Tehnici de Optimizare

- Seminar -

Probleme de optimizare cu constrângeri. Sistemul Kuhn-Tucker.

1 Probleme de programare convexă

Modelul general al problemelor de programare neliniară se rezumă la:

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

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

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

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

Funcţia Lagrangian:

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

Funcţia duală:

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

Problema duală:

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

Condiţiile suficiente de optimalitate Kuhn-Tucker:
Dacă f, h convexe şi $f^* = \phi^*$ atunci x^* soluţie pentru (2) 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 recuperează soluţia primală x^* din rezolvarea sistemului:

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

Exerciţii: Deduceti problema duala şi condiţiile de optimalitate Kuhn-Tucker pentru urmatoarele probleme convexe:

1. $\min_{x \in \mathbb{R}^2} \; 2x_1^2 + 3x_2^2$ s.l. $x_1 + x_2 \leq 1$
2. $\min_{x \in \mathbb{R}^2} \; x_1 - x_2$ s.l. $x_1 + x_2 = 1, -x_1 \leq 0, -x_2 \leq 0$
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. $Ax = b, \quad A \in \mathbb{R}^{m \times n}$. Pp. că există o soluţie a sistemului $Ax = b$.
6. Program patratic: $\min_x \; \frac{1}{2} x^T H x + q^T x$ s.l. $Ax = b$

Rezolvări:

1. Calculăm funcţia Lagrangian:

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

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

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

Calculăm funcţia duală:

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

De aici, avem problema duală

$$\max_{\mu \geq 0} \;\; -\mu^2/6 - \mu,$$

cu soluţia $\mu^* = 0$. Deci, $x^*(\mu^*) = x^* = 0$.

Condiţiile K-T sunt:

$$\begin{aligned}
&\begin{bmatrix} 4x_1^* + \mu^* \\ 6x_2^* + \mu^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} && \text{Optimalitate} \\
&x_1^* + x_2^* \le 1 \qquad \mu^* \ge 0 && \text{Fezabilitate} \\
&\mu^*(x_1^* + x_2^* - 1) = 0, && \text{Complementaritate}
\end{aligned}$$

De asemenea, putem deducem soluţia din următoarea contradicţie: dacă $\mu^* > 0$ atunci

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

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

   În acest caz, funcţia Lagrangian şi funcţia duală sunt:

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

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

   De aici, obţinem problema duală:

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

   şi sistemul K-T:

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

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

   Din funcţia Lagrangian:

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

   deducem funcţia duală $\phi(\lambda) = \min_x \frac{1}{2}\|x\|^2 + \lambda^T(Ax - b)$ şi observăm $x^*(\lambda) = -A^T\lambda$. De aici, obţinem problema duală:

$$\max_\lambda \phi(\lambda) = -\frac{1}{2}\|A^T\lambda\|^2 - \lambda^T b$$

   Hessiana $\nabla^2\phi(\lambda) = -AA^T \preceq 0$ ceea ce implică c a ϕ este concavă. Condiţiile de ordin I sunt suficiente pentru problema duală:

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

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

6. Program pătratic: $\min\limits_{x\in\mathbb{R}^n}\ \frac{1}{2}x^T Hx + q^T x$ s.l. $Ax = b$

   Sistemul Kuhn-Tucker:

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

   Dacă $H \succ 0$ atunci

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

2 Algoritmul de gradient dual

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

Metoda Gradient Proiectat Dual$(\lambda^0, \mu^0, \epsilon)$

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:

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

3. $k := k + 1$.

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

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

Exerciţiu. Fie problema pătratică:

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

- Calculaţi problema duală.
- Arătaţi forma explicită a iteraţiei MGD în acest caz particular.
- Calculaţi prima iteraţie cu pas $\alpha = 1$, pornind din $\lambda^0 = 1$ pentru $H = I_2, q = \begin{bmatrix}1\\1\end{bmatrix}$.

Rezolvare pe scurt: Din forma funcţiei Lagrange $\mathcal{L}(x,\lambda) = \frac{1}{2}x^T Hx + q^T x + \lambda(\|x\|^2 - 1)$ deducem expresia funcţiei duale $\phi(\lambda) = \min_x \ \mathcal{L}(x,\lambda))$. Condiţia de optimalitate de ordin I implică:

$$(H + 2\lambda I_2)x + q = 0 \Rightarrow x(\lambda) = \arg\min_x \mathcal{L}(x,\lambda)) = -(H + 2\lambda I_2)^{-1}q.$$

Funcţia şi problema duală:

$$\max_{\lambda\geq 0} \ \phi(\lambda) = \mathcal{L}(x(\lambda),\lambda)) = -\frac{1}{2}q^T(H + 2\lambda I_2)^{-1}q - \lambda.$$

Remarcăm $\nabla\phi(\lambda) = \|(H + 2\lambda I_2)^{-1}q\|_2^2 - 1$. Iteraţia MGD pentru problema duală de mai sus:

$$\lambda^{k+1} = \pi_{\mathbb{R}_+}\left(\lambda^k + \alpha_k[\|(H + 2\lambda^k I_2)^{-1}q\|_2^2 - 1]\right).$$

Pentru datele $\alpha = 1, \lambda^0 = 1, H = I_2, q = \begin{bmatrix}1\\1\end{bmatrix}$ avem

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

Deci $\phi(\lambda) = -\frac{1}{1+2\lambda} - \lambda$. Observăm că iteraţia MGD are forma:

$$\lambda^{k+1} = \max\{0, \lambda^k + (2/(1 + 2\lambda^k)^2 - 1)\}.$$