TEHNICI DE OPTIMIZARE
Curs 5

Andrei Pătraşcu

Departament Informatică
Universitatea din Bucureşti

- **Probleme de minimizare cu constrângeri (POC). Condiţii de optimalitate.**
- Algoritmi de ordinul I pentru POC
- Convergenţă
- Exemple şi aplicaţii

$$(POC): \quad \min_{x \in \mathbb{R}^n} \; f(x) \;\text{ supus la }\; x \in Q$$

- Q mulţime fezabilă (convexă, închisă)

$$(POC): \quad \min_{x \in \mathbb{R}^n} \; f(x) \; \text{ supus la } \; x \in Q$$

- Q mulţime fezabilă (convexă, închisă)
- se urmăreşte determinarea lui $x^* \in Q$ care asigură:

$$\begin{aligned} f(x^*) &\leq f(x), & \forall x \in Q, \|x - x^*\| \leq \epsilon \;\; \text{(minim local)} \\ f(x^*) &\leq f(x), & \forall x \in Q \;\; \text{(minim global)} \end{aligned}$$

$$(POC): \quad \min_{x \in \mathbb{R}^n} \; f(x) \;\text{ supus la }\; x \in Q$$

- Q mulţime fezabilă (convexă, închisă)
- se urmăreşte determinarea lui $x^* \in Q$ care asigură:

$$\begin{aligned} f(x^*) &\leq f(x), \qquad \forall x \in Q, \|x - x^*\| \leq \epsilon \;\text{ (minim local)} \\ f(x^*) &\leq f(x), \qquad \forall x \in Q \;\text{ (minim global)} \end{aligned}$$

- echivalent: $\min_x \; f(x) + \imath_Q(x)$, unde $\imath_Q(x) = \begin{cases} 0 & x \in Q \\ \infty & x \notin Q \end{cases}$

- LP: $\min_x \; c^T x \;$ s.l. $\; Ax = b, x \geq 0$

Exemple

- LP: $\min_x \ c^T x \ \text{ s.l. } \ Ax = b, x \geq 0$
- QP: $\min_x \ \frac{1}{2} x^T H x - q^T x \ \text{ s.l. } \ Ax \leq b, \ Gx = h$

Exemple

- LP: $\min_x \; c^T x \;$ s.l. $\; Ax = b, x \geq 0$
- QP: $\min_x \; \frac{1}{2} x^T Hx - q^T x \;$ s.l. $\; Ax \leq b, \; Gx = h$
- SOCP: $\min_x \; c^T x \;$ s.l. $\; \|A_i x - b_i\| \leq \beta_i, \;\; Gx = h \;\; i = 1, \cdots, m$

- LP: $\min_x \; c^T x$ s.l. $Ax = b, x \geq 0$
- QP: $\min_x \; \frac{1}{2}x^T Hx - q^T x$ s.l. $Ax \leq b, \; Gx = h$
- SOCP: $\min_x \; c^T x$ s.l. $\|A_i x - b_i\| \leq \beta_i, \;\; Gx = h \;\; i = 1, \cdots, m$
- QCQP: $\min_x \;\; \frac{1}{2}x^T Hx - q^T x$ s.l. $\|A_i x - b_i\| \leq \beta_i, \;\; Gx = h \;\; i = 1, \cdots, m$

Exemple

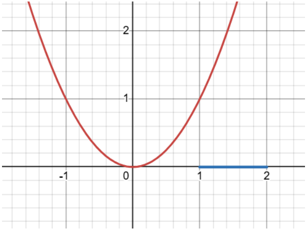
- LP: $\min_x \; c^T x$ s.l. $Ax = b, x \geq 0$
- QP: $\min_x \; \frac{1}{2}x^T Hx - q^T x$ s.l. $Ax \leq b, \; Gx = h$
- SOCP: $\min_x \; c^T x$ s.l. $\|A_i x - b_i\| \leq \beta_i, \;\; Gx = h \;\; i = 1, \cdots, m$
- QCQP: $\min_x \; \frac{1}{2}x^T Hx - q^T x$ s.l. $\|A_i x - b_i\| \leq \beta_i, \;\; Gx = h \;\; i = 1, \cdots, m$
- SDP: $\min_x \; c^T x$ s.l. $x_1 F_1 + \cdots + x_n F_n + G \preceq 0, \;\; Gx = h$

Numim mulţimea fezabilă simplă dacă următoarele obiecte pot fi calculate eficient:

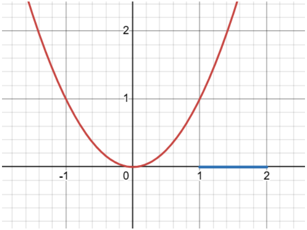
- punct fezabil: $x \in Q$
- proiecţia ortogonală: $\pi_Q(y) := \mathrm{argmin}_x \ \ \|x - y\|$ s.l. $x \in Q$
- oracol liniar: $\phi_Q(y) := \mathrm{argmin}_x \ \ x^T y$ s.l. $x \in Q$
- mulţimi active
- ...

Constrângeri simple

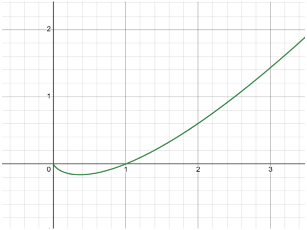
- $\min_x \; x^2$ s.l. $1 \leq x \leq 2$



- $\min_x \; x^2$ s.l. $1 \leq x \leq 2$



- $\min_x \; x \log(x)$ s.l. $\ell \leq x \leq u$



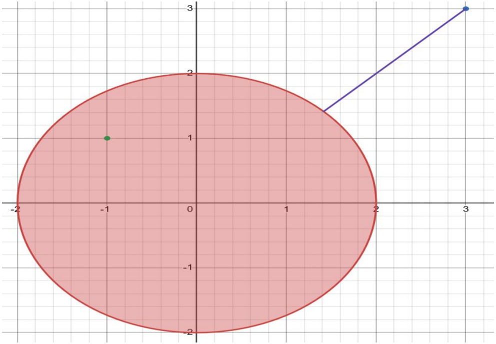
- $\min_x \; x_1^2 + x_2^2$ s.l. $1 \leq x_1 \leq 2, \; 1 \leq x_2 \leq 2$ (separabilitate)

- $\min_x \; x_1^2 + x_2^2$ s.l. $1 \leq x_1 \leq 2, \; 1 \leq x_2 \leq 2$ (separabilitate)
- $\min_x \; x_1^2 + x_2^2$ s.l. $x \geq 0, x_1 + x_2 = 1$ (proiecţie)

Problema proiecţiei ortogonale Euclidiene:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|x - y\|_2^2$$
$$\text{s.l. } x \in Q$$

Soluţia $\pi_Q(y)$: punctul cel mai apropiat de y, care aparţine mulţimii fezabile Q



Problema proiecţiei ortogonale Euclidiene:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|x - y\|_2^2$$
$$\text{s.l. } x \in Q$$

Soluţia $\pi_Q(y)$: punctul cel mai apropiat de y, care aparţine mulţimii fezabile Q

- Dacă Q este închisă atunci $\pi_Q(\cdot)$ există
- $y \in Q \to \pi_Q(y) = y$
- Condiţii de optimalitate: $(y - \pi_Q(y))^T(x - \pi_Q(y)) \leq 0 \quad \forall x \in Q.$

O problemă particulară

- bila $Q := B(0, r) = \{x \in \mathbb{R}^n : \ \|x\| \leq r\} \Rightarrow \pi_Q(y) = \left\{1, \frac{r}{\|y\|}\right\} y$

- bila $Q := B(0,r) = \{x \in \mathbb{R}^n : \ \|x\| \leq r\} \Rightarrow \pi_Q(y) = \left\{1, \frac{r}{\|y\|}\right\} y$
- hiperparalelipiped
  $Q := [l,u]^n = \{x \in \mathbb{R}^n : \ x \in [l,u]\} \Rightarrow \pi_Q(y)_i = \max\{l_i, \min\{u_i, y_i\}\}$

- bila $Q := B(0,r) = \{x \in \mathbb{R}^n : \ \|x\| \leq r\} \Rightarrow \pi_Q(y) = \left\{1, \frac{r}{\|y\|}\right\} y$
- hiperparalelipiped
  $Q := [l,u]^n = \{x \in \mathbb{R}^n : \ x \in [l,u]\} \Rightarrow \pi_Q(y)_i = \max\{l_i, \min\{u_i, y_i\}\}$
- hiperplan $Q := H(a,b) = \{x \in \mathbb{R}^n : \ a^T x = b\} \Rightarrow \pi_Q(y) = y - \frac{a^T y - b}{\|a\|_2^2} a$

O problemă particulară

- bila $Q := B(0,r) = \{x \in \mathbb{R}^n : \ \|x\| \leq r\} \Rightarrow \pi_Q(y) = \left\{1, \frac{r}{\|y\|}\right\} y$
- hiperparalelipiped
  $Q := [l,u]^n = \{x \in \mathbb{R}^n : \ x \in [l,u]\} \Rightarrow \pi_Q(y)_i = \max\{l_i, \min\{u_i, y_i\}\}$
- hiperplan $Q := H(a,b) = \{x \in \mathbb{R}^n : \ a^T x = b\} \Rightarrow \pi_Q(y) = y - \frac{a^T y - b}{\|a\|_2^2} a$
- semispaţiu
  $Q := S(a,b) = \{x \in \mathbb{R}^n : \ a^T x \leq b\} \Rightarrow \pi_Q(y) = y - \frac{\max\{0, a^T y - b\}}{\|a\|_2^2} a$

Alte proprietăţi

- $Q = Q_1 \times Q_2 \Rightarrow \pi_Q(x) = \begin{bmatrix} \pi_{Q_1}(x_1) \\ \pi_{Q_2}(x_2) \end{bmatrix}$

- $Q = Q_1 \times Q_2 \Rightarrow \pi_Q(x) = \begin{bmatrix} \pi_{Q_1}(x_1) \\ \pi_{Q_2}(x_2) \end{bmatrix}$
- $Q = Q_1 \times Q_2 \times \cdots \times Q_n \Rightarrow \pi_Q(x) = \begin{bmatrix} \pi_{Q_1}(x_1) \\ \vdots \\ \pi_{Q_n}(x_n) \end{bmatrix}$

- $Q = Q_1 \times Q_2 \Rightarrow \pi_Q(x) = \begin{bmatrix} \pi_{Q_1}(x_1) \\ \pi_{Q_2}(x_2) \end{bmatrix}$
- $Q = Q_1 \times Q_2 \times \cdots \times Q_n \Rightarrow \pi_Q(x) = \begin{bmatrix} \pi_{Q_1}(x_1) \\ \vdots \\ \pi_{Q_n}(x_n) \end{bmatrix}$
- $Q = Q_1 \cap Q_2 \cap \cdots \cap Q_n \Rightarrow \pi_Q(x) = ?$

$$(POC): \quad \min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s. l. } x \in Q.$$

Reformulare (nediferenţiabilă):

$$(POC): \quad \min_{x \in \mathbb{R}^n} \; f(x) + \iota_Q(x) = \begin{cases} f(x) & x \in Q \\ \infty & x \notin Q. \end{cases}$$

Condiţii de optimalitate?

Reformulare Th. Fermat ($Q = \mathbb{R}^n$):

$$
\begin{aligned}
x^* &= \arg\min_x \; f(x^*) + \nabla f(x^*)^T (x - x^*) + \frac{1}{2\alpha}\|x - x^*\|^2 \\
x^* &= x^* - \alpha \nabla f(x^*) \\
0 &= \nabla f(x^*).
\end{aligned}
$$

Reformulare Th. Fermat ($Q = \mathbb{R}^n$):

$$\begin{aligned} x^* &= \arg\min_{x} \; f(x^*) + \nabla f(x^*)^T (x - x^*) + \frac{1}{2\alpha}\|x - x^*\|^2 \\ x^* &= x^* - \alpha \nabla f(x^*) \\ 0 &= \nabla f(x^*). \end{aligned}$$

Intuiţie în caz constrâns ($Q \subseteq \mathbb{R}^n$):

$$\begin{aligned} x^* &= \arg\min_{x \in Q} \; f(x^*) + \nabla f(x^*)^T (x - x^*) + \frac{1}{2\alpha}\|x - x^*\|^2 \\ x^* &= \pi_Q(x^* - \alpha \nabla f(x^*)) \\ 0 &= x^* - \pi_Q(x^* - \alpha \nabla f(x^*)). \end{aligned}$$

$$(POC): \quad \min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s. l. } x \in Q.$$

Dacă $Q = \mathbb{R}^n$ atunci în toate punctele staţionare $\nabla f(x^*) = 0$.

Definiţie

Fie f diferenţiabilă, $x \in Q$ şi $\alpha > 0$ atunci denumim ***operator de gradient redus peste mulţimea Q*** *transformarea:*

$$\mathcal{G}_Q(x; \alpha) = \frac{1}{\alpha} [x - \pi_Q(x - \alpha \nabla f(x))]$$

Dacă $Q = \mathbb{R}^n$ atunci $\mathcal{G}_Q(x; \alpha) = \nabla f(x)$.

Definiţie

Fie f diferenţiabilă, $x \in Q$ şi $\alpha > 0$ atunci denumim ***operator de gradient redus peste mulţimea Q*** *transformarea:*

$$\mathcal{G}_Q(x;\alpha) = \frac{1}{\alpha}\left[x - \pi_Q(x - \alpha\nabla f(x))\right]$$

- $f(x) = \frac{1}{2}\|x - y\|^2$

$$\mathcal{G}(x;1) = x - \pi_Q(x - (x - y)) = x - \pi_Q(y) = x - x^*$$

Definiţie

Fie f diferenţiabilă, $x \in Q$ şi $\alpha > 0$ atunci denumim ***operator de gradient redus peste mulţimea Q*** *transformarea:*

$$\mathcal{G}_Q(x;\alpha) = \frac{1}{\alpha}\left[x - \pi_Q(x - \alpha\nabla f(x))\right]$$

- $f(x) = \frac{1}{2}\|x - y\|^2$

$$\mathcal{G}(x;1) = x - \pi_Q(x - (x - y)) = x - \pi_Q(y) = x - x^*$$

- $f(x) = \frac{1}{2}\|Ax - b\|^2, Q = B(0;1)$

$$\begin{aligned}\mathcal{G}(x;1) &= x - \pi_Q(x - A^T(Ax - b)) = x - \pi_Q((I - A^TA)x + A^Tb) \\ &= x - \frac{1}{\|(I - A^TA)x + A^Tb\|}[(I - A^TA)x + A^Tb] \\ &= (1 - \bar{\alpha})x + \bar{\alpha}\nabla f(x).\end{aligned}$$

$$(OPC): \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s. l. } \; x \in Q.$$

Pp. $x \in Q, \mathcal{G}(x) \neq 0$ atunci din definiţia lui $\mathcal{G}$ avem:

$$\begin{aligned}
\alpha^2 \|\mathcal{G}(x;\alpha) - \nabla f(x))\|_2^2 &= \|\pi_Q(x - \alpha\nabla f(x)) - (x - \alpha\nabla f(x))\|_2^2 \\
&= \min_{z \in Q} \|z - (x - \alpha\nabla f(x))\|_2^2 \\
&\leq \|x - (x - \alpha\nabla f(x))\|_2^2 = \alpha^2 \|\nabla f(x)\|_2^2.
\end{aligned}$$

Desfăşurând norma rezultă:

$$\frac{1}{2}\|\mathcal{G}(x;\alpha)\|_2^2 \leq \nabla f(x)^T \mathcal{G}(x;\alpha).$$

$$(OPC): \quad \min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s. l.} \; x \in Q.$$

Dacă f diferenţiabilă atunci:

$$\begin{aligned} f(x - \alpha \mathcal{G}(x;\alpha)) &= f(x) - \alpha \nabla f(x)^T \mathcal{G}(x;\alpha) + o(-\alpha \mathcal{G}(x;\alpha)) \\ &\leq f(x) - \alpha \frac{1}{2} \|\mathcal{G}(x;\alpha)\|_2^2 + o(-\alpha \mathcal{G}(x;\alpha)) \\ &\leq f(x) - \alpha \|\mathcal{G}(x;\alpha)\|_2 \left[\frac{1}{2} \|\mathcal{G}(x;\alpha)\|_2 - \frac{o(-\alpha \mathcal{G}(x;\alpha))}{\alpha \|\mathcal{G}(x;\alpha)\|_2} \right]. \end{aligned}$$

Un pas α suficient de mic asigură: $f(x - \alpha \mathcal{G}(x;\alpha)) < f(x)$, de aceea orice minim x^* satisface $\mathcal{G}(x^*;\alpha) = 0$.

$$(OPC): \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s. l.} \;\; x \in Q.$$

Teoremă

Fie f diferenţiabilă, Q mulţime convexă închisă şi α o constantă pozitivă. Orice minim x^* al problemei OPC satisface:

$$x^* = \pi_Q(x^* - \alpha \nabla f(x^*)) \;\; \Leftrightarrow \mathcal{G}(x^*; \alpha) = 0$$

Mai mult, dacă f convexă atunci x^* soluţia OPC dacă şi numai dacă $\mathcal{G}(x^*; \alpha) = 0$.

- Analog condiţiilor de ordin I pentru cazul neconstrâns (i.e. $Q = \mathbb{R}^n$)

$$(OPC): \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s. l.} \;\; x \in Q.$$

Teoremă

Fie f diferenţiabilă, Q mulţime convexă închisă şi α o constantă pozitivă. Orice minim x^* al problemei OPC satisface:

$$x^* = \pi_Q(x^* - \alpha \nabla f(x^*)) \quad \Leftrightarrow \mathcal{G}(x^*; \alpha) = 0$$

Mai mult, dacă f convexă atunci x^* soluţia OPC dacă şi numai dacă $\mathcal{G}(x^*; \alpha) = 0$.

- Analog condiţiilor de ordin I pentru cazul neconstrâns (i.e. $Q = \mathbb{R}^n$)
- Condiţii necesare în general, dar nec. şi suficiente în cazul f convexă

$$(OPC): \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s. l. } \; x \in Q.$$

Teoremă

Fie f diferenţiabilă, Q mulţime convexă închisă şi α o constantă pozitivă. Orice minim x^* al problemei OPC satisface:

$$x^* = \pi_Q(x^* - \alpha \nabla f(x^*)) \quad \Leftrightarrow \mathcal{G}(x^*; \alpha) = 0$$

Mai mult, dacă f convexă atunci x^* soluţia OPC dacă şi numai dacă $\mathcal{G}(x^*; \alpha) = 0$.

- Analog condiţiilor de ordin I pentru cazul neconstrâns (i.e. $Q = \mathbb{R}^n$)
- Condiţii necesare în general, dar nec. şi suficiente în cazul f convexă
- Cand π_Q este tractabil, condiţia de mai sus este verificabilă!

$$\max_{x} \ \|Ax\|_2^2 \ \text{ s.l. } \ \|x\| \le 1$$

$$\max_{x} \ \|Ax\|_2^2 \ \text{ s.l. } \ \|x\| \leq 1$$

$$\begin{aligned} x^* &= \pi_Q(x^* - \nabla f(x^*)) = \pi_Q(x^* - A^T A x^*) \\ x^* &= \frac{1}{\|(I - A^T A)x^*\|}(I - A^T A)x^* \\ x^* &= \frac{1}{1 - \|(I - A^T A)x^*\|} A^T A x^* = \lambda^* A^T A x^*. \end{aligned}$$

$$\max_x \ \|Ax\|_2^2 \ \text{ s.l. } \ \|x\| \leq 1$$

$$x^* = \pi_Q(x^* - \nabla f(x^*)) = \pi_Q(x^* - A^T A x^*)$$

$$x^* = \frac{1}{\|(I - A^T A)x^*\|}(I - A^T A)x^*$$

$$x^* = \frac{1}{1 - \|(I - A^T A)x^*\|}A^T A x^* = \lambda^* A^T A x^*.$$

Puncte staţionare: vectorii proprii asociaţi $A^T A$.
Soluţie: vectorul propriu x^* asociat $\lambda_{max}(A^T A)$.

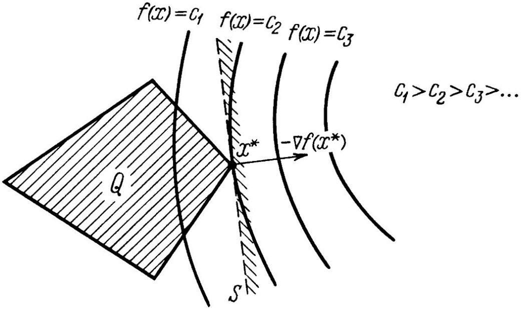
Condiţia de ordin I:

$$x^* = \pi_Q(x^* - \alpha \nabla f(x^*)) \quad \Leftrightarrow \mathcal{G}(x^*; \alpha) = 0$$

este echivalentă cu:

$$\nabla f(x^*)^T (x - x^*) \geq 0, \qquad \forall x \in Q.$$

Interpretare: *Direcţia* $-\nabla f(x^*)$ *face un unghi obtuz cu direcţiile fezabile*



[Polyak, page 229]

Exemplu curs 1:

$$
\begin{aligned}
\min \ f(x) :=& (x_1 + x_2)^2 \\
\text{s.t. } & 0 \leq x_1 \leq 1 \\
& 0 \leq x_2 \leq 1
\end{aligned}
$$

Exemplu curs 1:

$$\min \ f(x) := (x_1 + x_2)^2$$
$$\text{s.t. } 0 \le x_1 \le 1$$
$$0 \le x_2 \le 1$$

$$x^* = \pi_{[0,1]^2}(x^* - \nabla f(x^*))$$
$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} \pi_{[0,1]}(-x_2^*) \\ \pi_{[0,1]}(-x_1^*) \end{bmatrix}$$
$$x^* = 0.$$

$$\min_{x\in\mathbb{R}^3} \; f(x) = x_1^2 + x_2^2 + x_3^2 + (x_1 + 2x_2 - 4)^2$$
$$\text{s.l.} \; \|x\| \leq 1.$$

Care sunt condiţiile de optimalitate de ordin I?

Exemple

$$\begin{aligned} \min\ & f(x) := c^T x \\ \text{s.t. } & x \in [l, u]. \end{aligned}$$

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & x \in [l, u]. \end{aligned}$$

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & Ax = b. \end{aligned}$$

- Probleme de minimizare cu constrângeri (POC). Condiţii de optimalitate.
- **Algoritmi de ordinul I pentru POC. Convergenţă**
- Exemple şi aplicaţii

Din relaţiile de mai sus, pentru α suficient de mic obţinem descreşterea:

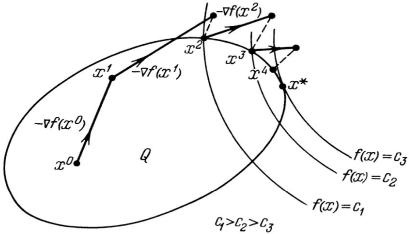
$$f(x - \alpha\mathcal{G}(x)) < f(x)$$

care sugerează iteraţia $x^+ = x - \alpha\mathcal{G}(x)$.

Calcularea operatorlui $\mathcal{G}(\cdot)$ foloseşte $\pi_Q(\cdot)$, care este tractabil doar în anumite cazuri particulare ”simple”: bilă, hipercub, hiperplan etc.

Metoda Gradientului Proiectat: inițial $x^0 \in Q$

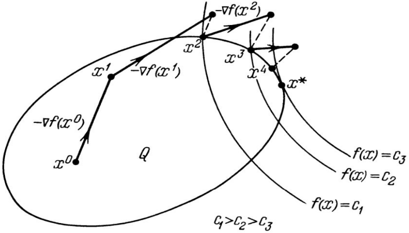
$$x^{k+1} := x^k - \alpha\mathcal{G}(x^k) = \pi_Q(x^k - \alpha\nabla f(x^k))$$



[Polyak, pg. 234]

Interpretare:

$$\begin{aligned} x^{k+1} &= \operatorname*{argmin}_z \; f(x^k) + \nabla f(x^k)^T(z - x^k) + \frac{1}{2\alpha_k}\|z - x^k\|^2 \\ &= \operatorname*{argmin}_z \; \frac{1}{2}\|z - (x^k - \alpha_k \nabla f(x^k))\|^2 \;\; \text{(exercitiu!)} \end{aligned}$$



[Polyak, pg. 234]

Algorithm 1: Metoda Gradientului Proiectat (MGP) $(x^0, \epsilon, \{\alpha_k\}_{k\geq 0})$:

Data: $k := 0$

while criteriu oprire = fals **do**
- Calculează: $\nabla f(x^k)$
- $x^{k+1} = \pi_Q(x^k - \alpha_k \nabla f(x^k))$
- $k := k + 1$

end

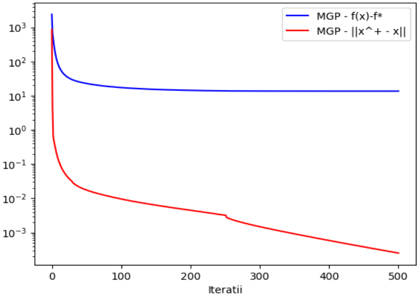
Teoremă

Fie f o funcţie diferenţiabilă cu gradient Lipschitz şi Q mulţimea fezabilă convexă. Pentru pasul $\alpha_k = \alpha \in (0, 2/L)$, şirul MGP satisface:

$$\|\mathcal{G}(x^k)\| \to 0 \ \text{ când } \ k \to \infty.$$

Mai mult, dacă f convexă şi X^* mulţimea optimă POC, atunci $x^k \to x^* \in X^*$.

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax-b\|^2 \text{ s.t. } l \leq x \leq u.$$



$$\min_{x\in\mathbb{R}^3} \; f(x) = x_1^2 + x_2^2 + x_3^2 + (x_1 + 2x_2 - 4)^2$$
$$\text{s.l.} \; \|x\| \leq 1.$$

Calculaţi prima iteraţie a MGP cu pas constant $\alpha = 1/L$, pornind din $x^0 = [0\ 0\ 0]$.

Teoremă (Nesterov, 2014)

Fie f o funcţie diferenţiabilă cu gradient Lipschitz cu constanta L, tare-convexă cu constanta σ, iar Q convexă şi închisă. Pentru pasul $\alpha \in \left(0, \frac{2}{L+\sigma}\right]$ avem:

$$\|x^k - x^*\| \leq \left(1 - \frac{\sigma}{L}\right)^k \|x^0 - x^*\|$$

- Performanţă similară cu cea a MG
- $x^k \in S_f(f^* + \epsilon)$ după $\mathcal{O}\left(\kappa \log(1/\epsilon)\right)$ iteraţii

Ipoteza: Este posibilă calcularea eficientă a operatorului cost liniar:

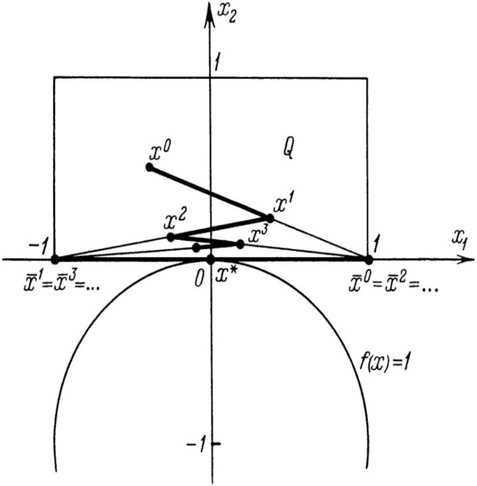
$$\phi_Q(y) := \operatorname*{argmin}_{z \in Q} \; y^T z$$

Ipoteza: Este posibilă calcularea eficientă a operatorului cost liniar:

$$\phi_Q(y) := \operatorname*{argmin}_{z \in Q} \; y^T z$$

Metoda Gradientului Condiţional realizează un model aproximativ liniar al costului şi foloseşte soluţia acestuia la fiecare iteraţie k:

$$\begin{aligned} \phi_Q(x^k) &: = \operatorname*{argmin}_{z \in Q} \; \nabla f(x^k)^T z \\ x^{k+1} &= x^k + \alpha_k(\phi_Q(x^k) - x^k). \end{aligned}$$


x_2
1
x^0
Q
x^1
x^2
x^3
-1
1
x_1
$\bar{x}^1 = \bar{x}^3 = \ldots$
0
x^*
$\bar{x}^0 = \bar{x}^2 = \ldots$
$f(x)=1$
-1

Metoda Gradientului Condiţional

Algorithm 2: Metoda Gradientului Condiţional (MGC) $(x^0, \epsilon, \{\alpha_k\}_{k\geq 0})$:

Data: $k := 0$

while criteriu oprire = fals **do**

  Calculează: $\phi_Q(x^k) := \underset{z\in Q}{\text{argmin}} \;\; \nabla f(x^k)^T z$

  $x^{k+1} = x^k + \alpha_k(\phi_Q(x^k) - x^k)$

  $k := k + 1$

end

Teoremă

Fie f o funcţie diferenţiabilă cu gradient Lipschitz şi Q mulţimea fezabilă convexă. Pentru pasul $\alpha_k = \text{argmin}_{\alpha\in[0,1]} f(x^k + \alpha(\phi_Q(x^k) - x^k))$, şirul MGC satisface:

$$\nabla f(x^k)^T(x^k - \phi_Q(x^k)) \to 0 \;\; \text{când} \;\; k \to \infty.$$

Mai mult, dacă f convexă atunci $f(x^k) - f^* = \mathcal{O}(1/k)$.

$$\min_{x\in\mathbb{R}^3} \; f(x) = x_1^2 + x_2^2 + x_3^2 + (x_1 + 2x_2 - 4)^2$$
$$\text{s.l.} \; \|x\| \leq 1.$$

Calculaţi prima iteraţie a MGC cu pas constant $\alpha = 1/L$, pornind din $x^0 = [0\ 0\ 0]$.

Calculaţi soluţia următoarelor probleme:

- $\arg\min_{x \in \mathbb{R}^n} \; c^T x \;$ s.l. $\; l \leq x \leq u$
- $\arg\min_{x \in \mathbb{R}^n} \; c^T x \;$ s.l. $\; \|x\| \leq r$
- $\arg\min_{x \in \mathbb{R}^n} \; c^T x \;$ s.l. $\; \frac{1}{2} x^T H x - q^T x \leq \beta \;\; (\beta > 0, H \succ 0)$

Metoda Gradientului Accelerat (Nesterov): fie $x^0 \in \mathbb{R}^n, y^0 = x^0, \beta = \frac{1-\sqrt{\kappa_f}}{1+\sqrt{\kappa_f}}$

1. $x^k = \pi_Q \left(y^k - \frac{1}{L} \nabla f(y^k) \right)$
2. $y^{k+1} = x^{k+1} + \beta(x^{k+1} - x^k)$

- **Accelerare**: $\mathcal{O}\left(\sqrt{\kappa_f} \log(1/\epsilon)\right)$ *vs* $\mathcal{O}\left(\kappa_f \log(1/\epsilon)\right)$ *(MG)*

Metoda Gradientului Accelerat (Nesterov): fie $x^0 \in \mathbb{R}^n, y^0 = x^0, \beta = \frac{1-\sqrt{\kappa_f}}{1+\sqrt{\kappa_f}}$

1. $x^k = \pi_Q\left(y^k - \frac{1}{L}\nabla f(y^k)\right)$
2. $y^{k+1} = x^{k+1} + \beta(x^{k+1} - x^k)$

- **Accelerare**: $\mathcal{O}\left(\sqrt{\kappa_f}\log(1/\epsilon)\right)$ *vs* $\mathcal{O}\left(\kappa_f \log(1/\epsilon)\right)$ (*MG*)
- Iteraţia MGA are acelaşi ordin de complexitate per iteraţie ca MG

Metoda Gradientului Accelerat (Nesterov): fie $x^0 \in \mathbb{R}^n, y^0 = x^0, \beta = \frac{1-\sqrt{\kappa_f}}{1+\sqrt{\kappa_f}}$

1. $x^k = \pi_Q\left(y^k - \frac{1}{L}\nabla f(y^k)\right)$
2. $y^{k+1} = x^{k+1} + \beta(x^{k+1} - x^k)$

- **Accelerare**: $\mathcal{O}\left(\sqrt{\kappa_f}\log(1/\epsilon)\right)$ *vs* $\mathcal{O}\left(\kappa_f \log(1/\epsilon)\right)$ (*MG*)
- Iteraţia MGA are acelaşi ordin de complexitate per iteraţie ca MG
- MGA este mai puţin robustă la erori

- Metoda Newton Proiectat
- Metoda Gradienţilor Conjugaţi (cu proiecţie)
- Metode Direcţii Descreştere