Tehnici de Optimizare

- Seminar 3 -

Probleme de optimizare cu constrângeri. Algoritmi.

1 Probleme de optimizare cu constrângeri

Fie mulţimea $Q \subset \mathbb{R}^n$, modelul general al problemei de optimizare constrânsă este:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \;\; & f(x) \\ \text{s.l.} \;\; & x \in Q. \end{aligned} \tag{1}$$

unde $f : \mathbb{R}^n \to \mathbb{R}$ reprezinta funcţia obiectiv, iar Q mulţimea fezabilă convexă în care se realizează căutarea. Desi Q poate fi exprimată printr-o reprezentare de tip subnivel (e.g. $Q = \{x \in \mathbb{R}^n : \; ||x|| \leq 1\}$), modelul (1) se foloseşte în probleme cu constrângeri simple, e.g. bilă, hiperplan, semiplan etc. Distingem un caz particular, cand minimul global al funţiei f sa afla in Q. In această situaţie, rezolvarea problemei (1) se reduce la rezolvarea unei probleme fără constrângeri.

2 Proiecţia ortogonală pe mulţimi convexe

Fie x^0 un punct oarecare în $\mathbb{R}^n$, atunci cel mai apropiat punct din mulţimea Q de punctul x^0 reprezintă proiecţia ortogonală $\pi_Q(x^0)$, i.e.:

$$\begin{aligned} \min \; & \|x - x^0\|^2 \\ \text{s.l.} \;\; & x \in Q. \end{aligned} \tag{2}$$

Funcţia obiectiv exprimă distanţa de la x^0 la variabila x, care este minimizată prin selecţia punctului optim din Q. Evident, dacă $x^0 \in Q$ atunci proiecţia este chiar x^0,i.e. $\pi_Q(x^0) = x^0$. Deci cazul netrivial este dat de $x^0 \notin Q$, care implică în general că $\pi_Q(x^0)$ stă pe frontiera lui Q.

Pentru mulţimi convexe "simple" Q, proiectia π_Q are forma explicită:

Exemple pe $\mathbb{R}$.

- $Q = \{x \in \mathbb{R} : x \geq 0\}$

$$x^0 \geq 0 \Rightarrow \pi_Q(x^0) = x^0$$
$$x^0 < 0 \Rightarrow \pi_Q(x^0) = 0$$
$$\pi_Q(x^0) = \max\{0, x^0\}$$

- $Q = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$

$$x^0 < 0 \Rightarrow \pi_Q(x^0) = 0$$
$$x^0 \in (0, 1) \Rightarrow \pi_Q(x^0) = x^0$$

$$x^0 > 1 \Rightarrow \pi_Q(x^0) = 1$$
$$\pi_Q(x^0) = \min\{\max\{0, x^0\}, 1\}$$

Exemple pe $\mathbb{R}^2$.

- $Q = \{x \in \mathbb{R}^2 : x \geq 0\}$

$$\pi_Q(x^0) = \begin{bmatrix} \max\{0, x_1^0\} \\ \max\{0, x_2^0\} \end{bmatrix}$$

- $Q = \{x \in \mathbb{R} : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\} = [0,1] \times [0,1]$

$$\pi_Q(x^0) = \begin{bmatrix} \min\{1, \max\{0, x_1^0\}\} \\ \min\{1, \max\{0, x_2^0\}\} \end{bmatrix}$$

- $Q = \{x \in \mathbb{R} : \|x\| \leq 1\}$

$$x^0 \notin Q(\|x^0\| > 1) \Rightarrow \pi_Q(x^0) = \frac{1}{\sqrt{(x_1^0)^2 + (x_2^0)^2}} \begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix}$$
$$x^0 \in Q(\|x^0\| \leq 1) \Rightarrow \pi_Q(x^0) = x^0$$

$$\pi_Q(x^0) = \min\{1, \frac{1}{\|x^0\|}\}x^0$$

Exemple pe $\mathbb{R}^n$.

- bila $B(0; r) = \{x : \|x\| \leq r\}$, $\pi_B(x^0) = \min\{1, \frac{r}{\|x^0\|}\}x^0$
- hiperplan $H_{a,b} = \{x : a^T x = b\}, \pi_H(x^0) = x^0 - \frac{a^T x^0 - b}{\|a\|^2} a$

$$\pi_H(x^0) \in S \Rightarrow x^0 - \pi_H(x^0) \in S^\perp$$
$$x^0 - \pi_H(x^0) = \alpha a \Rightarrow \pi_H(x^0) = x^0 - \alpha a$$
$$x^0 - \alpha a \in H \Leftrightarrow a^T(x^0 - \alpha a) = b \Leftrightarrow \alpha = \frac{a^T x^0 - b}{\|a\|^2}$$

- semiplan $S_{a,b} = \{x : a^T x \leq b\}$, $\pi_S(x^0) = x^0 - \frac{\max\{0, a^T x^0 - b\}}{\|a\|^2} a$
- box $T_{l,u} = \{x : l \leq x \leq u\}$, $\pi_T(x^0) = \max\{l, \min\{u, x^0\}\} = \begin{bmatrix} \max\{l_1, \min\{u_1, x_1^0\}\} \\ \cdots \\ \max\{l_n, \min\{u_n, x_n^0\}\} \end{bmatrix}$

Exercitiu. Aflati proiectia punctului $x^0 = \left[\ 2\ \frac{1}{2}\ \right]^T$ pe multimea:

$$Q = \{x \in \mathbb{R}^2 : \ \ x_1^2 \leq x_2\}$$

$$\pi_Q(x^0) = \arg\min_{x \in \mathbb{R}^2} \ \ \frac{1}{2}\|x - x^0\|^2$$
$$\text{s.l. } x_1^2 \leq x_2$$

$$\min_{x\in\mathbb{R}^2} \quad \frac{1}{2}\left[(x_1-x_1^0)^2+(x_2-x_2^0)^2\right]$$
$$\text{s.l. } x_1^2 = x_2$$

$$\min_{x\in\mathbb{R}^2} \quad \frac{1}{2}\left[(x_1-x_1^0)^2+(x_1^2-x_2^0)^2\right] = \frac{x_1^4}{2} + \left(\frac{1}{2}-x_2^0\right)x_1^2 - x_1^0 x_1 + c$$

$$\min_{x\in\mathbb{R}^2} \quad \frac{x_1^4}{2} - 2x_1 \Rightarrow 2x_1^3 - 2 = 0 \Rightarrow x_1 = 1, x_2 = 1 \Rightarrow \pi_Q(x^0) = [1 \;\; 1]^T$$

Cand Q este rezultatul unei intersectii de multimi simple, proiectia poate fi foarte dificila. Simpla aflare a unui punct oarecare dintr-o astfel de multime necesita un algoritm specific.

Problema de fezabilitate convexa. Fie multimile $\{Q_i\}$ convexe simple si:

$$Q = Q_1 \cap Q_2 \cap \cdots \cap Q_m,$$

gasiti $x \in Q$ folosind doar proiectiile pe $Q_i, \pi_{Q_i}(\cdot)$.
Algoritmul proiectiilor alternative:

1. Alege un indice i_k
2. Calculeaza: $x^{k+1} = \pi_{Q_{i_k}}(x^k)$

Exercitiu. Fie multimea:

$$Q = \underbrace{\{x\in\mathbb{R}^2 : \;\; x_1 = 1\}}_{H} \cap \underbrace{\{x\in\mathbb{R}^2 : \;\; x_1^2 + x_2^2 \le 1\}}_{B}$$

Fie $y^0 = [1 \;\; a]^T, |a| > 1$. Analizati convergenta urmatorului algoritm pentru o obtine un $x \in Q$.

$$x^k = \pi_H(y^k)$$
$$y^{k+1} = \pi_B(x^k)$$

$$x^0 = \pi_H(y^0) = y_0$$
$$y^1 = \pi_B(y^0) = \frac{1}{\sqrt{1+a^2}}\begin{bmatrix}1\\a\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{1+a^2}}\\ \frac{a}{\sqrt{1+a^2}}\end{bmatrix}$$

$$x^1 = \pi_H(y^1) = \begin{bmatrix} 1 \\ \frac{a}{\sqrt{1+a^2}} \end{bmatrix}$$

$$y^2 = \pi_B(x^1) = \frac{1}{\sqrt{\frac{1+2a^2}{1+a^2}}} \begin{bmatrix} 1 \\ \frac{a}{\sqrt{1+a^2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\frac{1+2a^2}{1+a^2}}} \\ \frac{a}{\sqrt{1+2a^2}} \end{bmatrix}$$

$$x^2 = \pi_H(y^2) = \begin{bmatrix} 1 \\ \frac{a}{\sqrt{1+2a^2}} \end{bmatrix}$$

$$\dots$$

$$x^k = \pi_H(y^k) = \begin{bmatrix} 1 \\ \frac{a}{\sqrt{1+ka^2}} \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} (k \to \infty)$$

$$\|x^k - \begin{bmatrix} 1 \\ 0 \end{bmatrix}\| \leq \frac{a}{\sqrt{1+ka^2}} = \frac{1}{\sqrt{k+1/a^2}} = O(1/\sqrt{k})$$

3 Algoritmi de optimizare

Presupunem ca mulţimea fezabiă este convexă. În acest caz majoritatea algoritmilor de optimizare pentru problemele fără constrângeri se pot adapta la modelul (1). Modificarea esenţială asupra unui algoritm, necesară pentru a rezolva problema (1), reprezintă realizarea proiecţiei ortogonale pe mulţimea fezabilă.

3.1 Metoda Gradient Proiectat

In principiu, metoda gradient reprezintă un algoritm de ordin I care generează un şir de puncte (vectori) $x^1, x^2, ...,$ pornind dintr-un punct iniţial ales x^0.

Metoda Gradient Proiectat(x^0, ϵ)

Initializeaza $k = 0$.
Cat timp *criteriu_stop* :
1. Calculeaza $\nabla f(x^k)$
2. Actualizeaza: $x^{k+1} = \pi_Q(x^k - \alpha_k \nabla f(x^k))$
3. Set $k := k + 1$.

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

Un aspect important al metodei este alegerea pasului $\alpha_k > 0$ dintre urmatoarele optiuni:

1. **Constant**: $\alpha_k = \frac{1}{L_f}$
$$f(x^{k+1}) \leq f(x^k) - \frac{L_f}{2}\|x^{k+1} - x^k\|^2.$$
2. **Ideal**: $\alpha_k = \arg\min_{\alpha \geq 0} f(\pi_Q(x^k - \alpha \nabla f(x^k))$
3. **Backtracking**: alege $c > 0$, ajustează pasul α_k astfel incat sa aiba loc relaţia de descrestere:
$$f(\pi_Q(x^k - \alpha_k \nabla f(x^k)) \leq f(x^k) - \frac{c}{\alpha_k}\|x^{k+1} - x^k\|^2. \tag{3}$$
procedura presupune alegerea unui parametru $\rho \in (0, 1]$ si actualizarea pasului, dupa cum urmeaza:
   - (i) Alegem $c, \rho \in (0, 1), \alpha_{k,0} > 0$
   - (ii) Cat timp $\alpha_{k,t}$ nu satisface (3) iteram: $\alpha_{k,t+1} = \rho\alpha_{k,t}$; $\quad t = t + 1$

3.2 Exemple iteratie MGP

Bila: $B(0;r) = \{x : \|x\| \leq r\}$

Iteratie MGP: $x^{k+1} = \pi_B(x^k - \alpha_k \nabla f(x^k)) = \min\left\{1, \frac{r}{\|x^k - \alpha_k \nabla f(x^k)\|}\right\}\left(x^k - \alpha_k \nabla f(x^k)\right)$

Hiperplan $H_{a,b} = \{x : a^T x = b\}$

Iteratie MGP: $x^{k+1} = \pi_H(x^k - \alpha_k \nabla f(x^k)) = \left(x^k - \alpha_k \nabla f(x^k)\right) - \frac{a^T(x^k - \alpha_k \nabla f(x^k)) - b}{\|a\|^2} a$

Cum variază convergenţa cu:

- *numărul de conditionare* al punctului de minim?
- pasul?

3.2.1 Probleme rezolvate

1. Fie problema

$$\begin{aligned} \min_x \;\; & f(x) = x_1^4 + x_2^4 + x_3^4 + (x_1 + 2x_2 - 4)^2 \\ \text{s.l.} \;\; & x_1 + x_2 + x_3 = 1. \;\; (a^T x = b) \end{aligned}$$

Calculati prima iteratie a MGP cu pas constant $\alpha = 1$, pornind din $x^0 = [1\ 0\ 0]$.

Rezolvare.

3.3 Metoda Newton Proiectat

Metoda Newton Proiectat(x^0, ϵ)
Initializeaza $k = 0$.
Cat timp *criteriu_stop* :
1. Calculeaza $\nabla f(x^k), \nabla^2 f(x^k)$
2. Actualizeaza: $x^{k+1} = \arg\min_{x \in Q} f(x^k) + \nabla f(x^k)^T(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k)$
3. Set $k := k + 1$.

Alegeri pas:

- constant: alegerea standard $\alpha = 1$
- ideal: $\alpha_k = \arg\min_{\alpha > 0} f(\pi_Q\left(x^k - \alpha[\nabla^2 f(x^k)]^{-1} \nabla f(x^k)\right))$

3.3.1 Probleme rezolvate

Fie problema

$$\begin{aligned} \min_x \ & x_1^4 + x_2^4 + (x_1 - 1)^2 + (x_2 - 4)^2 \\ \text{s.l.} \ & x_1^2 + x_2^2 \leq 1. \end{aligned}$$

Calculaţi prima iteraţie a metodei Newton pornind din $x^0 = [14\ 2]^T$ cu $\alpha_0 = 1$.
Rezolvare.

$$\nabla f(x) = \begin{bmatrix} 4x_1^3 + 2(x_1 - 1) \\ 4x_2^3 + 2(x_2 - 4) \end{bmatrix} \quad \nabla^2 f(x) = \begin{bmatrix} 12x_1^2 + 2 & 0 \\ 0 & 12x_2^2 + 2 \end{bmatrix}$$

$$[\nabla^2 f(x)]^{-1} = \begin{bmatrix} \frac{1}{12x_1^2+2} & 0 \\ 0 & \frac{1}{12x_2^2+2} \end{bmatrix}$$

$$d^N(x^0) = [\nabla^2 f(x^0)]^{-1} \nabla f(x^0) = \begin{bmatrix} \frac{1}{2354} & 0 \\ 0 & \frac{1}{50} \end{bmatrix} \begin{bmatrix} 2770 \\ 28 \end{bmatrix} \approx \begin{bmatrix} 1.18 \\ 0.5 \end{bmatrix}$$

$$y = x^0 - d^N(x^0) \approx [12\ 1.5]^T$$

$$x^1 = \pi_Q(y) = \frac{y}{\|y\|}$$