Tehnici de Optimizare

- Seminar 4 -

Probleme de programare neliniară. Algoritmi.

1 Probleme de programare convexă

Modelul general al problemelor de programare neliniara se rezuma la:

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.l.} \quad & g_i(x) = 0, \quad \forall i = 1, \cdots, m \\ & h_i(x) \le 0, \quad \forall i = 1, \cdots, p. \end{aligned} \tag{1}$$

unde $f : \mathbb{R}^n \to \mathbb{R}$ reprezinta funcţia obiectiv, g_i, h_i funcţiile asociate constrângerilor de egalitate, respectiv inegalitate. În particular, problema:

$$\begin{aligned} f^* = \min_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.l.} \quad & A_i x = b_i, \quad \forall i = 1, \cdots, m \\ & C_i x \le d_i, \quad \forall i = 1, \cdots, q. \\ & h_i(x) \le 0, \quad \forall i = q+1, \cdots, p. \end{aligned} \tag{2}$$

este convexă dacă $f, \{h_i\}_{1\le i\le p}$ sunt functii convexe.

Functia Lagrangian:

$$\begin{aligned} L(x,\lambda,\mu) &= f(x) + \lambda^T g(x) + \mu^T h(x) \\ L(x,\lambda,\mu) &= f(x) + \lambda^T (Ax-b) + \mu^T h(x) = f(x) + \sum_i \lambda_i (A_i x - b_i) + \sum_i \mu_i h_i(x) \end{aligned}$$

Functia Duala:

$$\phi(\lambda,\mu) = \min_x L(x,\lambda,\mu)$$

Problema duala:

$$\phi^* = \max_{\mu\ge 0,\lambda} \ \phi(\lambda,\mu)$$

Condiţiile suficiente de optimalitate Kuhn-Tucker:
Dacă f, h convexe şi $f^* = \phi^*$ atunci x^* solutie pentru (3) dacă şi numai dacă:

$$\begin{aligned} &\nabla_x L(x^*,\lambda^*,\mu^*) = 0 && \text{Optimalitate} \\ &Ax^* + b = 0, h(x^*) \le 0 \qquad \mu^* \ge 0 && \text{Fezabilitate} \\ &\mu_i^* h_i(x^*) = 0, && \text{Complementaritate} \end{aligned}$$

Pentru a avea **optimalitate tare** $f^* = \phi^*$ este suficientă condiţia Slater:

$$\exists x : \quad h(x) < 0, \quad Ax = b$$

Mai general, condiţia de mai sus se poate relaxa: fie h_i, unde $1 \leq i \leq q$, funcţii afine, iar restul neliniare, atunci

$$\exists x : \quad h_i(x) \leq 0 \ \ [1 \leq i \leq q] \qquad h_i(x) < 0, \ \ [q+1 \leq i \leq p] \qquad Ax = b$$

Multiplicatorii Lagrange reprezintă soluţia problemei duale:

$$(\lambda^*, \mu^*) = \arg \max_{\mu \geq 0, \lambda} \ \phi(\lambda, \mu)$$

Concluzie: rezolvând problema duală, i.e. calcularea (λ^*, μ^*), sub presupunerile condiţiilor Kuhn-Tucker, se recupereaza soluţia primală x^* din rezolvarea sistemului:

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

Exercitii: Deduceti problema duala şi condiţii le de optimalitate KT pentru urmatoarele probleme primale:

1. $\min_{x \in \mathbb{R}^2} \ 2x_1^2 + 3x_2^2$ s.l. $x_1 + x_2 \leq 1$

   Functia Lagrangian:

$$L(x, \mu) = 2x_1^2 + 3x_2^2 + \mu(x_1 + x_2 - 1) = \frac{1}{2} x^T \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} x + \mu [1 \ \ 1] x - \mu$$

$$\min_x L(x, \mu) = L(x^*(\mu), \mu) => \nabla_x L(x^*(\mu), \mu) = 0$$

$$\begin{bmatrix} 4x_1(\mu) + \mu \\ 6x_2(\mu) + \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x^*(\mu) = \begin{bmatrix} -\mu/4 \\ -\mu/6 \end{bmatrix} = -\mu \begin{bmatrix} 1/4 \\ 1/6 \end{bmatrix}$$

   Functia Duala:

$$\phi(\mu) = \min_x L(x, \mu) = L(x^*(\mu), \mu) = -\mu^2/6 - \mu$$

$$\max_{\mu \geq 0} \ -\mu^2/6 - \mu \qquad \mu^* = 0 \Rightarrow x^*(\mu^*) = x^* = 0$$

$$\begin{bmatrix} 4x_1^* + \mu^* \\ 6x_2^* + \mu^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{Optimalitate}$$

$$x_1^* + x_2^* \leq 1 \qquad \mu^* \geq 0 \qquad \text{Fezabilitate}$$

$$\mu^*(x_1^* + x_2^* - 1) = 0, \qquad \text{Complementaritate}$$

$$Pp. \ \ \mu^* > 0 \Rightarrow x_1^* = -\mu^*/4, x_2^* = -\mu^*/6 \Rightarrow \mu^* = -12/5 (contradictie!)$$

$$\mu^* = 0 \Rightarrow x^* = 0$$

2. $\min_{x\in\mathbb{R}^2}\ x_1 - x_2$ s.l. $x_1 + x_2 = 1, -x_1 \leq 0, -x_2 \leq 0$

Functia Lagrangian:

$$\begin{aligned} L(x,\lambda,\mu) &= x_1 - x_2 + \lambda(x_1 + x_2 - 1) + \mu_1(-x_1) + \mu_2(-x_2) \\ &= x_1(\lambda - \mu_1 + 1) + x_2(\lambda - \mu_2 - 1) - \lambda \end{aligned}$$

$$\phi(\lambda,\mu) = \begin{cases} -\infty, & \text{daca } \mu_1 \neq \lambda + 1, \mu_2 \neq \lambda - 1 \\ -\lambda & \text{daca } \mu_1 = \lambda + 1, \mu_2 = \lambda - 1 \end{cases}$$

Problema duala:

$$\max_{\lambda,\mu\geq 0}\ -\lambda \ \text{ s.l. } \ \mu_1 = \lambda + 1, \mu_2 = \lambda - 1$$

$$\begin{array}{lll} x_1^* - x_2^* = -\lambda^* & & \text{Optimalitate} \\ x_1^* + x_2^* = 1, x^* \geq 0 & \mu_1^* = \lambda^* + 1, \mu_2^* = \lambda^* - 1, \mu^* \geq 0 & \text{Fezabilitate} \\ \mu_1^* x_1^* = 0,\ \mu_2^* x_2^* = 0, & & \text{Complementaritate} \end{array}$$

3. Program liniar: $\min_x\ c^T x$ s.l. $Ax = b, x \geq 0$

4. $\min_x\ c^T x$ s.l. $\|x\|^2 \leq 1$

5. CMMP: $\min_x\ \frac{1}{2}x^T x$ s.l. $Ax = b, \quad A \in \mathbb{R}^{m\times n}$

Presupunere: Exista $x:$ a.i. $Ax = b$

Functia Lagrangian:

$$L(x,\lambda) = \frac{1}{2}x^T x + \sum_i \lambda_i(A_i x - b_i) = \frac{1}{2}\|x\|^2 + \lambda^T(Ax - b)$$

Functia duala:

$$\phi(\lambda) = \min_x \frac{1}{2}\|x\|^2 + \lambda^T(Ax - b)$$

$$x^*(\lambda) = -A^T\lambda$$

$$\max_\lambda \phi(\lambda) = -\frac{1}{2}\|A^T\lambda\|^2 - \lambda^T b \qquad Hess = -AA^T \preceq 0(\phi\ concava)$$

$$AA^T\lambda^* = -b => \lambda^* = -(AA^T)^{-1}b \quad (AA^T\ full - rank) => x^* = A^T(AA^T)^{-1}b$$

Sistemul Kuhn-Tucker: $x^* = -A^T\lambda^*(Optimalitate), Ax^* = b(Fezabilitate)$

6. Program patratic: $\min_x\ \frac{1}{2}x^T Hx + q^T x$ s.l. $Ax = b$

   Sistemul Kuhn-Tucker:

$$\begin{aligned} &Hx + q + A^T\lambda = 0, [n \text{ ecuatii}] \\ &Ax = b\ \ [m \text{ ecuatii}] \end{aligned}$$

   Daca $H \succ 0$ atunci

$$x^* = -H^{-1}(q + A^T\lambda^*) \Rightarrow AH^{-1}(A^T\lambda^* + q) = -b \Rightarrow \lambda^* = -AH^{-1}A^T(b + AH^{-1}q)$$

7. $\pi_H(x^0) = \arg\min_x\ \frac{1}{2}\|x - x^0\|^2$ s.l. $a^T x = b$

$$L(x, \lambda) = \frac{1}{2}\|x - x^0\|^2 + \lambda(a^T x - b)$$

   Sistemul Kuhn-Tucker:

$$\begin{aligned} &x^* = x^0 - \lambda^* a = 0, [Optimalitate] \\ &a^T x^* = b\ \ [Fezabilitate] \end{aligned}$$

$$a^T(x^0 - \lambda^* a) = b \Rightarrow \lambda^* = \frac{a^T x^0 - b}{\|a\|^2} \Rightarrow \pi_H(x^0) = x^0 - \frac{a^T x^0 - b}{\|a\|^2}a$$

2 Algoritmi

2.1 Metoda Gradient Proiectat cu Liniarizare

$$\begin{aligned} x^{k+1} = \arg\min_{x\in\mathbb{R}^n}\ & f(x^k) + \nabla f(x^k)^T(x - x^k) + \frac{1}{2\alpha}\|x - x^k\|^2 \\ \text{s.l. }\ & A_i x = b_i, \quad \forall i = 1, \cdots, m \\ & h_i(x) \approx h_i(x^k) + +\nabla h_i(x^k)^T(x - x^k) \le 0, \quad \forall i = 1, \cdots, p. \end{aligned} \tag{3}$$

Presupunem ca $A = 0$.

Metoda Gradient Proiectat cu Liniarizare(x^0, ϵ)

Initializeaza $k = 0$.
Cat timp *criteriu_stop* :
1. Calculeaza $\nabla f(x^k), \nabla h_1(x^k), \cdots, \nabla h_p(x^k)$
2. Actualizeaza:

$$x^{k+1} = \pi_{Q_k}(x^k - \alpha_k \nabla f(x^k)), \quad Q_k = \{x \ : \ Ax = b, h_i(x^k) + \nabla h_i(x^k)^T(x - x^k) \leq 0, \quad \forall i = 1, \cdots, p\}$$

3. Set $k := k + 1$.

Criteriu de stop: $\|x^{k+1} - x^k\| \leq \epsilon$

Aratati forma explicita a iteratiei MGPL pentru problema:

$$\min_{x \in \mathbb{R}^2} \ \frac{1}{2} x^T H x + q^T x \ \text{ s.l. } \ \|x\|^2 \leq 1$$

Calculati prima iteratie cu pas $\frac{1}{L_f}$, pornind din $x^0 = [0\ 1]$ pentru $H = I_2, q = [11]^T$.

$$y = x^0 - \frac{1}{L} \nabla f(x^0) = x^0 - \nabla f(x^0) = x^0 - (Hx^0 + q) = -q$$

$$h_1(x) = \|x\|^2 - 1, \nabla h_1(x^0) = 2x^0, h_1(x^0) = 0$$

$$\begin{aligned} Q_0 &= \{x : (x^0)^T(x - x^0) = x_1^0(x_1 - x_1^0) + x_2^0(x_2 - x_2^0) \leq 0\} \\ &= \{x : x_2 \leq 1\} \end{aligned}$$

$$x^1 = \pi_{Q_0}(y) = [-1 \quad -1]^T.$$

2.2 Algoritmul dual

$$\max_{\lambda, \mu \geq 0} \ \phi(\lambda, \mu) = - \min_{\lambda, \mu \geq 0} \ -\phi(\lambda, \mu)$$

Metoda Gradient Proiectat Dual(λ^0, ϵ)

Initializeaza $k = 0$.
Cat timp *criteriu_stop* :
1. Calculeaza $\nabla_\lambda \phi(\lambda^k, \mu^k), \nabla_\mu \phi(\lambda^k, \mu^k)$
2. Actualizeaza:

$$\begin{aligned} \lambda^{k+1} &= \lambda^k + \alpha_k \nabla_\lambda \phi(\lambda^k, \mu^k) \\ \mu^{k+1} &= \max(0, \mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)) \end{aligned}$$

3. Set $k := k + 1$.

Criteriu de stop: $\|\lambda^{k+1} - \lambda^k\| \leq \epsilon$

$$\nabla_x L(\hat{x}^*, \lambda_f, \mu_f) = 0 \quad [\hat{x}^* \;\; aproape - optim]$$

Aratati forma explicita a iteratiei MGPL pentru problema:

$$\min_{x \in \mathbb{R}^n} \quad \frac{1}{2} x^T H x + q^T x \;\; \text{s.l.} \;\; \|x\|^2 \leq 1$$

Calculati prima iteratie cu pas $\alpha = 1$, pornind din $\mu^0 = 1$ pentru $H = I_2, q = [1 \; 1]^T$.

$$\phi(\mu) = \min_x \; \frac{1}{2} x^T H x + q^T x + \mu(\|x\|^2 - 1) = -\frac{1}{1 + 2\mu} - \mu$$

$$(H + 2\mu I_2)x + q = 0 \Rightarrow x(\mu) = -(H + 2\mu I_2)^{-1} q = -\frac{1}{1 + 2\mu} q$$

$$\mu^{k+1} = \mu^k - \alpha_k (1 + 2/(1 + 2\mu^k)^2)$$