Tehnici de Optimizare

Facultatea de Matematica si Informatica

Universitatea Bucuresti

- Department Informatica-

2021

Proiect

- Alegere 1-2 lucrari; maxim 2 alegeri per lucrare
- Documentatie (rezultate ale lucrarii); limita pagini: 5-20
- Simulari (algoritmi din lucrare); grafice de analiza
- Echipa: maxim 2 - 3 studenti
- Pondere nota: 60%
- Termen: in sesiune
- Evaluare: online fata-in-fata

Algoritm de liniarizare (constrangeri egalitate)

$$\min_x f(x)$$
$$s.l.\, g_i(x) = 0, \quad i = 1, \cdots, m$$

Reamintim: ***g* neliniar face multimea fezabila neconvexa!**

Idee algoritm liniarizare:

1. Aproximeaza g liniar in jurul lui x^k; suprafata $\{x:\ g(x) = 0\}$ este aprox. cu un subspatiu liniar $\{x: g(x^k) + \nabla \mathrm{g}(\mathrm{x}^\mathrm{k})(\mathrm{x} - \mathrm{x}^\mathrm{k}) = 0\}$;

$$\min_x f(x)$$
$$s.l.\, g_i(x^k) + \nabla \mathrm{g}_i(\mathrm{x}^\mathrm{k})(\mathrm{x} - \mathrm{x}^\mathrm{k}) = 0, \quad i = 1, \ldots, m$$

Algoritm de liniarizare

$$\min_x f(x)$$
$$s.l.\, g(x) = 0$$

2. Aplica o iteratie de MG pe noul model creat:

$$x^{k+1} = arg\min_x f\left(x^k\right) + \nabla f\left(x^k\right)^T (x - x^k) + \frac{1}{\alpha_k}\left|\left|x - x^k\right|\right|^2$$
$$s.l.\, g\left(x^k\right) + \nabla \mathrm{g}\left(\mathrm{x^k}\right)\left(\mathrm{x} - \mathrm{x^k}\right) = 0$$

$$\boldsymbol{x^{k+1}} = \boldsymbol{\pi_{S_k}}[\boldsymbol{x^k} - \boldsymbol{\alpha_k \nabla f}(\boldsymbol{x^k})], \quad S_k = \{x\colon g\left(x^k\right) + \nabla \mathrm{g}\left(\mathrm{x^k}\right)\left(\mathrm{x} - \mathrm{x^k}\right) = 0\}$$

Algoritm de liniarizare

2. Aplica o iteratie de MG pe noul model creat:

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

$$s.l.\ g(x^k) + \nabla \mathrm{g}(\mathrm{x}^\mathrm{k})(\mathrm{x} - \mathrm{x}^\mathrm{k}) = 0$$

Subproblema convexa! $\nabla g(x^k) \in R^{m \times n}$

C.N. – E. ne spune ca solutia rezulta din sistemul linear de ecuatii :

$$\nabla f(x^k) + \frac{1}{\alpha_k}(x - x^k) + \nabla g(x^k)^T \lambda = 0$$

$$g(x^k) + \nabla \mathrm{g}(\mathrm{x}^\mathrm{k})(\mathrm{x} - \mathrm{x}^\mathrm{k}) = 0$$

Algoritm de liniarizare (constrangeri egalitate)

THEOREM 1. Let x^* be a nonsingular minimum point, and let $\nabla^2 f(x)$, $\nabla^2 g_i(x)$ satisfy a Lipschitz condition in a neighborhood of x^*. Then we can find a $\bar{\gamma} > 0$ such that for $0 < \gamma < \bar{\gamma}$ method (1) is well-defined and converges locally to x^* with the rate of geometric progression.

- Presupuneri: pentru oricare x, y

$$\left\| \nabla^2 f(x) - \nabla^2 f(y) \right\| \leq L||x - y||$$

$$\left\| \nabla^2 g_i(x) - \nabla^2 g_i(y) \right\| \leq L_i||x - y||$$

- **Convergenta locala (tipic)!**
- x^* nesingular, i.e. $\nabla^2 f(x^*) \succ 0$

Algoritm dual

$$\min_x f(x)$$
$$s.l.\, g(x) = 0$$

Functia Lagrangian $L(x, \lambda)$ este definita:

$$L(x, \lambda) = f(x) + \lambda^T g(x) = f(x) + \sum_i \lambda_i g\_i(x)$$

Observatii:

1. **C.N.-E. se reduc la C.N. pe functia Lagrangian!**
2. Intuitiv: Pentru $\mathrm{sgn}(\lambda) = \mathrm{sgn}\big(g(x)\big)$, $||\lambda|| \to \infty$, avem $\lambda^T g(x)$ pozitiv si dominant, deci solutia $\operatorname{argmin}_x L(x, \lambda)$ va, in primul rand, asigura valori mici pentru $g(x)$

Idee algoritm dual:

Aplicati Metoda Gradient (fara cons.) pentru minimizarea $L(\cdot, \lambda)$ si maximizarea $L(x,\cdot)$

Algoritm dual

$$\min_x f(x)$$
$$s.l.\, g(x) = 0$$

Functia Lagrangian $\boldsymbol{L(x, \lambda)}$ este definita:

$$L(x, \lambda) = f(x) + \lambda^T g(x) = f(x) + \sum_i \lambda_i g_i(x)$$

Algoritm dual:

1. $x^{k+1} = x^k - \alpha \nabla_x L(x^k, \lambda^k)$ (pas spre $\min_x L(x, \lambda)$, cu λ fixat)
2. $\lambda^{k+1} = \lambda^k + \alpha \nabla_\lambda L\left(x^k, \lambda^k\right)$ (pas spre $max_\lambda\, L(x, \lambda)$, cu x fixat)

Algoritm dual

$$\min_x f(x)$$
$$s.l.\, g(x) = 0$$

Functia Lagrangian $\boldsymbol{L(x,\lambda)}$ este definita:

$$L(x,\lambda) = f(x) + \lambda^T g(x) = f(x) + \sum_i \lambda_i g\_i(x)$$

Algoritm dual:

1. $x^{k+1} = x^k - \alpha \nabla_x L\left(x^k, \lambda^k\right) = x^k - \alpha[\nabla f\left(x^k\right) + \sum_i \lambda_i^{\mathrm{k}} \nabla g_i\left(x^k\right)]$
2. $\lambda^{k+1} = \lambda^k + \alpha \nabla_\lambda L\left(x^k, \lambda^k\right) = \lambda^k + \alpha g(x^k)$

Algoritm dual

Functia Lagrangian $\boldsymbol{L(x,\lambda)}$ este definita:

$$L(x,\lambda) = f(x) + \lambda^T g(x) = f(x) + \sum_i \lambda_i g\_i(x)$$

Daca cunoastem *a priori* multiplicatorii Lagrange λ^**, atunci:*

Algoritm dual:

1. $x^{k+1} = x^k - \alpha \nabla_x L\left(x^k, \lambda^*\right) = x^k - \alpha\left[\nabla f\left(x^k\right) + \sum_i \lambda_i^* \nabla g_i\left(x^k\right)\right]$

Echivalent, de fapt, cu MG pentru $\min_x L(x, \lambda^*)$

Algoritm dual

THEOREM 2. Let x^* be a nonsingular minimum point, let $L''_{xx}(x^*, y^*) > 0$ and let the second derivatives $\nabla^2 f(x)$, $\nabla^2 g_i(x)$ satisfy a Lipschitz condition in a neighborhood of x^*. Then we can find a $\bar{\gamma} > 0$ such that for $0 < \gamma < \bar{\gamma}$ method (6) converges locally to x^*, y^* with the rate of geometric progression.

Similar cu rezultatul AL, cu exceptia $\nabla^2_{xx} L(x^*, \lambda^*) \succ 0$!

Algoritm de penalitate

O alta reformulare a problemei cu constrangeri: $\rho > 0$

$$\min_x f(x) + \frac{\rho}{2}||g(x)||^2$$

cu solutia x_ρ, unde ρ parametru de penalitate.

Observati:

- Cand $\rho \to 0$ se ignora constrangerile
- Cand $\rho \to \infty$ avem fezabilitate $g(x_\rho) \to 0$

Idee AP: creaza $x^k = arg\min_x f(x) + \frac{\rho_k}{2}||g(x)||^2; \rho_{k+1} = 2\rho_k$

unde sirul $\rho_k \to \infty$. Deci, sub anumite conditii, vom avea $x^k \to x^*$

Algoritm de penalitate

- Alegem Q_0 o multime initiala in jurul punctului de optim
- Notam: $f_k(x) = f(x) + \frac{\rho_k}{2}||g(x)||^2$

THEOREM 6. Let problem (A) have solutions, and let X^* denote the set of all these solutions. Let f and g_i be continuous, and let Q_0 be closed and bounded, $Q_0 \cap X^* \neq \emptyset$. Then every limit point of method (13) (by x^k we mean the global minimum of $f_k(x)$ on Q_0) is a global minimum for problem (A).

- Rezultat mai general decat restul metodelor de pana acum! (nu necesita diferentiabilitate)
- Se bazeaza ca putem rezolva subproblema!

Probleme de programare neliniara

$$
\begin{aligned}
&\min_x f(x) \\
s.l.\ & g_i(x) = 0, \qquad i = 1, \cdots, m \\
& h_i(x) \le 0, \qquad i = 1, \cdots, p
\end{aligned}
$$

unde f, g, h sunt functii diferentiabile.

Problema convexa: $\boldsymbol{f}$ convexa + $\boldsymbol{h_i}$ convexe + $\boldsymbol{g_i}$ liniare!

Exemplu convex:

$$
\begin{aligned}
&\min_x \frac{1}{2} \left\| Ax - b \right\|^2 \\
s.l.\ & \quad Cx \le d
\end{aligned}
$$

Probleme de programare neliniara

$$\min_x f(x)$$
$$s.l.\ g_i(x) = 0, \qquad i = 1, \cdots, m$$
$$h_i(x) \leq 0, \qquad i = 1, \cdots, p$$

Notam $I(x) = \{i: h_i(x) = 0 (< 0)\}$, indicii inegalitatilor **active (inactive)**.

Functia Lagrangian: $L: R^n \times R^m \times R_+^p$

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x); L(x^*, \lambda^*, \mu^*) = \cdots + \sum \mu_i^* h_i(x^*)$$

Egalitatile sunt penalizate cand $||\lambda|| \to \infty$

Inegalitatile sunt penalizate cand $\mu \geq 0\ (\mu^T h(x) \geq 0)\ si\ ||\mu|| \to \infty$

$$h(x) = [h_1(x) \quad h_2(x)] = [(< 0) \quad (> 0)] \Rightarrow \mu^T h(x) = \mu_1 h_1(x) + \mu_2 h_2(x)$$

Probleme de programare convexa

$$\min_x f(x)$$
$$s.l.\ Ax = b,$$
$$h(x) \leq 0,$$

unde f, h_i sunt functii convexe.

Teorema Kuhn-Tucker. Fie f, h_i convexe si conditia Slater:

$$h_i(x^*) < 0, \qquad i = 1, \cdots, m, \qquad pentru\ h_i\ neliniare$$

Atunci x^* este o solutie a problemei daca si numai daca: gasim $\mu^* \geq 0$

$\mu_i^* h_i(x^*) = 0, \quad i = 1, \cdots, m$ si $L(x, \lambda^*, \mu^*) \geq L(x^*, \lambda^*, \mu^*)$ sau $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$

Probleme de programare convexa

$$\min_x f(x)$$
$$s.l.\ Ax = b,$$
$$h(x) \leq 0,$$

Regularitate: (x^*, μ^*) minim regulat daca $\mu^* \geq 0, \mu_i^* h_i(x^*) = 0$ si conditia Slater:

$$h_i(x^*) < 0, \qquad i = 1, \cdots, m$$

Concluzia teoremei: **Daca problema are minim regulat, atunci se poate reduce la o problema fara constrangeri (obiectivul este functia Lagrangian)!**

Probleme de programare convexa

$$\min_{x} f(x)$$
$$s.l.\ Ax = b,$$
$$h(x) \leq 0,$$

Regularitate: (x^*, μ^*) minim regulat daca $\mu^* \geq 0, \mu_i^* h_i(x^*) = 0$ si conditia Slater:

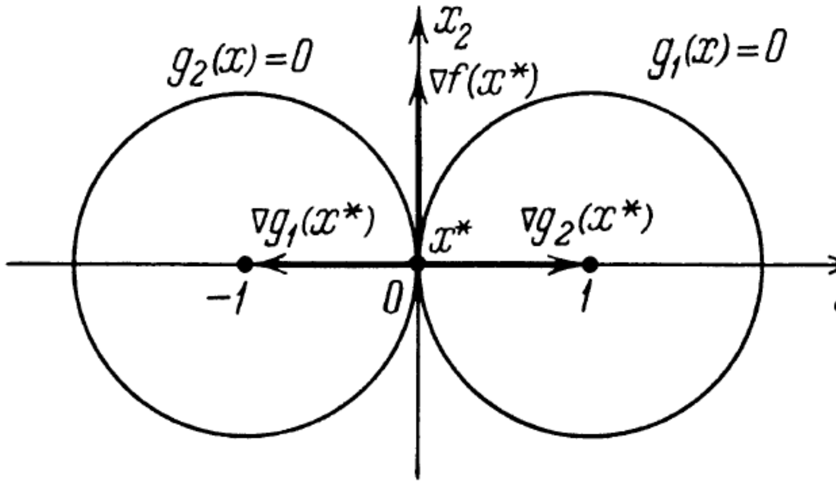
$$h_i(x^*) < 0, \qquad i = 1, \cdots, m$$

Concluzia teoremei: **Daca problema are minim regulat,**

$$\boldsymbol{x^*} \Rightarrow \boldsymbol{\nabla_{\mathbf{x}} \mathbf{L}(\mathbf{x}^*, \lambda^*, \mu^*) = 0}$$

Probleme de programare convexa

$$\min_{x} x_2$$
$$s.l.\ (x_1 - 1)^2 + x_2^2 \leq 1,$$
$$(x_1 + 1)^2 + x_2^2 \leq 1,$$



Conditia Slater nu are loc:

$[\ 0\ 0\ ]^T$ singurul punct fezabil!

Se arata usor ca nu exista $\mu_1^*, \mu_2^* > 0$

Incat $\nabla L(0, \mu^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\mu_1^* \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 2\mu_2^* \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0.$

Probleme de programare convexa

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

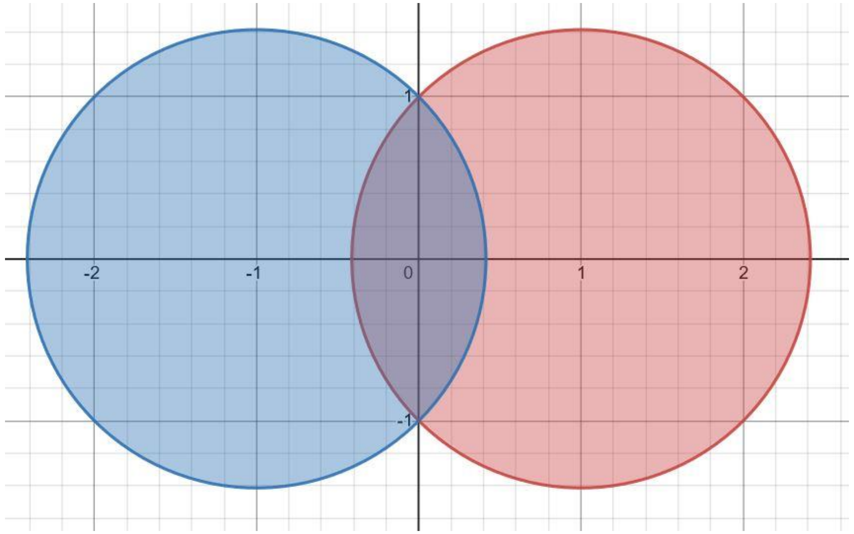
Conditia Slater are loc!

C.N.: $\nabla_x L(x, \mu) = 0$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \mu_1 2 \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + \mu_2 2 \begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mu_1[(x_1 - 1)^2 + x_2^2 - 2] = 0$$

$$\mu_2[(x_1 + 1)^2 + x_2^2 - 2] = 0$$



Probleme de programare convexa

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

Cazul $\boldsymbol{\mu_1 = \mu_2 = 0}$

C.N.: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (Exclus)

Probleme de programare convexa

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

Cazul $\boldsymbol{\mu_1 = 0, \mu_2 > 0}$

C.N.: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\mu_2 \begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = -1\ (violeaza\ prima\ ineg.)$

$$(x_1 - 1)^2 + x_2^2 < 2$$

$$(x_1 + 1)^2 + x_2^2 = 2$$

Probleme de programare convexa

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

Cazul $\boldsymbol{\mu_1 > 0, \mu_2 = 0}$

C.N.: $\begin{bmatrix}0\\1\end{bmatrix} + 2\mu_1 \begin{bmatrix}x_1-1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \Rightarrow x_1 = 1\ (violeaza\ a\ doua\ ineg.)$

$$(x_1-1)^2+x_2^2 = 2$$
$$(x_1+1)^2+x_2^2 < 2$$

Probleme de programare convexa

$$\min_{x} x_2$$
$$s.l.\ (x_1 - 1)^2 + x_2^2 \leq 2,$$
$$(x_1 + 1)^2 + x_2^2 \leq 2$$

Cazul $\boldsymbol{\mu_1 > 0, \mu_2 > 0}$

C.N.: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\mu_1 \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + 2\mu_2 \begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = \frac{\mu_2 - \mu_1}{\mu_1 + \mu_2}, x_2 = -\frac{1}{\mu_1 + \mu_2}$

$$(x_1 - 1)^2 + x_2^2 = 2$$
$$(x_1 + 1)^2 + x_2^2 = 2$$

Probleme de programare convexa

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

Cazul $\boldsymbol{\mu_1 > 0, \mu_2 > 0}$

C.N.: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\mu_1 \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + 2\mu_2 \begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = \frac{\mu_2 - \mu_1}{\mu_1 + \mu_2}, x_2 = -\frac{1}{\mu_1 + \mu_2}$

$$(x_1 - 1)^2 + x_2^2 = 2 \quad \Rightarrow \mu_1 = \mu_2 = \frac{1}{2}$$
$$(x_1 + 1)^2 + x_2^2 = 2$$

Probleme de programare convexa

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

Cazul $\boldsymbol{\mu_1 > 0, \mu_2 > 0}$

C.N.: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2\mu_1 \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + 2\mu_2 \begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = \frac{\mu_2 - \mu_1}{\mu_1 + \mu_2}, x_2 = -\frac{1}{\mu_1 + \mu_2}$

$$(x_1 - 1)^2 + x_2^2 = 2 \quad \Rightarrow \mu_1 = \mu_2 = \frac{1}{2} \Rightarrow x_1 = 0, x_2 = -1$$
$$(x_1 + 1)^2 + x_2^2 = 2$$

Dualitate

Functia Lagrangian: $L: R^n \times R^m \times R_+^p$

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x)$$

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

Problema duala: $$\max_{\mu \geq 0, \lambda} \phi(\lambda, \mu)$$

- Solutiile problemei duale: μ^*, λ^* reprezinta multiplicatorii Lagrange!
- Soluta problemei primale: x^* se numeste solutie primala!

Dualitate

Teorema dualitate. Pentru oricare $\boldsymbol{x}, \boldsymbol{\mu}$ fezabili avem:

$$f(x) \geq \phi(\lambda, \mu).$$

Daca x^* minim regulat si (λ^*, μ^*) multiplicatori Lagrange, atunci

$$f^* = \phi^*$$

- **Sub regularitate, $\boldsymbol{f}$ si $\boldsymbol{\phi}$ au minime, iar valorile optime coincid.**
- **Mai mult, rezolvarea problemei duale poate mai simpla decat primala!**

Dualitate-exemple

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

Dualitate-exemple

$$\min_x c^T x$$

$$s.l. \left|\left|x\right|\right|^2 \leq 1$$

Functia Lagrangian: $L(x, \mu) = c^T x + \mu \left(\left|\left|x\right|\right|^2 - 1\right)$

Functia duala: $\phi(\mu) = \min_x L(x, \mu)$

Dualitate-exemple

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

Functia Lagrangian: $L(x, \mu) = c^T x + \mu \left(\left||x|\right|^2 - 1 \right)$

Functia duala: $\phi(\mu) = \min_x L(x, \mu)$

$$x(\mu) = arg\min_x c^T x + \mu \left(\left||x|\right|^2 - 1 \right) \Rightarrow c + 2\mu x(\mu) = 0$$

Dualitate-exemple

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

Functia Lagrangian: $L(x,\mu) = c^T x + \mu\left(\left||x|\right|^2 - 1\right)$

Functia duala: $\phi(\mu) = \min_x L(x,\mu)$

$$x(\mu) = arg\min_x c^T x + \mu\left(\left||x|\right|^2 - 1\right) \Rightarrow c + 2\mu x(\mu) = 0$$

Deci: $x(\mu) = -\frac{1}{2\mu} c \Rightarrow \phi(\mu) = L(x(\mu),\mu) = c^T x(\mu) + \mu\left(\left||x(\mu)|\right|^2 - 1\right)$

Dualitate-exemple

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

Functia Lagrangian: $L(x, \mu) = c^T x + \mu \left(\left||x|\right|^2 - 1 \right)$

Functia duala: $\phi(\mu) = \min_x L(x, \mu)$

$$x(\mu) = arg\min_x c^T x + \mu \left(\left||x|\right|^2 - 1 \right) \Rightarrow c + 2\mu x(\mu) = 0$$

Deci: $x(\mu) = -\frac{1}{2\mu} c \Rightarrow \phi(\mu) = L(x(\mu), \mu) = c^T x(\mu) + \mu \left(\left||x(\mu)|\right|^2 - 1 \right)$

$$\phi(\mu) = -\frac{1}{4\mu} \left||c|\right|^2 - \mu$$

Dualitate - exemple

$$\min_{x} c^T x$$
$$s.l. \left|\left|x\right|\right|^2 \leq 1$$

Problema duala:

$$\max_{\mu \geq 0} -\frac{1}{4\mu}\left|\left|c\right|\right|^2 - \mu$$

Dualitate-exemple

$$\min_x \frac{1}{2} x^T Q x + q^T x \qquad (Q \succ 0)$$
$$s.l. \left||x|\right|^2 \leq 1$$