{aligned}$$

Aceste relaţii formează *sistemul KKT* sau *condiţiile de optimalitate Karush-Kuhn-Tucker*. În cazul nedegenerat, condiţiile suficiente de optimalitate se enunţă după cum urmează: dacă x^* satisface sistemul KKT iar Hessiana Lagrangianului dată de expresia:

$$\nabla_x^2 L(x^*, \lambda, \mu) = \nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla^2 h_i(x^*)$$

este pozitiv definită pe subspaţiul definit de:

$$M = \left\{ d: \ \nabla h(x^*)d = 0, \nabla g_j(x^*)^T d = 0 \ \forall j \in \mathcal{A}(x^*) \right\} \qquad (5.2)$$

unde $\mathcal{A}(x^*)$ este mulţimea constrângerilor active, atunci x^* este un punct de minim strict local.

5.2 Probleme rezolvate de laborator

În acest subcapitol vom utiliza mediul de programare Matlab pentru rezolvarea unui număr de exemple de probleme de optimizare constrânsă.

5.2.1 Formularea unei probleme de optimizare în formă standard

Exemplul 19. Un avion cargo are trei compartimente pentru transport mărfuri, pe care le notăm C_1, C_2 şi C_3. Presupunem că toate cele trei compartimente sunt limitate din punct de vedere al greutăţii şi volumului admis aşa cum se specifică în următorul tabel:

Compartiment	Limită greutate $[t]$	Limită volum $[m^3]$
C_1	10	6.8
C_2	16	8.6
C_3	8	5.3

Pentru a menţine echilibrul avionului, încărcătura fiecărui compartiment trebuie să respecte un raport prestabilit cu limita de greutate corespunzătoare. Mărfurile transportate sunt descrise, în plus, şi de profitul adus transportatorului. În cele ce urmează, precizăm pentru fiecare tip de mărfă, notat de la M_1 la M_4, datele corespunzătoare:

Tip marfă	Total disponibil $[t]$	Volum $[m^3/t]$	Profit $[\$/t]$
M_1	18	0.48	3100
M_2	15	0.65	3800
M_3	23	0.58	3500
M_4	12	0.39	2850

Ţinând cont că avionul ne permite să stocăm mai multe tipuri de marfă într-un singur compartiment, se doreşte determinarea cantităţilor corespunzătoare fiecarui tip de marfă astfel încât profitul pentru zbor să fie maxim.
(i) Să se formuleze această problemă ca una de programare liniară.
(ii) Să se rezolve problema liniară rezultată la punctul (i) cu ajutorul mediului Matlab utilizând funcţia `linprog` şi limbajul CVX.

Rezolvare. Evident, vectorul variabilelor de decizie pentru această problemă este dat de cantităţile de marfă ce trebuie alocate diferitelor compartimente. Din moment ce putem stoca mai multe tipuri de marfă într-un singur compartiment, avem posibilitatea de a stoca între 1 şi 4 tipuri de marfă în fiecare compartiment. Astfel, avem următoarea variabilă de decizie:

$$x = [x_{11}\ x_{12}\ x_{13}\ x_{14}\ \ \dots x_{34}\ x_{44}]^T = \{x_{ij}\}_{1\leq i\leq 3, 1\leq j\leq 4},$$

unde $x \in \mathbb{R}^{12}$, indicele $i = 1, \dots, 3$ reprezintă numărul compartimentului, iar $j = 1, \dots, 4$ reprezintă tipul de marfă. Funcţia obiectiv este liniară $f(x) = c^T x$ definită:

$$c = [3100\ 3800\ 3500\ 2850\ 3100\ 3800\ 3500\ 2850 \dots]^T.$$

Pe de altă parte, formularea constrângerilor în formă standard presupune:

- constrângerea cantităţii disponibile per tip de marfă (nu putem livra mai mult decât este disponibil). Constrângerile în acest caz au forma: $x_{11} + x_{21} + x_{31} \leq 18$ pentru primul tip de marfă, $x_{12} + x_{22} + x_{32} \leq 15$ pentru al doilea tip ş.a.m.d;
- limitarea masei totale dintr-un compartiment la o cantitate maximă

admisă, i.e. $x_{11} + x_{12} + x_{13} + x_{14} \leq 10$, $x_{21} + x_{22} + x_{23} + x_{24} \leq 16$ etc;
- volumul încărcăturilor este limitat la o anumită valoare în fiecare compartiment, i.e. $0.48x_{11} + 0.65x_{12} + 0.58x_{13} + 0.39x_{14} \leq 6.8$, $0.48x_{21} + 0.65x_{22} + 0.58x_{23} + 0.39x_{24} \leq 8.6$ etc;
- prin definiţia de pozitivitate a masei restricţionăm cantităţile din fiecare marfă la valori nenegative: $x \geq 0$;
- secţiunea din enunţ: "încărcăturile din compartimente trebuie să aibă acelaşi raport cu limitele lor de greutate" furnizează o serie de constrângeri de egalitate:

$$\frac{1}{10}\left(x_{11} + x_{12} + x_{13} + x_{14}\right) = \frac{1}{16}\left(x_{21} + x_{22} + x_{23} + x_{24}\right)$$
$$\frac{1}{16}\left(x_{21} + x_{22} + x_{23} + x_{24}\right) = \frac{1}{8}\left(x_{31} + x_{32} + x_{33} + x_{34}\right).$$

În final, formularea de programare liniară este definită astfel:

$$\begin{aligned} &\max_{x \in \mathbb{R}^{12}} c^T x \\ &\text{s.l.: } Ax = b, \quad Cx \leq d, \end{aligned}$$

unde parametrii $A \in \mathbb{R}^{2\times 12}$ şi $b \in \mathbb{R}^2$ cuprind constrângerile de egalitate, iar cele de inegalitate sunt cuprinse în $Cx \leq d$ unde $C \in \mathbb{R}^{22\times 12}$ şi $b \in \mathbb{R}^{22}$.
(ii) Pentru simplitate, furnizăm doar codul în limbaj CVX pentru rezolvarea acestei probleme, deoarece sintaxa rezolvării cu funcţia Matlab `linprog` este evidentă din apelarea comenzii `helplinprog` :

```
cvx_begin
variable x(12,1)
maximize ((c')*x)
subject to
A*x=b
C*x<=d
x>=0
cvx_end
```

5.2.2 Calcularea proiecţiei ortogonale a unui punct pe o mulţime convexă

Exemplul 20. Pentru fiecare din următoarele mulţimi convexe, să se calculeze proiecţia ortogonală a punctului asociat utilizând CVX:

(i) $\mathcal{B} = \{x \in \mathbb{R}^2 : \|x\| \leq 2\}$, $x_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$;

(ii) $\mathcal{E} = \{x \in \mathbb{R}^2 : (x-c)^T P^{-1}(x-c) \leq 1\}$, unde $P = \begin{bmatrix} 25 & 0 \\ 0 & 16 \end{bmatrix}$, $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $x_0 = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$;

(iii) $\mathcal{T} = \{x \in \mathbb{R}^3 : Ax \leq b\}$, unde $A = \begin{bmatrix} -I_3 \\ \mathbf{1}_3^T \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, $x_0 = \begin{bmatrix} 1.6 \\ 1.6 \\ 1.5 \end{bmatrix}$.

Să se figureze grafic mulţimile, punctele asociate şi proiecţiile ortogonale ale acestor puncte.

Rezolvare. În general, calculul proiecţiei ortogonale unui punct a_0 pe o mulţime X se realizează prin rezolvarea următoarei probleme de optimizare:

$$[a_0]_{(I_n,X)} = \arg\min_{x \in X} \|x - a_0\|^2. \tag{5.3}$$

(i) Secvenţa de cod CVX ce calculează proiecţia $[x_0]_{(I_n,\mathcal{B})}$ unui punct x_0 pe mulţimea $\mathcal{B} \subset \mathbb{R}^2$ (vezi Fig. 5.1) este dată în cele ce urmează:

```
x0=[3;1];
cvx_begin
 variable xB(2);
 minimize (xB'*xB-2*xB'*x0+x0'*x0)
 subject to
  norm(xB)<=2;
cvx_end
```

Observând că mulţimea $\mathcal{B}$ este de tip bilă cu rază 2 şi centrul în origine, reprezentarea grafică este realizată prin:

```
%Mai intai gasim mai multe puncte din bila:
[x1,x2]=meshgrid(-2:0.1:2,-2:0.1:2);
z=sqrt(x1.^2+x2.^2);
d=(z<=2); ball1=[]; ball2=[];
for i=1:size(d,1)
    for j=1:size(d,2);
        if(d(i,j)==1)
```

```
            ball1=[ball1; x1(i,j)];
            ball2=[ball2; x2(i,j)];
        end
    end
end
```

Mai departe, trasăm grafic toate elementele obţinute:

```
figure(1)
hold on
axis([-4 4 -4 4]); scatter(ball1,ball2,'+');
plot(x0(1),x0(2),'ro','LineWidth',2);
plot(xB(1),xB(2),'r+','LineWidth',2);
legend('Bila B','x_0','x_B');hold off;
```

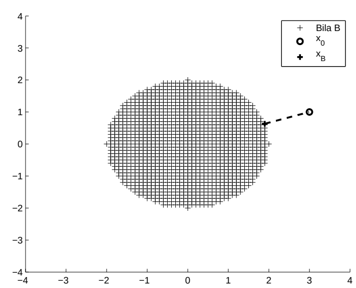

Figura 5.1: *Proiecţia unui punct exterior pe mulţime tip bilă.*

(ii) Observăm ca mulţimea $\mathcal{E} \subset \mathbb{R}^2$ este un elipsoid definit de matricea P şi centrat în punctul $(1, 1)$. Codul CVX pentru determinarea proiecţiei punctului x_0 pe mulţimea $\mathcal{E}$ (vezi Fig. 5.2), este redat de următoarea secvenţă:

```
Pinv=inv([25 0; 0 16]);
x0=[-4;-4]; xc=[1;1];
cvx_begin
 variable xE(2);
 minimize(xE'*xE-2*xE'*x0+x0'*x0)
 subject to
  (xE-xc)'*Pinv*(xE-xc)<=1;
cvx_end
```

Trasarea grafică a mulţimii $\mathcal{E}$ este realizată de secvenţa de cod:

```
[x1,x2]=meshgrid(-6:0.1:6,-6:0.1:6);
y=Pinv*xc; const=xc'*Pinv*xc;
z=x1.^2*Pinv(1,1)+x2.^2*Pinv(2,2)+2.*x1.*x2*Pinv(1,2)-...
2*(x1*y(1)+x2*y(2))+const;
d=(z<=1); ellipse1=[]; ellipse2=[];
for i=1:size(d,1)
    for j=1:size(d,2);
        if(d(i,j)==1)
            ellipse1=[ellipse1; x1(i,j)];
            ellipse2=[ellipse2; x2(i,j)];
        end
    end
end
figure(1)
hold on
axis([-7 7 -7 7]); scatter(ellipse1,ellipse2,'+');
plot(x0(1),x0(2),'ro','LineWidth',2);
plot(xE(1),xE(2),'r+','LineWidth',2);
legend('Elipsoidul E','x_0','x_E'); hold off;
```

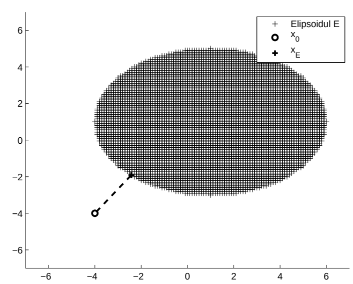

Figura 5.2: *Proiecţia unui punct exterior pe mulţime tip elipsoid.*

(iii) Calcularea proiecţiei unui punct x_0 pe mulţimea politopică $\mathcal{T}$ (vezi Fig. 5.3) utilizând CVX se realizează în mod asemănător:

```
A=[-eye(3); ones(1,3)]; b=[zeros(3,1);1]; x0=[1;1;1];
cvx_begin
 variable xT(3,1);
 minimize (xT'*xT-2*xT'*x0+x0'*x0)
 subject to
  A*xT<=b;
```

```
cvx_end
```

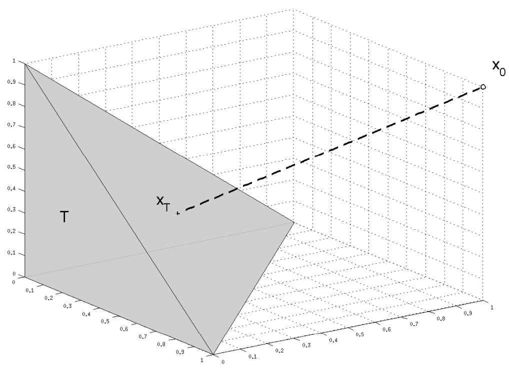

Figura 5.3: *Proiecţia unui punct exterior pe mulţime tip politop.*

Pentru a facilita reprezentarea grafică, observăm că mulţimea $\mathcal{T}$ poate fi reprezentată ca acoperirea convexă a vârfurilor (vertex) sale. Astfel, obţinându-se proiecţia x_T a punctului x_0 pe mulţimea $\mathcal{T}$, reprezentarea grafică se realizează prin următoarea secvenţă de cod:

```
x=[1; 0; 0; 0]; y=[0 ; 1; 0 ; 0]; z=[0 ; 0; 1; 0];
figure(1)
hold on
tetramesh(tesselation,[x(:) y(:) z(:)]);
x0hp = plot3(x0(1),x0(2),x0(3), 'ro');
set(x0hp, 'markersize', 10);
xThp = plot3(xT(1),xT(2),xT(3), 'r+');
plot3(points(1,:),points(2,:),points(3,:), 'r--');
set(xThp, 'markersize', 10); cameramenu; grid;
hold off
```

5.2.3 Metoda Gauss-Newton

Exemplul 21. Considerăm următoarea problemă de optimizare constrânsă:

$$\max_{x\in\mathbb{R}^n} \; -\sum_{i=1}^{n} x_i \, \ln(x_i)$$
$$\text{s.l.:} \; \sum_{i=1}^{n} x_i = 1, \quad \sum_{i=1}^{n} x_i c_i = m, \; i = 1, 2, \ldots, n,$$

unde c_i şi m sunt valori cunoscute.
Sistemul KKT al acestei probleme este definit de un sistem neliniar de ecuaţii $F(y) = 0$, unde vectorul y este compus din variabila primală x şi multiplicatorii Lagrange μ asociaţi constrângerilor de egalitate, i.e.

$$y = \begin{bmatrix} x \\ \mu \end{bmatrix}$$

Rezolvarea sistemului KKT poate fi abordată prin două metode:

(i) *metoda Gauss-Newton*;

(ii) *metoda Newton clasică*.

Să se rezolve problema anterioară în Matlab prin aplicarea celor două metode enumerate.

Rezolvare. Metoda *Gauss-Newton* este o metodă destinată rezolvării de probleme CMMP neliniare, i.e. de exemplu:

$$\min_{y\in\mathbb{R}^n} \frac{1}{2}\|F(y)\|^2, \tag{5.4}$$

a cărei soluţie rezolvă sistemul de ecuaţii $F(y) = 0$ în sens CMMP.
Procedura urmată de metoda Gauss-Newton presupune ca într-un punct dat y_k la iteraţia k, $F(y)$ este liniarizat:

$$F(y) \approx F(y_k) + J(y_k)(y - y_k),$$

unde $J(y)$ este Jacobianul lui $F(y)$. La următoarea iteraţie, y_{k+1}, poate fi determinată ca o soluţie a problemei liniare aproximative:

$$y_{k+1} \;=\; \arg\min_{y\in\mathbb{R}^n} \frac{1}{2}\|F(y_k) + J(y_k)(y - y_k)\|^2.$$

Pentru simplitatea expunerii, renotăm $J(y_k)$ cu J_k, iar în loc de $F(y_k)$ folosim F_k. Dacă presupunem că $J_k^T J_k$ este inversabilă, atunci rezultă iteraţia metodei Gauss-Newton definită de:

$$\begin{aligned} y_{k+1} &= \arg\min_{y\in\mathbb{R}^n} \frac{1}{2}\|F_k + J_k(y-y_k)\|^2 \\ &= y_k + \arg\min_{d\in\mathbb{R}^n} \frac{1}{2}\|F_k + J_k d\|^2 \\ &= y_k - (J_k^T J_k)^{-1} J_k^T F_k. \end{aligned}$$

Astfel, în cazul particular în care $x \in \mathbb{R}^3$, punctul iniţial este dat de $x^0 = [\frac{1}{9}\ \frac{3}{7}\ \frac{3}{10}]^T$, $m = 6$ şi $c = [3\ 6\ 9]^T$, scriem funcţia Matlab auxiliară cu sintaxa `[F,J]=Fgauss(y,c,m)` ce returnează în F şi J valoarea vectorului $F(y)$ şi a Jacobianului $J(y)$. Utilizăm criteriul de oprire $\|F(y)\| < \epsilon$, unde ϵ este toleranţă prestabilită. Astfel, codul Matlab corespunzător implementării metodei Gauss-Newton este:

```
function y=gauss_newt
x0=[1/9; 3/7; 3/10];
c=[3;6;9];mu1=0; mu2=0;
y=[x0;mu1;mu2]; eps=0.0001;m=6;
[F,J]=Fgauss(y,p,m);
while(norm(F)>eps)
    y=y-(J'*J)^(-1)*J'*F;
    [F,J]=Fgauss(y,p,m);
end
```

În Fig. 5.4 observăm descreşterea lui $\|F(y)\|$.

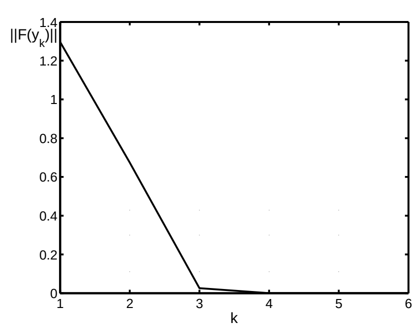

Figura 5.4: *Evoluţia $\|F(y)\|$ pentru metoda Gauss-Newton.*

Implementarea metodei Newton clasice o lăsăm ca exerciţiu cititorului.

5.2.4 Metoda gradientului proiectat

Exemplul 22. Fie problema de optimizare:

$$\begin{aligned}\min_{x\in\mathbb{R}^2}\ & (x_1-x_2)^2+6(x_1-x_2)^4+4x_1+8x_2\\ \text{s.l.: } & 4x_1+3x_2\le 5, \quad 2x_1+x_2\le 3,\\ & x_1+x_2=3.\end{aligned}$$

Să se rezolve problema în Matlab cu metoda gradientului proiectat.

Rezolvare. Metoda gradient proiectat presupune proiecţia ortogonală a unui punct auxiliar (definit de deplasarea în direcţia antigradientului) pe mulţimea constrângerilor notată cu X, i.e.

$$x_{k+1}=\left[x_k-\alpha_k\nabla f(x_k)\right]_{(I_n,X)},$$

unde α_k poate fi ales de exemplu prin metoda backtracking. Definiţia proiecţiei ortogonale a unui punct y pe o mulţime X, notată $[y]_{(I_n,X)}$ este dată de soluţia problemei:

$$\begin{aligned}[y]_{(I_n,X)}= \ & \arg\min_x \frac{1}{2}\left\|x-y\right\|^2\\ & \text{s.l.: } x\in X.\end{aligned} \tag{5.5}$$

În general, proiecţia pe mulţimea descrisă de:

$$X=\left\{x\in\mathbb{R}^n\colon Ax\le b, Cx=d\right\},$$

se realizează observând formularea echivalentă a problemei (5.5):

$$\begin{aligned}\min_x\ & \frac{1}{2}x^TI_nx-(x_k)^Tx\\ \text{s.l.: } & Ax\le b, \quad Cx=d,\end{aligned}$$

ce poate fi abordată, de exemplu, cu funcţia `quadprog` din Matlab.
În cazul particular prezentat, mulţimea constrângerilor este definită de:

$$X=\left\{x\in\mathbb{R}^2\colon 4x_1+3x_2\le 5, 2x_1+x_2\le 3, x_1+x_2=3\right\}$$

Creăm funcţiile auxiliare:

```
[f,g]=f_proj(x)    f=phi_proj(alpha,x,d),
```

unde prima funcţie returnează $f(x)$ şi $\nabla f(x)$, iar cea de-a doua returnează $f(x+\alpha d)$. Pentru criteriul de oprire, vom folosi $\|x_k - \bar{x}_k\| < \epsilon$, unde $\bar{x}_k$ este vectorul $x_k - \alpha_k \nabla f(x_k)$ proiectat pe mulţimea constrângerilor. Astfel, codul Matlab pentru implementarea metodei gradientului proiectat este:

```
function x=projected_gradient(x0,eps)
x=x0;
%Multimea constrangerilor este poliedru
%definit de matricele si vectorii
A=[4 3; 2 1]; b=[5;-3];
C=[1 1]; d=3.5; xbar=x0+1; alpha2=0.5;
while (norm(xbar-x)>=eps)
    [f,g]=f_proj(x);
    alpha=fminsearch(@(alpha) phi_proj(alpha,x,-g),1);
    grad_step=x-alpha*g;
    xbar=quadprog(2*eye(2),-2*grad_step,A,b,C,d);
    x=x+alpha2*(xbar-x);
end
```

În Fig. 5.5 observăm punctele obţinute de metoda gradientului proiectat, pornind din $x_0 = [-5 \quad 2]$.

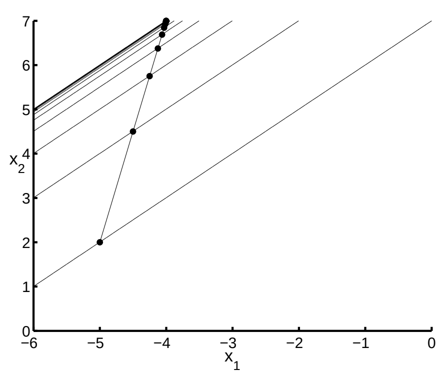

Figura 5.5: *Puncte obţinute de metoda gradientului proiectat.*

5.3 Probleme rezolvate de seminar

5.3.1 Problema duală, condiţii şi puncte Karush-Kuhn- Tucker

Problema 1. Fie problema de optimizare:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} \ \|Ax\|_\infty \\ &\quad \text{s.l: } \|x\|_1 \leq 1, \ e^{x_1+x_2} = 1, \ (a^T x - b)^2 \leq 1, \end{aligned}$$

unde $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $a = [3\ 4]^T$ şi $b = 1$. Să se scrie această problemă de optimizare în formă LP explicită.

Rezolvare. Dacă considerăm matricea $A = \begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix}$ şi definiţia normei infinit, atunci funcţia obiectiv se poate scrie drept:

$$\|Ax\|_\infty = \left\| \begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} x \right\|_\infty = \max_x \left\{ |a_1^T x|, |a_2^T x| \right\}.$$

Introducem o variabilă artificială t pentru a muta funcţia obiectiv în constrângeri, obţinându-se astfel o problemă echivalentă cu cea originală:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2, t\in\mathbb{R}} \quad t \\ &\quad \text{s.l: } |a_1^T x| \leq t, \ |a_2^T x| \leq t, \\ &\qquad \|x\|_1 \leq 1, \ e^{x_1+x_2} = 1, \ (a^T x - b)^2 \leq 1. \end{aligned}$$

Din definiţia modulului, observăm că putem scrie primele două constrângeri drept:

$$\begin{cases} -t \leq a_1^T x \leq t \\ -t \leq a_2^T x \leq t \end{cases} \Rightarrow \begin{cases} a_1^T x \leq t \\ a_2^T x \leq t \\ -a_1^T x \leq t \\ -a_2^T x \leq t \end{cases} \Rightarrow \begin{bmatrix} A \\ -A \end{bmatrix} x \leq t\mathbf{1}_4$$

unde $\mathbf{1}_4$ este vectorul de dimensiune 4 cu toate elementele egale cu 1. Dacă introducem două variabile adiţionale s_1, s_2, şi ne reamintim definiţia normei 1, putem atunci rescrie constrângerea $\|x\|_1 \leq 1$ drept:

$$\|x\|_1 \leq 1 \Rightarrow |x_1| + |x_2| \leq 1 \Rightarrow \begin{cases} s_1 + s_2 \leq 1 \\ |x_1| \leq s_1 \\ |x_2| \leq s_2 \end{cases}$$

Observăm acum faptul că egalitatea $e^{x_1+x_2} = 1$ se reduce simplu la $x_1 + x_2 = 0$, iar constrângerea $(a^T x - b)^2 \leq 1$ se reduce la:

$$-1 \leq a^T x - b \leq 1 \Rightarrow \begin{cases} a^T x - b \leq 1 \\ -a^T x + b \leq 1 \end{cases}$$

Luând astfel o variabilă de decizie $z = [t\ x_1\ x_2\ s_1\ s_2]^T$, problema noastră originală se poate rescrie ca un LP:

$$\begin{aligned} \min_{z \in \mathbb{R}^5}\ & c^T z \\ \text{s.l } & A_z z = b_z \\ & C_z z \leq d_z \end{aligned}$$

unde $c = [1\ 0\ 0\ 0\ 0]^T$, $A_z = [0\ 1\ 1\ 0\ 0]$ şi $b_z = 0$, iar

$$C_z = \begin{bmatrix} -1 & 2 & 1 & 0 & 0 \\ -1 & 1 & 2 & 0 & 0 \\ -1 & -2 & -1 & 0 & 0 \\ -1 & -1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 3 & 4 & 0 & 0 \\ 0 & -3 & -4 & 0 & 0 \end{bmatrix} \text{ şi } d_z = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{bmatrix}.$$

Problema 2. Fie problema de optimizare următoare:

$$\begin{aligned} \min_{x \in \mathbb{R}^n}\ & \|x\|_1 \\ \text{s.l.: } & Ax = b. \end{aligned}$$

(i) Arătaţi că această problemă se poate reformula ca una de programare liniară.
(ii) Aduceţi problema la forma de programare liniară (LP) *standard*.
(iii) Determinaţi problema duală corespunzătoare.

Rezolvare. (i) Utilizând definiţia normei 1 obţinem reformularea:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & \sum_{i=1}^n |x_i| \\ \text{s.l.: } & Ax = b. \end{aligned}$$

Mai departe, folosim procedura de reformulare a unei probleme de optimizare generale într-o problemă cu funcţie obiectiv liniară, coborând în constrângeri componentele sumei:

$$\begin{aligned} \min_{x, t \in \mathbb{R}^n} \; & \sum_{i=1}^n t_i \\ \text{s.l.: } & |x_i| \leq t_i, \; i = 1, \dots, n, \\ & Ax = b. \end{aligned}$$

Din definiţia modulului rezultă echivalenţa inegalităţii $|y| \leq t$ cu $-t \leq y \leq t$. Aplicând în problema noastră această transformare obţinem o problemă de programare liniară:

$$\begin{aligned} \min_{x, t \in \mathbb{R}^n} \; & \sum_{i=1}^n t_i \\ \text{s.l.: } & x_i \leq t_i, \; i = 1, \dots, n, \\ & -x_i \leq t_i, \; i = 1, \dots, n, \\ & Ax = b, \end{aligned}$$

(ii) Forma standard a problemelor de programare liniară este definită de:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & c^T x \\ \text{s.l.: } & Ax = b, \\ & x \geq 0. \end{aligned}$$

Pentru a aduce problema de la punctul (i) la forma standard, introducem un set de variabile "virtuale", rescriind fiecare variabilă de decizie din problema de optimizare iniţială:

$$x_i = x_i^+ - x_i^-, \text{ unde } x_i^+ \geq 0, x_i^- \geq 0.$$

În plus, reformulăm inegalităţile în următorul mod:

$$
\begin{aligned}
t_i - x_i &= s_{2i-1}, \ \text{unde } s_{2i-1} \geq 0, \\
t_i + x_i &= s_{2i}, \quad \text{unde } s_{2i} \geq 0,
\end{aligned}
$$

$i = 1, \ldots, n$.
Cu ajutorul acestei reformulări obţinem o problemă de programare liniară în formă standard:

$$
\begin{aligned}
\min_{x^+, x^-, t \in \mathbb{R}^n, s \in \mathbb{R}^{2n}} \ & \sum_{i=1}^{n} t_i \\
\text{s.l.: } & t_i - x_i = s_{2i-1}, \\
& t_i + x_i = s_{2i}, \ i = 1, \ldots, n, \\
& A(x^+ - x^-) = b, \\
& s \geq 0, x^+ \geq 0, x^- \geq 0, t \geq 0.
\end{aligned}
$$

unde $x^+ = [x_1^+ \ldots x_n^+]^T, x^- = [x_1^- \ldots x_n^-]^T$.
Mai compact avem:

$$
\begin{aligned}
\min_{z \in \mathbb{R}^{5n}} \ & c^T z \\
\text{s.l.: } & Dz = d \\
& z \geq 0,
\end{aligned}
$$

unde $z = \begin{bmatrix} x^+ \\ x^- \\ t \\ s \end{bmatrix}, D = \begin{bmatrix} M \\ N \end{bmatrix}, M \in \mathbb{R}^{m \times 5n}, N \in \mathbb{R}^{2n \times 5n}$.

(iii) Pentru a determina problema duală, calculăm Lagrangianul asociat problemei primale:

$$
\begin{aligned}
\mathcal{L}(x, t, \lambda, \mu^+, \mu^-) &= \sum_{i=1}^{n} t_i + \lambda^T(Ax - b) + {\mu^+}^T(t - x) - {\mu^-}^T(t + x) \\
&= (e + \mu^+ - \mu^-)^T t - {\mu^+}^T x - {\mu^-}^T x + \lambda^T(Ax - b) \\
&= (e + \mu^+ - \mu^-)^T t + \left(A^T\lambda - \mu^+ - \mu^-\right)^T x - \lambda^T b,
\end{aligned}
$$

unde vectorul e are componentele egale cu 1. Obţinem funcţia duală prin minimizarea Lagrangianului în raport cu variabilele primale:

$$
q(\lambda, \mu^+, \mu^-) = \min_{x, t \in \mathbb{R}^n} (e + \mu^+ - \mu^-)^T t + \left(A^T\lambda - \mu^+ - \mu^-\right)^T x - \lambda^T b.
$$

Observând că $q(\lambda, \mu^+, \mu^-)$ ia valori finite doar în cazul în care

$$\begin{aligned} &e + \mu^+ - \mu^- = 0 \\ &A^T\lambda - \mu^+ - \mu^- = 0, \end{aligned}$$

aceste relaţii au rolul de constrângeri în componenţa problemei duale. În concluzie, problema duală asociată problemei din enunţ este dată de:

$$\begin{aligned} \max_{\lambda,\mu^+,\mu^-} \quad & -\lambda^T b \\ & e + \mu^+ - \mu^- = 0 \\ & A^T\lambda - \mu^+ - \mu^- = 0 \\ & \mu^+ \geq 0, \mu^- \geq 0. \end{aligned}$$

Remarca 3. Observaţi că problema din enunţ reprezintă proiecţia ortogonală a punctului 0 pe subspaţiul $Ax = b$, în raport cu norma 1.

Problema 3. Fie problema de programare liniară:

$$\begin{aligned} \min_{x_1,x_2} \; & 50x_1 + 80x_2 \\ \text{s.l.: } & 3x_1 \geq 6, \quad 2x_1 + 4x_2 \geq 10, \\ & 2x_1 + 5x_2 \geq 8, \quad x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

Determinaţi problema duală corespunzătoare.

Rezolvare. Pentru o expunere simplă notăm funcţia obiectiv şi constrângerile problemei precedente după cum urmează:

$$\begin{aligned} \min_{x_1,x_2} \quad & c^T x \\ \text{s.l.: } & Ax \leq b, \quad x \geq 0, \end{aligned}$$

unde $c = \begin{bmatrix} 50 \\ 80 \end{bmatrix}$, $A = \begin{bmatrix} -3 & 0 \\ -2 & -4 \\ -2 & -5 \end{bmatrix}$, $b = \begin{bmatrix} -6 \\ -10 \\ -8 \end{bmatrix}$. În continuare, observăm că Lagrangianul problemei este dat de relaţia:

$$\mathcal{L}(x, \lambda_1, \lambda_2) = c^T x + \lambda_1^T (Ax - b) - \lambda_2^T x,$$

unde $\lambda_1 \in \mathbb{R}^3$, $\lambda_1 = [\lambda_1^1 \quad \lambda_1^2 \quad \lambda_1^3]^T$, iar $\lambda_2 \in \mathbb{R}^2$.

Funcţia duală $q(\lambda_1, \lambda_2) = \min_{x \in \mathbb{R}^2} \left(c + A^T\lambda_1 - \lambda_2\right) x - \lambda_1^T b$ ia valori finite atunci când $c + A^T\lambda_1 - \lambda_2 = 0$. De aceea, impunând această constrângere variabilelor duale, rezultă problema duală:

$$\begin{aligned} \max_{\lambda_1 \in \mathbb{R}^3, \lambda_2 \in \mathbb{R}^2} \quad & -\lambda_1^T b \\ \text{s.l.: } & c + A^T\lambda_1 - \lambda_2 = 0, \\ & \lambda_1 \geq 0, \lambda_2 \geq 0, \end{aligned}$$

sau echivalent,

$$\begin{aligned} \max_{\lambda_1 \in \mathbb{R}^3, \lambda_2 \in \mathbb{R}^2} \quad & -\lambda_1^T b \\ \text{s.l.: } & c + A^T\lambda_1 \geq 0, \\ & \lambda_2 \geq 0. \end{aligned}$$

Mai exact, substituind valorile parametrilor obţinem:

$$\begin{aligned} \max_{\lambda_1 \in \mathbb{R}^3, \lambda_2 \in \mathbb{R}^2} \quad & 6\lambda_1^1 + 10\lambda_1^2 + 8\lambda_1^3 \\ \text{s.l.: } & 3\lambda_1^1 + 2\lambda_1^2 + 2\lambda_1^3 \leq 50, \\ & 4\lambda_1^2 + 5\lambda_1^3 \leq 80, \\ & \lambda_2 \geq 0. \end{aligned}$$

Problema 4. Numeroase probleme din statistică, învăţare automată, recunoaştere de imagini/feţe, presupun clasificarea de obiecte sau imagini prin determinarea unui model matematic capabil să recunoască şi să încadreze într-o clasă o instanţă nouă de obiect/imagine. Una dintre cele mai renumite tehnici de recunoaştere/clasificare este SVM (*Support Vector Machine*), ce se formulează sub forma unei probleme pătratice de optimizare:

$$\begin{aligned} \min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \quad & \frac{1}{2}\|a\|^2 \\ \text{s.l.: } & c_i \left(a^T y_i - b\right) \geq 1, \ i = 1, \dots, m. \end{aligned}$$

Variabilele de decizie a, b reprezintă parametrii unui hiperplan de separare a claselor de obiecte/imagini, aşa cum se observă în Fig. 5.6 (vezi [1]). Să se determine problema duală.

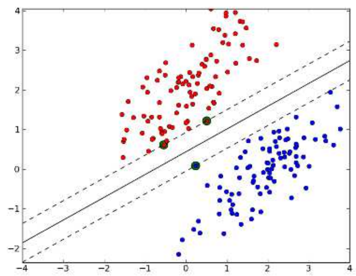

Figura 5.6: *Hiperplan de separare a două clase de obiecte.*

Rezolvare. Lagrangianul problemei este dat de relaţia:

$$\mathcal{L}(x,\lambda) = \frac{1}{2}\|a\|^2 - \sum_{i=1}^{n} \lambda_i \left[c_i(a^T y_i - b) - 1\right].$$

Forma explicită a funcţiei duale:

$$q(\lambda) = \min_{a\in\mathbb{R}^n, b\in\mathbb{R}} \frac{1}{2}\|a\|^2 - \sum_{i=1}^{n} \lambda_i \left[c_i(a^T y_i - b) - 1\right],$$

este definită de condiţiile de optimalitate de ordin I:

$$a^* - \sum_{i=1}^{n} c_i\lambda_i y_i = 0, \qquad \sum_{i=1}^{n} c_i\lambda_i = 0.$$

Ambele relaţii se obţin din derivarea Lagrangianului în raport cu a, respectiv cu b, anularea expresiei rezultate şi egalarea cu 0. Deoarece a doua egalitate asigură că funcţia duală ia valori finite, aceasta devine o constrângere necesară a problemei duale. De aceea, prin substituirea valorii optime a^* în expresia Lagrangianului avem:

$$\begin{aligned}
q(\lambda) &= \frac{1}{2}\|\sum_{i=1}^{n} c_i\lambda_i y_i\|^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i c_i c_j \lambda_j (y_i^T y_j) + \sum_{i=1}^{n} \lambda_i \\
&= \frac{1}{2}\|\sum_{i=1}^{n} c_i\lambda_i y_i\|^2 - \|\sum_{i=1}^{n} c_i\lambda_i y_i\|^2 + \sum_{i=1}^{n} \lambda_i \\
&= -\frac{1}{2}\|\sum_{i=1}^{n} c_i\lambda_i y_i\|^2 + \sum_{i=1}^{n} \lambda_i,
\end{aligned}$$

putem obţine problema duală asociată, care observăm că este concavă:

$$\max_{\lambda\in\mathbb{R}^m} \; -\frac{1}{2}\|\sum_{i=1}^{n} c_i\lambda_i y_i\|^2 + \sum_{i=1}^{n}\lambda_i$$
$$\text{s.l.:} \; \sum_{i=1}^{n} c_i\lambda_i = 0, \;\; \lambda_i \geq 0, \forall i = 1,\ldots,n.$$

Utilizând notaţiile $\lambda = [\lambda_1 \ldots \lambda_m]^T, c = [c_1 \ldots c_m]^T, e = [1 \ldots 1]^T$ şi observând forma pătratică a problemei duale obţinem reformularea:

$$\max_{\lambda\in\mathbb{R}^m} \; -\frac{1}{2}\lambda^T Q\lambda - e^T\lambda$$
$$\text{s.l.:} \; c^T\lambda = 0, \quad \lambda \geq 0,$$

unde matricea $Q \in \mathbb{R}^n, Q = Q^T$, are elementele $Q_{ij} = c_i c_j a_i^T a_j$.

Problema 5. Fie problema celor mai mici pătrate (CMMP), ce presupune calculul unei soluţii corespunzătoare unui sistem liniar de ecuaţii subdeterminat. Se observă uşor că problema se reduce la determinarea unui vector de normă minimă ce satisface sistemul liniar studiat:

$$\min_{x\in\mathbb{R}^n} \; \frac{1}{2}\|x\|^2$$
$$\text{s.l.:} \; Ax = b,$$

unde $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. Să se calculeze problema duală a problemei de minimizare.

Rezolvare. Deoarece problema de minimizare prezintă doar constrângeri de tip egalitate, definim Lagrangianul:

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

Astfel, funcţia duală este dată de soluţia problemei următoare:

$$q(\mu) = \min_{x\in\mathbb{R}^n} \; \frac{1}{2}x^T x + \mu^T(Ax - b).$$

Se obţine uşor soluţia $x^* = -A^T\mu$. Substituind în problema precedentă, obţinem funcţia duală pătratică concavă:

$$q(\mu) = -\frac{1}{2}\mu^T AA^T\mu - \mu^T b.$$

În concluzie, problema duală poate fi formulată ca o problemă pătratică convexă:

$$q^* = \max_{\mu \in \mathbb{R}^m} -\frac{1}{2}\mu^T AA^T \mu - \mu^T b.$$

Mai mult decât atât, din condiţiile de optimalitate ale problemei duale avem: $\nabla q(\mu) = -AA^T\mu - b = 0$. În cazul în care $\text{rang}(A) = m$, obţinem $\mu^* = -(AA^T)^{-1}b$ şi $x^* = A^T(AA^T)^{-1}b$, cu remarca importantă că x^* este soluţia clasică CMMP.

Problema 6. Fie problema de optimizare:

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \frac{1}{2}x^T x \\ &\text{s.l.: } Ax \leq b, \end{aligned}$$

unde $A \in \mathbb{R}^{m \times n}$, iar $\text{rang}(A) = m$. Să se determine problema duală şi punctul de optim.

Rezolvare. Pentru determinarea problemei duale calculăm Lagrangianul problemei primale:

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

De asemenea, funcţia duală este dată de:

$$q(\lambda) = \min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda) = \min_{x \in \mathbb{R}^n} \frac{1}{2}x^T x + \lambda^T(Ax - b).$$

Deoarece funcţia obiectiv a acestei probleme este convexă, condiţiile de optimalitate de ordin I sunt suficiente pentru determinarea optimului $x^*(\lambda) = -A^T\lambda$, iar prin substituţia în $q(\lambda)$ obţinem:

$$\begin{aligned} q(\lambda) &= \frac{1}{2}\lambda^T AA^T\lambda + \lambda^T(A(-A^T\lambda) - b) \\ &= -\frac{1}{2}\lambda^T AA^T\lambda - \lambda^T b. \end{aligned}$$

În final, problema duală este definită de:

$$q^* = \max_{\lambda \geq 0} q(\lambda) \qquad \left(:= -\frac{1}{2}\lambda^T AA^T\lambda - \lambda^T b\right). \tag{5.6}$$

Dacă matricea A are rangul maxim (implicit AA^T inversabilă şi pozitiv definită), observăm că soluţia problemei (5.6) este dată de:

$$\lambda^* = \arg\min_{\lambda \geq 0} \frac{1}{2}\|A^T\lambda + A^T(AA^T)^{-1}b\|_2^2 = \arg\min_{\lambda \geq 0} \frac{1}{2}\|A^T(\lambda + (AA^T)^{-1}b)\|_2^2.$$

În concluzie, observăm un compromis între problema primală şi cea duală: în cazul problemei primale funcţia obiectiv are o formă foarte simplă, iar constrângerile reprezintă dificultatea centrală; în cazul problemei duale, funcţia obiectiv este relativ complicată, iar constrângerile au formă simplă.

Problema 7. Fie problema de optimizare constrânsă:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & c^T x - \sum_{i=1}^n x_i \ln x_i \\ \text{s.l.} \; & \sum_{i=1}^n x_i = 1, \quad x \geq 0, \end{aligned}$$

unde $c \in \mathbb{R}^n$. (i) Să se determine dacă problema este convexă.
(ii) Să se determine un punct KKT pentru această problemă.

Rezolvare. (i) Pentru a determina convexitatea problemei, observăm în primul rând că toate constrângerile sunt liniare, deci mulţimea constrângerilor este convexă. În cazul funcţiei obiectiv însă, calculăm matricea Hessiană:

$$\nabla^2 f(x) = \begin{bmatrix} -\frac{1}{x_1} & 0 & \dots & 0 \\ 0 & -\frac{1}{x_2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & -\frac{1}{x_n} \end{bmatrix}$$

şi observăm că este negativ semidefinită pe tot domeniul de definiţie al funcţiei obiectiv $\mathrm{dom} f = \mathbb{R}^n_+$. În concluzie, funcţia obiectiv este concavă, rezultând o problemă de optimizare neconvexă.

(ii) Indicăm o procedură relativ inedită pentru determinarea unui punct KKT. Şi anume, folosim structura problemei observând condiţiile de optimalitate de ordin I:

$$c_i - \ln x_i - 1 = 0, \forall i = 1, \dots, n.$$

Aceste condiţii furnizează punctul de optim neconstrâns definit de componentele $x_i^* = e^{c_i - 1}$.
Dacă proiectăm punctul de optim neconstrâns pe hiperplanul:

$$H = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1 \right\}$$

obţinem vectorul y^* definit de componentele:

$$y_i^* = \frac{e^{c_i - 1}}{\sum_{j=1}^n e^{c_j - 1}} \geq 0, \quad \sum_i y_i^* = 1.$$

În final, verificăm dacă y^* obţinut satisface următorul sistem KKT:

$$\begin{cases} c_i - \ln y_i^* - 1 + \lambda_i^* + \mu^* = 0, \\ \sum\limits_{i=1}^n y_i^* = 1, \; y_i^* \geq 0, \\ \lambda_i^* \geq 0, \\ \lambda_i^* y_i^* = 0, i = 1, \ldots, n. \end{cases}$$

Observând că $y_i^* > 0$ atunci condiţiile de complementaritate impun $\lambda_i^* = 0$ pentru orice $i = 1 \ldots n$. De asemenea, condiţia de optimalitate $c_i - (c_i - 1) + \ln \sum_{j=1}^n e^{c_j} - 1 + \mu = 0$ este satisfăcută dacă impunem $\mu = -\ln \sum_j e^{c_j}$. În aceste condiţii, y^* satisface condiţiile KKT, fiind un punct staţionar pentru problema din enunţ.

Problema 8. Fie problema de minimizare:

$$\begin{aligned} &\min_{x_1, x_2 \in \mathbb{R}} \; x_2 \\ &\text{s.l.:} \quad 1 - x_1^2 - x_2^2 \geq 0. \end{aligned}$$

(i) Determinaţi punctele KKT pentru această problemă.

(ii) Găsiţi punctele de minim strict locale.

Rezolvare. (i) Funcţia Lagrange corespunzătoare funcţiei obiectiv are următoarea formă:

$$\mathcal{L}(x, \mu) = x_2 + \lambda(x_1^2 + x_2^2 - 1).$$

Derivând Lagrangianul în funcţie de variabilele x_1 şi x_2 rezultă următorul sistem KKT asociat problemei:

$$\begin{cases} 2\lambda x_1 = 0 \\ 1 + 2\lambda x_2 = 0 \\ x_1^2 + x_2^2 - 1 \leq 0 \\ \lambda \geq 0 \\ \lambda(x_1^2 + x_2^2 - 1) = 0. \end{cases}$$

Prima relaţie $\lambda x_1 = 0$ impune $x_1 = 0$ sau $\lambda = 0$. Constatăm uşor că pentru $\lambda = 0$, sistemul nu are soluţie. De aceea, din $x_1 = 0$ avem $x_2 = -\frac{1}{2\lambda}$ şi substituind în constrângerea de tip inegalitate obţinem:

$$\frac{1}{4\lambda^2} - 1 \leq 0 \Rightarrow \lambda^2 \geq \frac{1}{4}.$$

Din moment ce $\lambda \geq 0$, avem $\lambda \geq \frac{1}{2}$. Iar din relaţia de complementaritate avem:

$$\frac{\lambda - 4\lambda^2}{4\lambda^2} = 0 \Rightarrow \lambda = \frac{1}{2}.$$

În concluzie, $x_2 = -1$, iar punctele KKT în acest caz sunt definite de: $x^\star = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$şi $\lambda^* = \frac{1}{2}$.

(ii) Pentru a arăta că punctul KKT găsit este punct de minim, calculăm mai întâi planul tangent:

$$M = \left\{d \colon \nabla g_j(x^*)^T d = 0 \ \forall j \in \mathcal{A}(x^*)\right\}$$

unde $\mathcal{A}(x^*) = \{j : g_j(x^*) = 0\}$. Evident pentru $x^* = [0 \quad -1]^T$, constrângerea noastră $g(x^*) = (x_1^*)^2 + (x_2^*)^2 - 1 = 0$, deci este activă, iar $\lambda^* = \frac{1}{2} \geq 0$. Gradientul funcţiei de constrângere este:

$$\nabla g(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \tag{5.7}$$

iar $\nabla g(x^*) = [0 \ -2]^T$. Astfel, planul tangent este:

$$M = \left\{d \in \mathbb{R}^2 \colon d_2 = 0\right\}$$

Pentru ca x^* să fie punct de minim strict, atunci $\nabla_x^2 L(x^\star, \lambda^*)$ trebuie să fie pozitiv definită pe M. Printr-un calcul simplu, observăm că:

$$\nabla_x^2 L(x^\star, \lambda^*) = \begin{bmatrix} 2\lambda^* & 0 \\ 0 & 2\lambda^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Deci pentru $d \in M$ avem:

$$d^T \nabla_x^2 L(x^\star, \lambda^*) d = d_1^2 > 0$$

şi concluzionăm că $x^* = [0 \ -1]^T$ este punct de minim strict local.

Problema 9. Fie problema de optimizare constrânsă:

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

(i) Să se arate că problema este convexă.
(ii) Să se rezolve sistemul KKT asociat problemei. Verificaţi dacă punctul obţinut este punct de minim global.

Rezolvare. (i) Funcţia obiectiv se formulează matriceal $f(x) = \frac{1}{2}x^T I_2 x$. Observăm că Hessiana acesteia este pozitiv definită, deci funcţia obiectiv este convexă. Pe de altă parte, constrângerea $(x_1 + x_2 - 1)^2 = 0$ este evident echivalentă cu egalitatea liniară $x_1 + x_2 - 1 = 0$. Deoarece singura constrângere este liniară, avem că mulţimea fezabilă este convexă.
(ii) Derivând Lagrangianul asociat problemei:

$$\mathcal{L}(x, \mu) = \frac{1}{2}(x_1^2 + x_2^2) + \mu(x_1 + x_2 - 1)$$

obţinem sistemul KKT:

$$\begin{cases} x_1 + \mu = 0 \\ x_2 + \mu = 0 \\ x_1 + x_2 - 1 = 0 \end{cases}$$

din care rezultă $\mu^* = -\frac{1}{2}$ şi $x_1^* = x_2^* = \frac{1}{2}$. Pentru a verifica dacă x^* este punct de minim strict local, atunci calculăm mai întâi planul tangent

$$M = \left\{ d \in \mathbb{R}^2 : \ \nabla h(x^*) d = 0 \right\} = \left\{ d \in \mathbb{R}^2 : \ d_1 + d_2 = 0 \right\}. \qquad (5.8)$$

Remarcând că:

$$d^T \nabla^2 f(x^*) d = d_1^2 + d_2^2 > 0, \quad \forall d \in M,$$

concluzionăm că x^* este punct de minim global strict.

Problema 10. Fie următoarea problemă:

$$\begin{aligned} \min_{x\in\mathbb{R}^3} f(x)\colon &= \frac{1}{2}(x_1^2+x_2^2+x_3^2) \\ \text{s.l } g_1(x)\colon &= x_1+x_2+x_3 \le 0 \\ g_2(x)\colon &= x_1 \le 0. \end{aligned}$$

Să se găsească punctele KKT asociate.

Rezolvare. Observăm că orice punct fezabil pentru această problemă este un *punct regulat*, iar din condiţiile ale sistemului KKT obţinem:

Optimalitate	*Fezabilitate*	*Complementaritate*
$x_1^*+\lambda_1^*+\lambda_2^*=0,$	$x_1^*+x_2^*+x_3^*\le 0,\ x_1^*\le 0,$	$\lambda_1^*(x_1^*+x_2^*+x_3^*)=0,$
$x_2^*+\lambda_1^*=0,$	$\lambda_1^*\ge 0,\ \lambda_2^*\ge 0,$	$\lambda_2^*x_1^*=0.$
$x_3^*+\lambda_1^*=0,$		

Din condiţiile de optimalitate avem $x_1^*=-\lambda_1^*-\lambda_2^*, x_2^*=-\lambda_1^*, x_3^*=-\lambda_2^*$. Substituind aceste expresii în condiţiile de complementaritate rezultă sistemul:

$$\begin{aligned} 3(\lambda_1^*)^2+\lambda_1^*\lambda_2^*=0, \\ (\lambda_2^*)^2+\lambda_1^*\lambda_2^*=0, \end{aligned}$$

ce elimină posibilităţile ca $\lambda_1^*=0,\lambda_2^*>0$ sau $\lambda_1^*>0,\lambda_2^*=0$. De aceea, rămâne să verificăm cazurile când $\lambda_1^*>0,\lambda_2^*>0$ sau $\lambda_1^*=0,\lambda_2^*=0$. Primul din cazurile rămase nu este posibil deoarece $x_1^*=-\lambda_1^*-\lambda_2^*$ şi a doua condiţie de complementaritate nu este satisfăcută. În concluzie, doar punctul nul satisface condiţiile KKT.

Problema 11. Fie problema de optimizare constrânsă:

$$\begin{aligned} \min_{x\in\mathbb{R}^2} x_1^2-x_2^2 \\ x_1+x_2\ge 0. \end{aligned}$$

(i) Să se determine punctele KKT;
(ii) Aflaţi funcţia duală şi problema duală.

Rezolvare. (i) Lagrangianul acestei probleme este definit de:

$$\mathcal{L}(x,\lambda)=x_1^2-x_2^2-\lambda(x_1+x_2),$$

deci sistemul KKT este dat de:

$$\begin{cases} \nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \\ -x_1^* - x_2^* \leq 0, \\ \lambda^* \geq 0, \\ -\lambda^*(x_1^* + x_2^*) = 0. \end{cases}$$

Din prima ecuaţie a sistemului KKT rezultă:

$$x_1^* = \frac{\lambda^*}{2}, \ x_2^* = -\frac{\lambda^*}{2}.$$

Înlocuind în ultima ecuaţie şi remarcând că fezabilitatea primală este satisfăcută, deducem că λ^* poate lua o infinitate de valori.

(ii) Remarcăm o expresie neobişnuită a funcţiei duale:

$$q(\lambda) = \min_{x \in \mathbb{R}^2} L(x, \lambda) = \min_{x \in \mathbb{R}^2} x_1^2 - x_2^2 - \lambda(x_1 + x_2) = 0,$$

de aceea constatăm că problema duală nu este o problemă de optimizare propriu-zisă: $\max\limits_{\lambda \geq 0} 0$.

5.4 Probleme propuse

Problema 1. Să se arate că $x^* = [1 \ 1/2 \ -1]^T$ este punctul minim global pentru următoarea problemă de optimizare:

$$\begin{aligned} &\min_{x \in \mathbb{R}^3} \frac{1}{2} x^T P x + q^T x + r \\ &\text{s.l.:} -1 \leq x_i \leq 1, \ i = 1, 2, 3, \end{aligned}$$

unde

$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \quad q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13 \end{bmatrix}, \quad r = 1.$$

Problema 2. Să se determine condiţiile necesare şi suficiente în cea mai simplă formă pentru ca $x \in \mathbb{R}^n$ să reprezinte punctul de optim corespunzător unei funcţii convexe diferenţiabile f peste mulţimea de tip simplex, i.e.: $\{x : \sum_{i=1}^{n} x_i = 1, x_i \geq 0,\ i = 1, \ldots, n\}$.

Problema 3. Să se înlocuiască funcţia obiectiv din problema 1 (Probleme rezolvate) cu $\|x\|_\infty$ şi să se rezolve noua problemă.

Problema 4. Să se determine sistemul KKT pentru următoarea problemă de optimizare:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} x_1^2 + x_2^2 \\ &\quad \text{s.l } (x_1-1)^2 + (x_2-1)^2 \leq 1, \\ &\qquad\ (x_1-1)^2 + (x_2+1)^2 \leq 1. \end{aligned}$$

Există multiplicatori Lagrange λ_1^* şi λ_2^* care să demonstreze că punctul x^* este optim?

Problema 5. Fie problema de optimizare:

$$\begin{aligned} &\min_{x\in\mathbb{R}^3} x_1^4 + x_2^4 + x_3^4 + (x_1 + 2x_2 - 4)^2 \\ &\quad \text{s.l } x_1 + x_2 + x_3 = 1,\ x \geq 0. \end{aligned}$$

Să se scrie primul pas al metodei gradient proiectat, pornind din punctul iniţial $x_0 = [1\ 0\ 0]^T$ şi alegând pasul $s_k = 1$.

Problema 6. Fie problema de optimizare:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} x_1^4 + x_2^4 + (x_1-1)^2 + (x_2-4)^2 \\ &\quad \text{s.l } x_1 + x_2 = 16 \end{aligned}$$

Să se scrie aproximarea Taylor de ordin II pentru funcţia obiectiv şi să se scrie primul pas al metodei Newton proiectat pornind din punctul iniţial $x_0 = [14 \quad 2]^T$.

Problema 7. Fie problema de optimizare:

$$\begin{aligned} &\min_{x\in\mathbb{R}^2} x_1^2 + 1 \\ &\quad \text{s.l } (x-1)(x-4) \leq 0 \end{aligned}$$

Să se determine mulţimea fezabilă, punctul de minim şi valoarea minimă pentru această problemă. Să se traseze pe acelaşi grafic funcţia obiectiv, mulţimea fezabilă, punctul de minim şi Lagrangianul $\mathcal{L}(x, \lambda)$ pentru două valori pozitive ale lui λ luate la alegere. Să se formuleze problema duală, să se demonstreze că este concavă şi să se găsească punctul optim dual λ^*.

Problema 8. Considerăm problema celor mai mici pătrate cu constrângeri de egalitate:

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2 \\ &\text{s.l.: } Cx = d \end{aligned}$$

unde $A \in \mathbb{R}^{m \times n}$, $\text{rang}(A) = n$, $C \in \mathbb{R}^{p \times n}$, cu $\text{rang}(C) = p$. Să se expliciteze condiţiile KKT ale acestei probleme şi să se determine expresia soluţiei primale x^* şi a celei duale μ^*.

Problema 9. Fie problema de optimizare:

$$\begin{aligned} &\min_{x \in \mathbb{R}^2} \; -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3) \\ &\text{s.l.: } x_1^2 + x_2^2 + x_3^2 = 1 \end{aligned}$$

Să se determine punctele punctele KKT pentru această problemă. Care pereche de puncte îi va corespunde punctului de optim?.

Problema 10. Fie problema de optimizare:

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} \sum_{i=1}^{n} \log(c_i + x_i) \\ &\text{s.l.: } x \geq 0, \quad \sum_{i=1}^{n} x_i = 1, \end{aligned}$$

unde c_i reprezintă parametrii cunoscuţi.

(i) Să se determine problema duală.

(ii) Să se determine un punct KKT.