Tehnici de Optimizare

Facultatea de Matematica si Informatica

Universitatea Bucuresti

- Department Informatica-

2021

Metoda Newton – convergenta globala

- [**Convergenta globala**]. Fie f dublu diferentiabila, convexa, cu $\nabla^2 f(x) \succ 0$ si $S(x^0) = \{x \colon f(x) \leq f(x^0)\}$ marginita. Atunci Metoda Newton cu pas **ideal** genereaza un sir x^k convergent la minimul global unic x^***.**
- Rezultate de **convergenta globala** pentru MN cu pas backtracking!
- Convergenta locala (vezi cursul trecut) se refera la comportamentul MN intr-o vecinatate a minimului local. In acest caz, trebuie asigurata atingerea acestei vecinatati de catre traiectoria MN.

Probleme de optimizare cu constrangeri

$$\min_{\mathbf{x}} \boldsymbol{f}(\boldsymbol{x})$$
$$\boldsymbol{x} \in \boldsymbol{Q}$$

- f functie cost/obiectiv
- Q multime fezabila
- Presupunem Q convexa si simpla (anumite “obiecte” se calculeaza usor, e.g. proiectia ortogonala)

Probleme de optimizare cu constrangeri

Conditii necesare de optimalitate

THEOREM 1 (necessary first-order minimum condition). Let $f(x)$ be differentiable at the minimum point x^*, and let Q be a convex set. Then

$$(\nabla f(x^*), x - x^*) \geq 0 \qquad \text{for all } x \in Q . \tag{1}$$

Probleme de optimizare cu constrangeri

Conditii necesare de optimalitate

PROOF. Let $(\nabla f(x^*), x^0 - x^*) < 0$ for some $x^0 \in Q$. Then $x(\alpha) = x^* + \alpha(x^0 - x^*) \in Q$ for $0 \le \alpha \le 1$ by the convexity of Q and

$$f(x(\alpha)) = f(x^*) + \alpha(\nabla f(x^*), x^0 - x^*) + o(\alpha) < f(x^*)$$

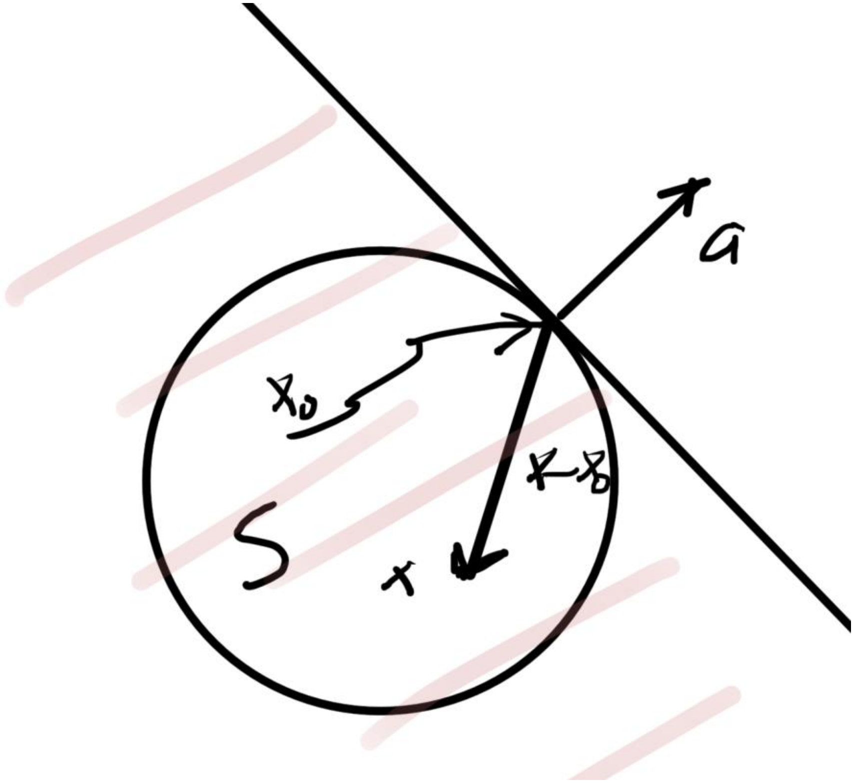
for sufficiently small $\alpha > 0$, which contradicts the local optimality of x^*. □

Reamintim: f diferentiabila, $f(x) = f(x^*) + \nabla f(x^*)^T(x - x^*) + o\left(||x - x^*||\right)$

$$\alpha\left(\nabla f(x^*)^T(x^0 - x^*) - \frac{o(\alpha)}{\alpha}\right) < 0 \quad pentru\ \alpha\ mic$$

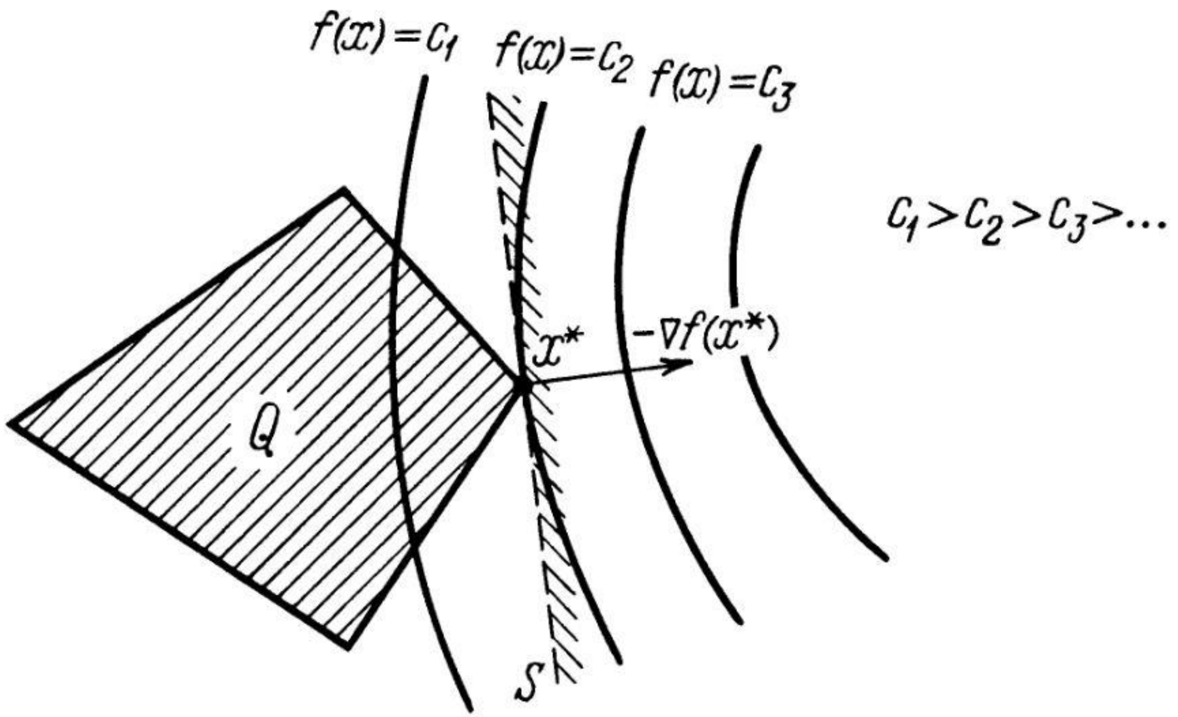
Probleme de optimizare cu constrangeri

Un vector $a \in R^n$ care satisfice $a^T(x - x^*) \leq 0$ pentru toti $x \in Q$ se numeste **hiperplan de suport** al multimii Q in punctul x^*



Probleme de optimizare cu constrangeri

Conditii necesare: $-\nabla f(x^*)$ reprezinta un vector de suport al multimii Q in x^*



Probleme de optimizare cu constrangeri

Conditii suficiente de optimalitate

THEOREM 2 (sufficient first-order minimum condition). Let $f(x)$ be differentiable at the point $x^* \in Q$, let Q be convex and let the condition

$$(\nabla f(x^*), x - x^*) \geq \alpha \|x - x^*\|, \qquad \alpha > 0, \tag{2}$$

be satisfied for all $x \in Q$, $\|x - x^*\| \leq \varepsilon$, $\varepsilon > 0$. Then x^* is a local minimum point of $f(x)$ on Q.

Exemple

$$\min_{x} \quad f(x)$$
$$s.l. \, \mathrm{x} \in \mathrm{Q} = \{x \in R^n : l \le x \le u\}$$

Conditii necesare: $\boldsymbol{\nabla f(x^*)^T (x - x^*) = \sum_k \nabla_k f(x^*)(x_k - x_k^*) \ge 0}, \quad \forall \boldsymbol{x} \in \boldsymbol{Q}$

Alegem $x \in Q$ astfel: $x_j = x_j^*$ pentru $j \ne i$, atunci avem

$$\boldsymbol{\nabla_i f(x^*)(x_i - x_i^*) \ge 0}, \qquad \forall \boldsymbol{x_i} \in [\boldsymbol{l_i}, \boldsymbol{u_i}], \forall \boldsymbol{i} = \boldsymbol{1}, \cdots, \boldsymbol{n}$$

- $x_i^* \in (l_i, u_i)$, implica $\nabla_i f(x^*) = 0$
- $x_i^* = l_i$ implica $\nabla_i f(x^*) \ge 0$
- $x_i^* = u_i$ implica $\nabla_i f(x^*) \le 0$

Exemple

$$\min_{x} \quad f(x)$$
$$s.l.\, \mathrm{x} \in \mathrm{Q} = \{x \in R^n : l \leq x \leq u\}$$

Conditii necesare: $\nabla_i f(x^*) = \begin{cases} = 0, x_i^* \in (l_i, u_i) \\ \geq 0, x_i^* = l_i \\ \leq 0, x_i^* = u_i \end{cases}$

Conditii suficiente: Daca $\mathrm{x_i^*} = \mathrm{l_i}$ sau $\mathrm{u_i}$, $\quad \nabla_i f(x^*) = \begin{cases} > 0, x_i^* = l_i \\ < 0, x_i^* = u_i \end{cases},$

$$atunci\ x^*\ minim\ local$$

Probleme convexe

$$\min_{x} \quad f(x)$$
$$s.l.\, \mathrm{x} \in \mathrm{Q}$$

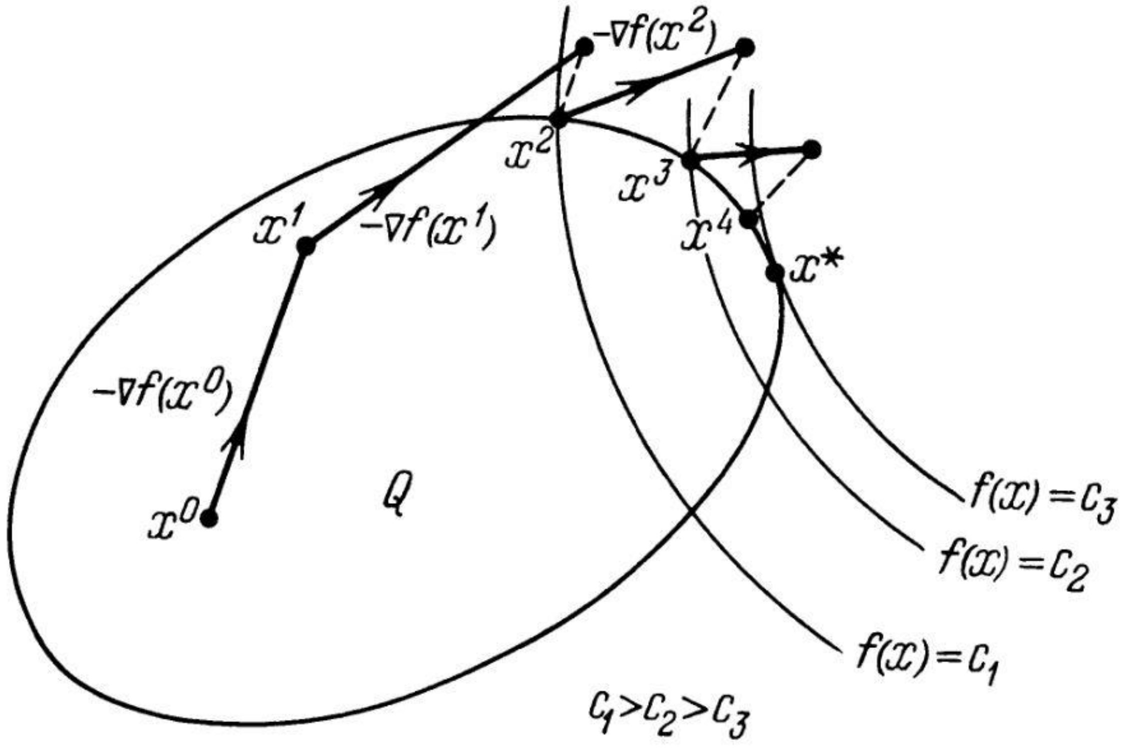
Conditii necesare si suficiente: $\nabla f(x^*)^T(x - x^*) \geq 0, \forall x \in Q$

- In cazul convex conditiile necesare sunt si suficiente
- Diferentiabilitatea nu este necesara (gradientul se generalizeaza prin subgradient)

Metoda Gradient Proiectat

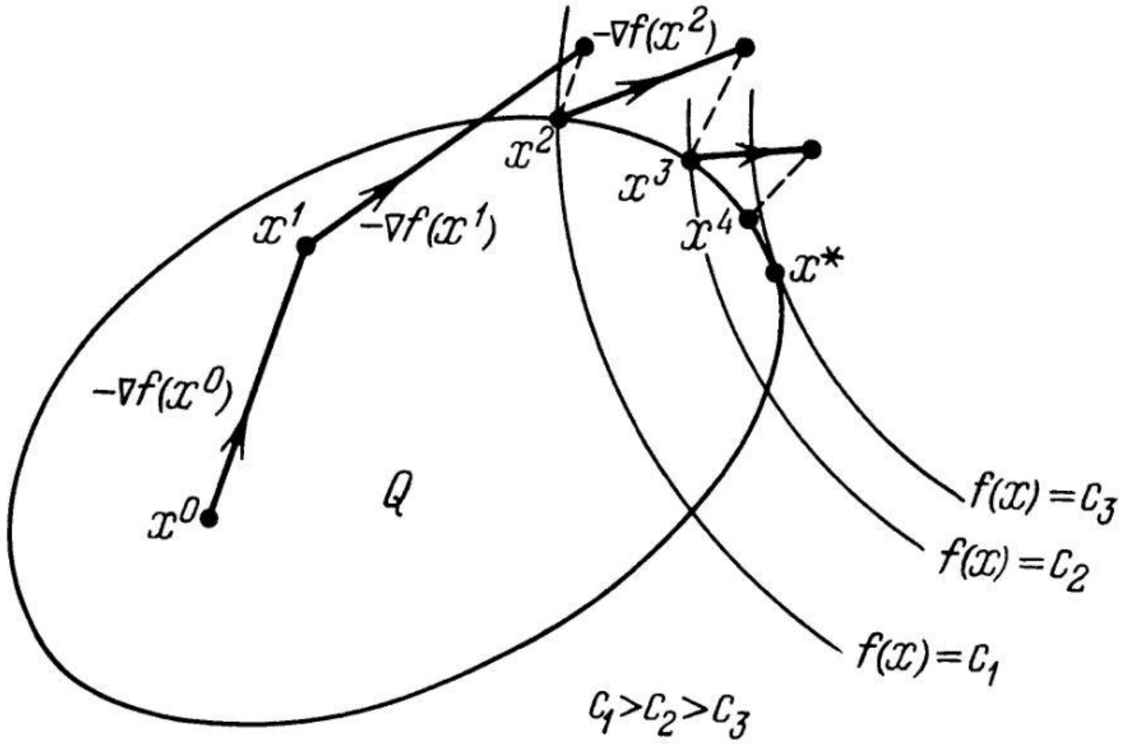
$$x^{k+1} = \pi_Q(x^k - \alpha_k \nabla f(x^k))$$

unde $\pi_Q(\cdot)$ reprezinta operatorul de proiectie ortogonala pe Q



Metoda Gradient Proiectat

$$x^{k+1} = \arg\min_{x \in Q} f\left(x^k\right) + \nabla f\left(x^k\right)^T \left(x - x^k\right) + \frac{1}{\alpha_k} \left\|x - x^k\right\|^2$$



Metoda Gradient Proiectat

THEOREM 1. Let $f(x)$ be a convex differentiable function in $\mathbf{R}^n$ whose gradient satisfies a Lipschitz condition with constant L on Q. Let Q be convex and closed, $x^* = \operatorname{Argmin}_{x \in Q} f(x) \neq \emptyset$ and $0 < \gamma < 2/L$. Then

(i) $x^k \to x^* \in X^*$;

(ii) if $f(x)$ is strongly convex, then $x^k \to x^*$ with the rate of geometric progression;

(iii) if $f(x)$ is twice differentiable and $\ell I \leq \nabla^2 f(x) \leq LI$, $x \in Q$, $\ell > 0$, then the progression ratio is $q = \max\{|1 - \gamma\ell|, |1 - \gamma L|\}$;

Metoda Gradient Proiectat - exemple

- $Q = \{x \in R^n : x \geq 0\},\ x^{k+1} = \max\left\{\mathbf{0}_{\boldsymbol{n}}, x^k - \alpha_k \nabla f\left(x^k\right)\right\}$

- $Q = \left\{x \in R^n : \left||x|\right| \leq r\right\},$

$$x^{k+1} = \begin{cases} x^k - \alpha_k \nabla f\left(x^k\right), & pt. \left\|x^k - \alpha_k \nabla f\left(x^k\right)\right\| \leq r \\ r \dfrac{x^k - \alpha_k \nabla f\left(x^k\right)}{\left||x^k - \alpha_k \nabla f(x^k)|\right|}, & pt. \left\|x^k - \alpha_k \nabla f\left(x^k\right)\right\| > r \end{cases}$$

Metoda Gradient Conditional

$$y^k = \arg\min_{x \in Q} \nabla f\left(x^k\right)^T x \; = \arg\min_{\mathrm{x} \in \mathrm{Q}} f\left(x^k\right) + \nabla f\left(x^k\right)^T \left(x - x^k\right)$$
$$x^{k+1} = x^k + \alpha_k (y^k - x^k)$$

- Se aproximeaza functia obiectiv $f(x)$ cu $f\left(x^k\right) + \nabla f\left(x^k\right)^T (x - x^k)$
- Se minimizeaza la fiecare iteratie modelul linear
- Pentru f convexa, L.c.g., se arata: $f\left(x^k\right) - f^* = O\left(\frac{1}{k}\right)$
- Convergenta pentru orice multime convexa Q?

Metoda Gradient Conditional

- Cand exista solutie pentru:
- $min_{x\geq 0}\, c^T x = min_{x\geq 0}\, \sum_i c_i x_i = \sum_i min_{x_i\geq 0} c_i x_i = 0, daca\ c \geq 0$
- Daca c $\geq$ 0 atunci $\min_{x\geq 0} cx = 0, altfel\ nu\ are\ solutie$
- $\min_{l\leq x\leq u} c^T x = \sum_i min_{l_i\leq x_i\leq u_i} c_i x_i$
- $\min_{l_i\leq x\leq u_i} cx = cl_i\ (daca\ c \geq 0), altfel \min_{l_i\leq x\leq u_i} cx = cu_i$
- MGC are sens pentru multimi fezabile marginite!

Metoda Newton Proiectat

Se aproximeaza functia obiectiv cu modelul patratic Taylor de ordin II:

$$x^{k+1} = \min_{x \in Q} f\left(x^k\right) + \nabla f\left(x^k\right)^T \left(x - x^k\right) + \frac{1}{2}\left(x - x^k\right)^T \nabla^2 f(x^k)(x - x^k)$$

THEOREM 4. Let $f(x)$ attain a minimum on a closed convex set Q at a point x^*, at which $f(x)$ is twice differentiable on Q in a neighborhood of x^*, let $\nabla^2 f(x)$ satisfy a Lipschitz condition, and let

$$\nabla^2 f(x^*) > 0 . \tag{18}$$

Then method (17) converges locally to x^* with quadratic rate.

Metoda Newton Proiectat

Se aproximeaza functia obiectiv cu modelul patratic Taylor de ordin II:

$$x^{k+1} = \min_{x \in Q} f\left(x^k\right) + \nabla f\left(x^k\right)^T \left(x - x^k\right) + \frac{1}{2}\left(x - x^k\right)^T \nabla^2 f(x^k)(x - x^k)$$

- Q poliedru, exista algoritmi in timp finit pentru subproblema Newton
- Q box, se rezolva cu metoda gradientilor conjugati
- Q bila sau subspatiu liniar, subproblema are solutie simpla