TEHNICI DE OPTIMIZARE
Curs 8-9

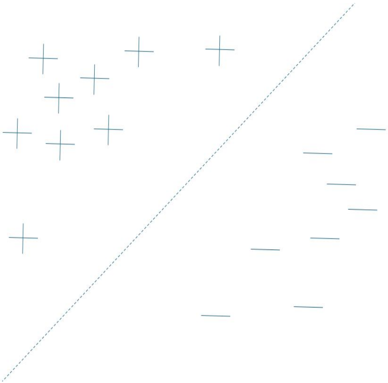
Andrei Pătraşcu

Departament Informatică
Universitatea din Bucureşti

-

Fie seturile de "obiecte": $X^+ = \{x_1^+, x_2^+, \cdots, x_m^+\}, X^- = \{x_1^-, x_2^-, \cdots, x_m^-\}$. Determinaţi hiperplanul $H(w, 0)$ care separă (distinge) cele două seturi.

$$w^T x_i^+ > 0 \quad i = 1, \cdots, m$$
$$w^T x_i^- < 0 \quad i = 1, \cdots, m.$$



Fie seturile de "obiecte": $X^+ = \{x_1^+, x_2^+, \cdots, x_m^+\}, X^- = \{x_1^-, x_2^-, \cdots, x_m^-\}$.
Determinaţi hiperplanul $H(w, 0)$ care separă (distinge) cele două seturi.

Echivalent, căutăm w astfel:

$$y_i(w^T x_i) > 0 \quad \forall x_i \in X^+ \cup X^-.$$

unde $y_i = \pm 1$ dacă $x_i \in X^+$ sau X^-.
Printr-o simplă schimbare de variabilă, problema se reduce la:

$$-y_i(\hat{w}^T x_i) \leq -1 \quad \forall x_i.$$

- Problema de clasificare liniară
- **Constrângeri de inegalitate. Condiţii de optimalitate (convexitate)**
- Algoritmi pentru (POCi) convexe

Programare Neliniară: probleme de minimizare supuse la constrângeri de inegalitate

$$(POCi:) \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s.l.} \;\; h(x) \leq 0, g(x) = 0.$$

- Mulţimea fezabilă este definită $Q = \{x \in \mathbb{R}^n : \; h_i(x) \leq 0, g_j(x) = 0, i = 1, \cdots, p; j = 1, \cdots, m\}$.
- Optim-ul global: $f(x^*) \leq f(x), \; \forall x \in Q$.
- POCi convexă dacă f, h_j convexe pentru $j = 1, \cdots, p$ + g_i liniare pentru $i = 1, \cdots, m$.

$$(POCi\ \ convexa:)\quad \min_{x\in\mathbb{R}^n}\ f(x)$$
$$\text{s.l.}\ \ h(x)\leq 0, Ax=b.$$

unde f, h_i funcţii convexe, $h:\mathbb{R}^n\rightarrow\mathbb{R}^p, A\in\mathbb{R}^{m\times n}, b\in\mathbb{R}^m$.

$$(POCi\ \ convexa:) \quad \min_{x\in\mathbb{R}^n} \ f(x)$$
$$\text{s.l.}\ \ h(x) \leq 0, Ax = b.$$

unde f, h_i funcţii convexe, $h : \mathbb{R}^n \to \mathbb{R}^p, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$.

$$(SVM:) \quad \min_{w,b,\xi} \ \frac{1}{2}\|w\|_2^2 + \rho\sum_i \xi_i$$
$$\text{s.l.}\ \ y_i(w^T x_i - b) \geq 1 - \xi_i, \xi \geq 0.$$

$$(POCi\ \ convexa:)\quad \min_{x\in\mathbb{R}^n}\ f(x)$$
$$\text{s.l. } h(x)\leq 0, Ax=b.$$

unde f, h_i funcţii convexe, $h:\mathbb{R}^n\to\mathbb{R}^p, A\in\mathbb{R}^{m\times n}, b\in\mathbb{R}^m$.

$$(Problema\ \ Google:)\quad \min_{x}\ \frac{1}{2}\|Ex-x\|_2^2$$
$$\text{s.l. } \sum_i x_i = 1, x\geq 0.$$

$$(POCi:) \quad \min_{x\in\mathbb{R}^n} \; f(x)$$
$$\text{s.l.} \;\; h(x) \leq 0, Ax = b,$$

unde f, h_i funcţii convexe, $h : \mathbb{R}^n \to \mathbb{R}^p, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$.

$$(POCi:) \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s.l.} \;\; h(x) \leq 0, Ax = b,$$

unde f, h_i funcţii convexe, $h : \mathbb{R}^n \to \mathbb{R}^p, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

- Funcţia Lagrange $\mathcal{L} : \mathrm{dom}(f) \times \mathbb{R}^m \times \mathbb{R}^p_+$:

$$\mathcal{L}(x, \mu) = f(x) + \mu^T(Ax - b) + \lambda^T h(x)$$

- μ, λ multiplicatori Lagrange (variabile duale)

$$Q = \{x : Ax = b, h_i(x) \leq 0 \quad \forall i = 1, \cdots, p\}$$

$$\min_x f(x) + \iota_Q(x) = \begin{cases} f(x) & x \in Q \\ \infty & x \notin Q \end{cases}$$

Reformulare funcţie indicator (multiplicatori Lagrange):

$$\iota_Q(x) = \max_{\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^p_+} \mu^T(Ax - b) + \lambda^T(h(x)) \left(= \sum_i \mu_i(A_i x - b_i) + \lambda_i h_i(x) \right)$$
$$= \begin{cases} 0 & x \in Q \\ \infty & x \notin Q. \end{cases}$$

$$\min_x \quad f(x) + \iota_Q(x) = \begin{cases} f(x) & x \in Q \\ \infty & x \notin Q \end{cases}$$

Echivalent (multiplicatori Lagrange):

$$\min_x f(x) + \iota_Q(x) = \min_{x \in \mathbb{R}^n} \underbrace{\max_{\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^p_+} \quad f(x) + \; \mu^T (Ax - b) + \lambda^T (h(x))}$$

$$= \begin{cases} f(x) & x \in Q \\ \infty & x \notin Q \end{cases}$$

$$\min_x f(x) + \iota_Q(x) = \begin{cases} f(x) & x \in Q \\ \infty & x \notin Q \end{cases}$$

Echivalent (multiplicatori Lagrange):

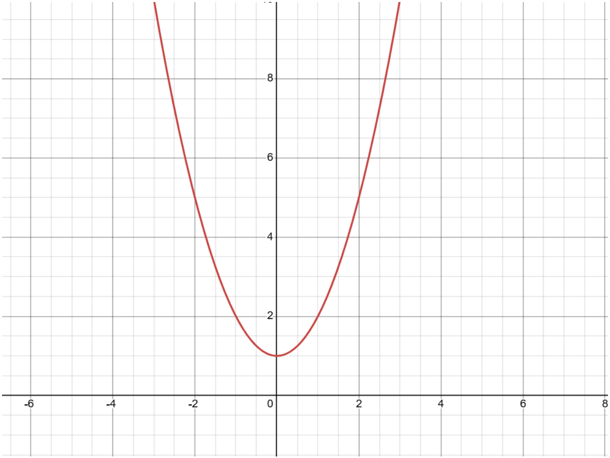
$$\min_x f(x) + \iota_Q(x) = \min_{x \in \mathbb{R}^n} \max_{\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^p_+} f(x) + \mu^T (Ax - b) + \lambda^T (h(x))$$

[Pentru $\mu = \mu^*, \lambda = \lambda^*$ mărginiţi] $= \min_{x \in \mathbb{R}^n} f(x) + (\mu^*)^T (Ax - b) + (\lambda^*)^T h(x).$

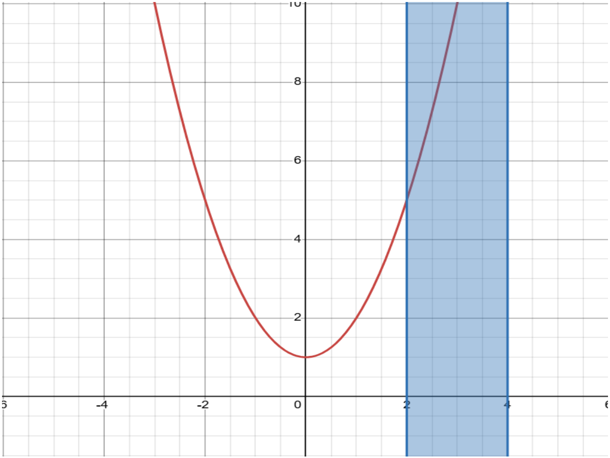
- Problema de clasificare liniară
- **Constrângeri de inegalitate. Condiţii de optimalitate (convexitate)** ‘
- Algoritmi pentru (POCi) convexe

$$\begin{aligned} \min_{x \in \mathbb{R}} \ & x^2 + 1 \\ \text{s.l.} \ & (x-2)(x-4) \leq 0. \end{aligned}$$

$$\min_{x \in \mathbb{R}} \; x^2 + 1$$
$$\text{s.l.} \; (x-2)(x-4) \leq 0.$$

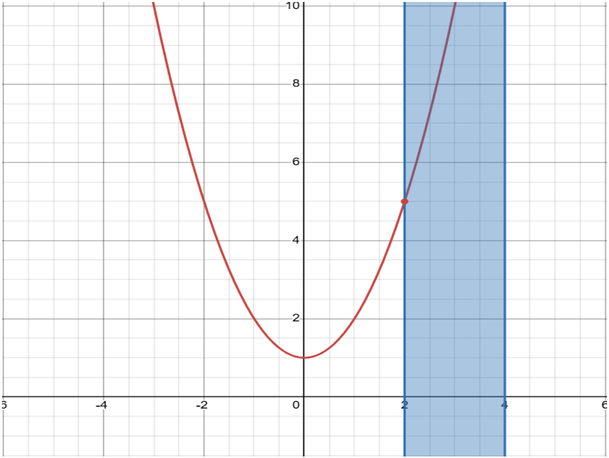


$$\min_{x \in \mathbb{R}} \; x^2 + 1$$
$$\text{s.l.} \; (x-2)(x-4) \le 0.$$



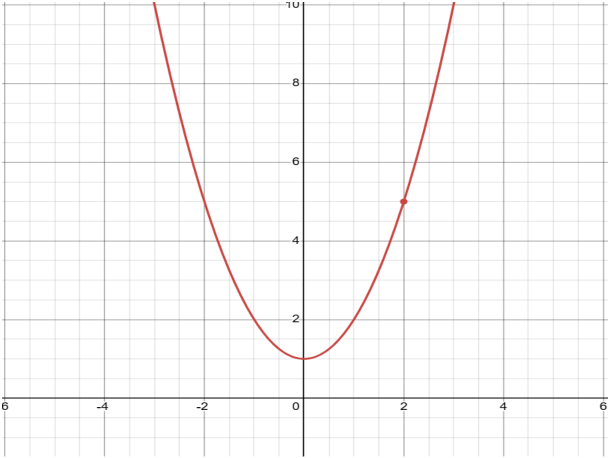
$$\min_{x \in \mathbb{R}} \; x^2 + 1 \;\; \text{s.l.} \;\; (x-2)(x-4) \leq 0.$$

$$x^* = 2.$$



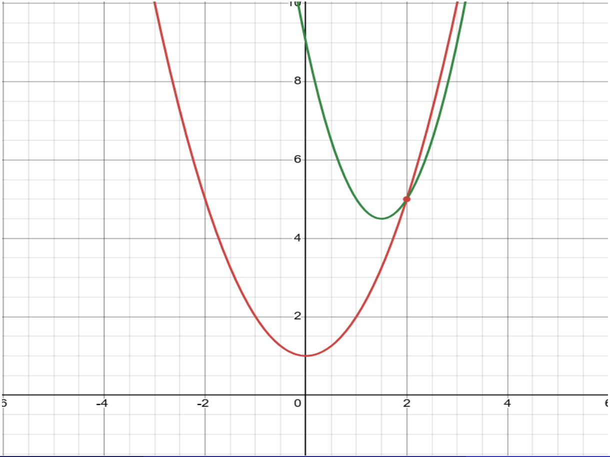
$$\min_{x \in \mathbb{R}} \; x^2 + 1 \;\; \text{s.l.} \;\; (x-2)(x-4) \leq 0$$

$$x^* = 2.$$



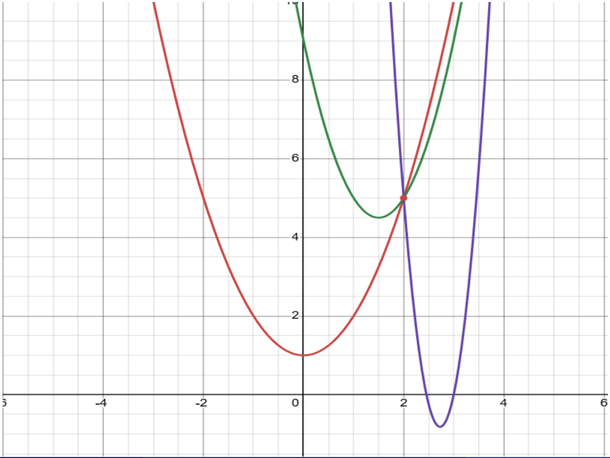
$$\min_{x\in\mathbb{R}} \; x^2+1 \;\text{ s.l. }\; (x-2)(x-4)\le 0.$$

(verde) $\mathcal{L}(x,1) = x^2+1+(x-2)(x-4)$.



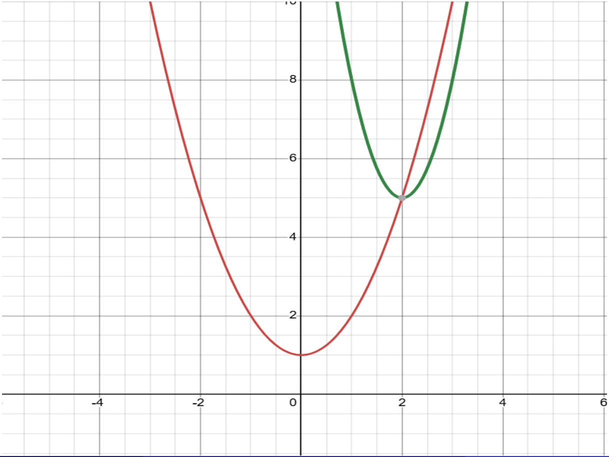
$$\min_{x \in \mathbb{R}} \; x^2 + 1 \;\; \text{s.l.} \;\; (x-2)(x-4) \leq 0.$$

(albastru) $\mathcal{L}(x, 10) = x^2 + 1 + 10(x-2)(x-4)$.



$$\min_{x \in \mathbb{R}} \; x^2 + 1 \;\; \text{s.l.} \;\; (x-2)(x-4) \leq 0.$$

(verde) $\mathcal{L}(x, \lambda^* = 2) = x^2 + 1 + 2(x-2)(x-4)$. verificaţi!



$$(POCi:) \quad \min_{x\in\mathbb{R}^n} \; f(x)$$
$$\text{s.l.} \quad Cx \le d, Ax = b,$$

unde $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$.

Teoremă (Kuhn-Tucker)

Fie f, h_i funcţii convexe. Punctul fezabil x^* este optim dacă şi numai dacă:

$$\nabla_x \mathcal{L}(x^*, \mu^*, \lambda^*) := \nabla f(x^*) + A^T\mu^* + C^T\lambda^* = 0 \; [opt.]$$
$$Cx^* \le d, Ax^* = b \quad [fezabilitate]$$
$$\lambda_i^*(C_i x^* - d_i) = 0 \quad \forall i = 1, \cdots, p \quad [complementaritate].$$

$$(POCi:) \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s.l.} \;\; h(x) \leq 0, Ax = b,$$

unde $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Teoremă (Kuhn-Tucker)

*Fie f, h_i funcţii convexe. Sub **condiţia Slater**: există z astfel*

$$h_i(z) < 0 \qquad \forall i = 1, \cdots, p,$$

punctul fezabil x^* este optim dacă şi numai dacă:

$$\nabla_x \mathcal{L}(x^*, \mu^*, \lambda^*) := \nabla f(x^*) + A^T \mu^* + \nabla h(x^*)^T \lambda^* = 0 \; [opt.]$$
$$h(x^*) \leq 0, Ax^* = b \quad [fezabilitate]$$
$$\lambda_i^* h_i(x^*) = 0 \quad \forall i = 1, \cdots, p \quad [complementaritate].$$

$$(POCi:) \quad \min_{x \in \mathbb{R}^n} \; f(x)$$
$$\text{s.l.} \quad h(x) \leq 0, Ax = b,$$

unde $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Teoremă (Kuhn-Tucker)

Fie f, h_i funcţii convexe. Sub ***condiţia Slater****: există z astfel*

$$h_i(z) < 0 \qquad \forall i = 1, \cdots, p,$$

punctul fezabil x^* este optim dacă şi numai dacă:

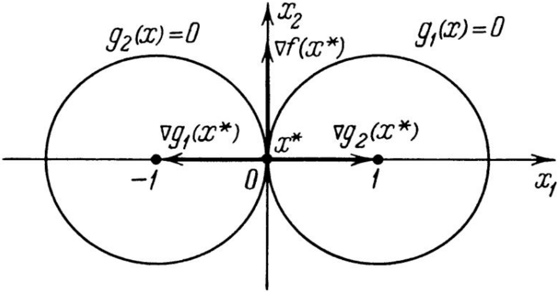
$$\mathcal{L}(x^*, \mu, \lambda) \leq \mathcal{L}(x^*, \mu^*, \lambda^*) \leq \mathcal{L}(x, \mu^*, \lambda^*)$$

Echivalent, teorema Kuhn-Tucker garantează:

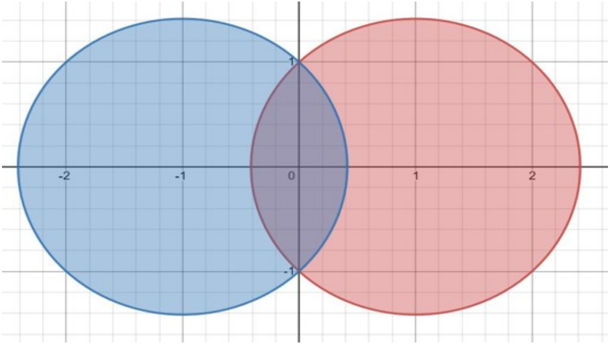
$$\min_{x} \max_{\mu,\lambda\in\mathbb{R}_{+}^{p}} \mathcal{L}(x,\mu,\lambda) = \max_{\mu,\lambda\in\mathbb{R}_{+}^{p}} \min_{x} \mathcal{L}(x,\mu,\lambda)$$

Perechea $(x^{*},[\mu^{*},\lambda^{*}])$ este punct-şa al funcţie Lagrangian.

$$\begin{aligned} \min_{x\in\mathbb{R}^2} \quad & x_2 \\ \text{s.l.} \quad & (x_1-1)^2+x_2^2 \le 1 \\ & (x_1+1)^2+x_2^2 \le 1 \end{aligned}$$



$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & x_2 \\ \text{s.l.} \quad & (x_1 - 1)^2 + x_2^2 \leq 2 \\ & (x_1 + 1)^2 + x_2^2 \leq 2 \end{aligned}$$



- Dualitate. Exemple
- Algoritmi pentru (POCi) convexe

Teorema Kuhn-Tucker garantează:

$$\min_{x} \max_{\mu,\lambda \in \mathbb{R}_+^p} \mathcal{L}(x,\mu,\lambda) = \max_{\mu,\lambda \in \mathbb{R}_+^p} \min_{x} \mathcal{L}(x,\mu,\lambda)$$

Perechea $(x^*, [\mu^*, \lambda^*])$ este punct-şa al funcţiei Lagrangian.

Definiţie

Punctul $(x^*, [\mu^*, \lambda^*])$ este minim regulat (sau problema este regulată) dacă condiţia Slater are loc şi $\lambda_i^* h_i(x^*) = 0$.

Funcţia Lagrange $\mathcal{L} : \mathrm{dom}(f) \times \mathbb{R}^m \times \mathbb{R}^p_+$:

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \mu^T (Ax - b) + \lambda^T h(x).$$

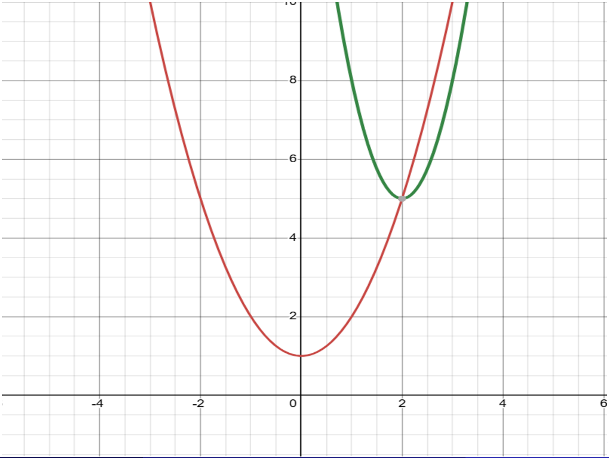
Funcţia duală $\phi : \mathbb{R}^m \times \mathbb{R}^p_+$:

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

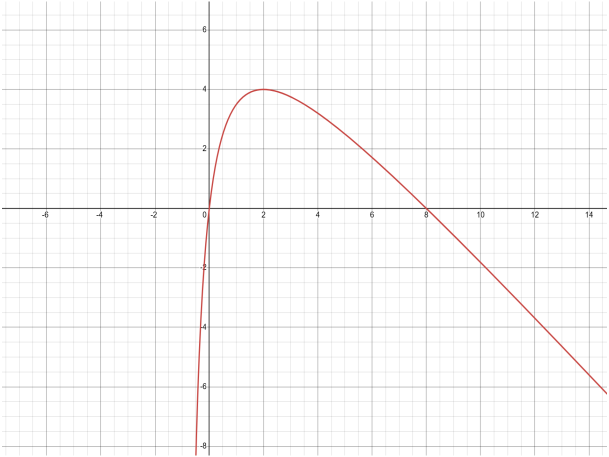
Observaţie: $\mathcal{L}(\cdot, \mu, \lambda)$ este funcţie convexă!

$$\min_{x\in\mathbb{R}} \; x^2+1 \;\text{ s.l. }\; (x-2)(x-4)\le 0.$$

$$\mathcal{L}(x,\lambda^*=2)=x^2+1+2(x-2)(x-4)$$



$$\phi(\lambda) = -\frac{9\lambda^2}{1+\lambda} + 8\lambda$$



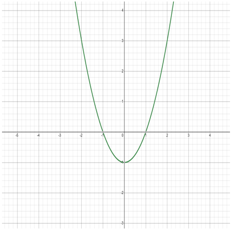
$$\begin{aligned}\phi(\mu,\lambda) &= \min_{x} \ \mathcal{L}(x,\mu,\lambda)\\ &= \min_{x,t,u,v} \ t + u^T\mu + v^T\lambda \ \text{ s.l. } \ f(x) \leq t, \ \ g(x) = u, \ \ h(x) = v\end{aligned}$$

Notăm:

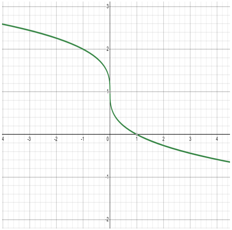
$$\mathcal{F} = \{(f(x), g(x), h(x)) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^p \mid x \in \text{ dom}(f)\}$$

$$\mathcal{F} = \{(f(x), g(x), h(x)) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^p \mid x \in \operatorname{dom}(f)\}$$

- min x $s.l.$ $x^2 \leq 1$



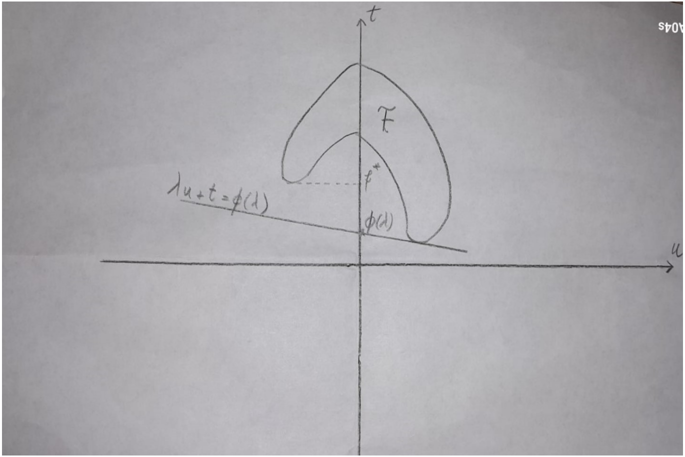
- min x^3 $s.l.$ $x \geq 1$



$$\begin{aligned}\phi(\mu,\lambda) = &= \min_{t,u,v}\; t + u^T\mu + v^T\lambda \;\text{ s.l. }\; u,v,t \in \mathcal{F}\\ &\leq t + u^T\mu + v^T\lambda \;\text{ s.l. }\; u,v,t \in \mathcal{F}.\end{aligned}$$

Hiperplanul $H(-[\mu^T\ \lambda^T\ 1]^T, \phi(\mu,\lambda))$ reprezintă hiperplan de suport al $\mathcal{F}$.

Dualitate

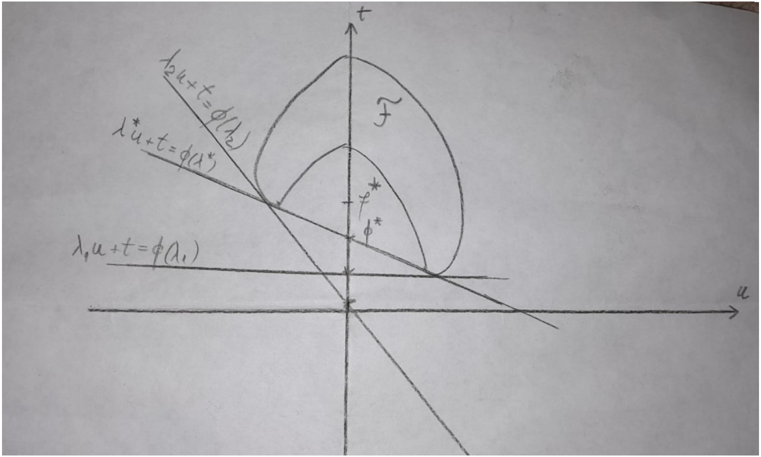


S. Boyd, L. Vandanberghe, Convex Optimization, 2004.

Dualitate slabă:

$$
\begin{aligned}
\phi(\mu,\lambda) &= \min_{t,u,v}\ t + u^T\mu + v^T\lambda \ \text{ s.l. } \ u,v,t \in \mathcal{F} \\
&\leq \min_{t,u,v}\ t + u^T\mu + v^T\lambda \ \text{ s.l. } \ u,v,t \in \mathcal{F}, u = 0, v \leq 0 \\
&\leq \min_{t,u,v}\ t \ \text{ s.l. } \ u,v,t \in \mathcal{F}, u = 0, v \leq 0 \\
&= f^*
\end{aligned}
$$

Dualitate



S. Boyd, L. Vandanberghe, Convex Optimization, 2004.

Definim: $v(u) = \min f(x)$ s.l. $h(x) \leq u$

$$\text{Condiţia Slater}: \quad \exists x : h(x) < 0.$$

Dacă condiţia Slater este satisfăcută, atunci (din Kuhn-Tucker) pentru $\forall \mu, \lambda \geq 0, x \in dom(f)$

$$\mathcal{L}(x, \mu^*, \lambda^*) \geq \phi^* = \mathcal{L}(x^*, \mu^*, \lambda^*) \geq \mathcal{L}(x^*, \mu, \lambda)$$

(*i*) **Dualitate slabă:** $\phi^* \leq f(x^*) = f^*$

(*ii*) Fie $\lambda = 0$, atunci $\phi^* \geq \mathcal{L}(x^*, \mu, 0) = f(x^*) + \mu^T h(x^*) + 0^T(g(x^*)) = f^*$

Dualitate tare: $\phi^* = f^*$.

”Patologii” în lipsa regularităţii:

$$\min_x x_1 \quad \text{s.l.} \quad x_2 \leq 0.$$

$$\min_x \frac{1}{x} \quad \text{s.l.} \quad -x \leq 0.$$

$$\min_x \quad x \quad \text{s.l.} \quad x^2 \leq 0.$$

$$\min_{x} \quad c^T x \;\; \text{s.l.} \;\; \|x\| \leq 1.$$

Când dualitatea este utilă?

- Condiţia de tip Slater este vizibil satisfăcută.
- m, p mici (astfel problema duală are dimensiune mică, i.e $m + p$)

Când dualitatea este utilă?

- Condiţia de tip Slater este vizibil satisfăcută.
- m, p mici (astfel problema duală are dimensiune mică, i.e $m + p$)
- constrângeri primale complicate (constrângeri duale $\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^p_+$)

Exemple

$$\min_x \quad \|x - y\|_2^2 \ \text{ s.l. } \ Ax \leq b.$$

$$\begin{aligned} \mathcal{L}(x, \lambda) &= \ \|x - y\|_2^2 + \lambda^T (Ax - b) \\ \phi(\lambda) &= \min_x \mathcal{L}(x, \lambda) \end{aligned}$$

Problema duală: $\displaystyle \max_{\lambda \geq 0} \ -\frac{1}{4}\|A^T \lambda\|_2^2 + \lambda^T (Ay - b)$

LP dimensiune n:

$$\min_x \; c^T x \;\text{ s.l. }\; Ax \leq b.$$

Problema este convexă, m constrângeri liniare $\Rightarrow$ dualitate tare

$$\begin{aligned}
\mathcal{L}(x,\lambda) &= \; c^T x + \lambda^T(Ax - b) = (c + A^T\lambda)^T x - \lambda^T b \\
\phi(\lambda) &= \min_x \; (c + A^T\lambda)^T x - \lambda^T b \\
&= \begin{cases} -\lambda^T b & \text{dacă } \; c + A^T\lambda = 0, \lambda \geq 0 \\ -\infty & \text{altfel.} \end{cases}
\end{aligned}$$

Problema duală: LP dimensiune m

$$\max_{\lambda \geq 0} \; -\lambda^T b \;\text{ s.l. }\; c + A^T\lambda = 0.$$

QP dimensiune n:

$$\min_x \quad \frac{1}{2}x^T Hx + c^T x \ \text{ s.l. } \ Ax \leq b.$$

Problema este convexă $H \succ 0$, m constrângeri liniare $\Rightarrow$ dualitate tare

$$\mathcal{L}(x, \lambda) = \frac{1}{2}x^T Hx + c^T x + \lambda^T (Ax - b)$$

$$\phi(\lambda) = \min_x \quad \frac{1}{2}x^T Hx + (c + A^T \lambda)^T x - \lambda^T b$$

Kuhn-Tucker:

$$
\begin{aligned}
&Hx + c + A^T\lambda = 0\\
&\lambda_i(A_i x - b_i) = 0 \quad \forall i = 1, \cdots, m\\
&Ax \leq b, \lambda \geq 0.
\end{aligned}
$$

Forma soluţiei primale:

$$
\begin{aligned}
x(\lambda) &= H^{-1}(c + A^T\lambda)\\
\phi(\lambda) &= -\frac{1}{2}(c + A^T\lambda)^T H^{-1}(c + A^T\lambda) - \lambda^T b.
\end{aligned}
$$

Problema duală: QP dimensiune m

$$
\max_{\lambda \geq 0} \; -\frac{1}{2}(c + A^T\lambda)^T H^{-1}(c + A^T\lambda) - \lambda^T b
$$

- Dualitate. Exemple
- **Algoritmi pentru (POCi) convexe**

$$\min_x \; f(x) \; \text{ s.l. } \; g(x) \leq 0.$$

Metoda Gradientului Dual:

$$x^{k+1} = \arg\min_x \mathcal{L}(x, \lambda^k)$$
$$\lambda^{k+1} = \pi_{\geq 0}(\lambda^k + \alpha g(x^{k+1}))$$

$$\min_x \; f(x) \;\; \text{s.l.} \;\; g(x) \leq 0.$$

Problema duală:

$$\max_{\lambda \geq 0} \; \phi(\lambda), \qquad \phi(\lambda) = \min_x \mathcal{L}(x, \lambda)$$

- notăm $x(\lambda) = \arg\min_x \mathcal{L}(x, \lambda)$
- Dacă f este $\sigma-$tare convexă, atunci $\nabla \phi(\lambda) = g(x(\lambda))$ este $\frac{1}{\sigma}-$continuu Lipschitz

$$\min_{x} \; f(x) \;\; \text{s.l.} \;\; g(x) \leq 0.$$

Metoda Gradientului Dual:

$$\lambda^{k+1} = \pi_{\geq 0}(\lambda^k + \alpha \nabla \phi(\lambda^k))$$

- Iteraţia este identică cu MGP pentru problema duală!
- $\lambda^k \rightarrow \lambda^* \Rightarrow x(\lambda^k) \rightarrow x^*$

$$\min_x \; f(x) \;\text{ s.l. }\; g(x) \leq 0.$$

Problema proiecţiei ortogonale:

$$\min_x \; \|x - y\|_2^2 \;\text{ s.l. }\; Ax \leq b.$$

$$\nabla \phi(\lambda) = -\frac{1}{2} AA^T \lambda + (Ay - b)$$

Metoda Gradientului Dual:

$$\lambda^{k+1} = \pi_{\geq 0} \left([I - \alpha \frac{1}{2} AA^T] \lambda^k + \alpha(Ay - b) \right)$$

- Iteraţia este $O(mn)$
- Din Teorema de convergenţă MGP: $\phi^* - \phi(\lambda^k) = O\left(\frac{\|A\|^2 \|\lambda^0 - \lambda^*\|^2}{k}\right)$
- Dacă $AA^T \succeq \sigma_{\min} I_m$ atunci $\phi^* - \phi(\lambda^k) = O\left(\left(1 - \frac{\sigma_{\min}}{\|A\|^2}\right)^k \|\lambda^0 - \lambda^*\|^2\right)$

$$\min_x \; f(x) \;\; \text{s.l.} \;\; g(x) \leq 0.$$

$$\min_x \; f(x) \;\; \text{s.l.} \;\; g(x) = s, \;\; s \leq 0.$$

Funcţia Lagrangian Augmentat $\mathcal{L}_\mu : \mathbb{R}^n \times \mathbb{R}^n_- \times \mathbb{R}^m_+$:

$$\mathcal{L}_0(x, s, \lambda) := f(x) + \lambda^T (g(x) - s)$$
$$\mathcal{L}_\rho(x, s, \lambda) := f(x) + \lambda^T (g(x) - s) + \frac{\rho}{2} \|g(x) - s\|^2$$

$$\min_x \; f(x) \; \text{ s.l. } \; g(x) \leq 0.$$

$$\min_x \; f(x) \; \text{ s.l. } \; g(x) = s, \; s \leq 0.$$

Eliminăm s şi obţinem funcţia duală modificată

$$\begin{aligned}\phi_\rho(\lambda) &:= \min_x \min_{s \leq 0} f(x) + \lambda^T (g(x) - s) + \frac{\rho}{2} \|g(x) - s\|^2 \\ &:= \min_x f(x) + \frac{\rho}{2} \|[g(x) + (1/\rho)\lambda]_+\|^2 - \frac{1}{2\rho} \|\lambda\|^2\end{aligned}$$

$$\min_x \; f(x) \;\; \text{s.l.} \;\; g(x) \leq 0.$$

$$\min_x \; f(x) \;\; \text{s.l.} \;\; g(x) = s, \;\; s \leq 0.$$

Problema duală:

$$\max_{\lambda \geq 0} \; \phi_\rho(\lambda), \qquad \phi_\rho(\lambda) = \min_{x,s} \mathcal{L}_\rho(x, s\lambda)$$

- notăm $x_\rho(\lambda) = \arg\min_{x,s} \mathcal{L}_\rho(x, s, \lambda)$
- $\nabla \phi_\rho(\lambda) = [g(x_\rho(\lambda)) + (1/\rho)\lambda]_+ - (1/\rho)\lambda$ este $\frac{1}{\mu}$−continuu Lipschitz
- MGDm: $\lambda^{k+1} = [\lambda^k + \alpha \nabla \phi_\rho(\lambda^k)]_+$

$$\min_x \ \|x - y\|_2^2 \ \text{ s.l. } \ Ax = b$$

$$\begin{aligned}
\mathcal{L}(x, \mu) &:= f(x) + \mu^T(Ax - b) \\
\mathcal{L}_\rho(x, \mu) &:= f(x) + \mu^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|^2 \\
\phi_\rho(\mu) &:= \min_x f(x) + \mu^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|^2 \\
x_\rho(\mu) &:= \arg\min_x f(x) + \mu^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|^2
\end{aligned}$$

Se arată că:

$$\nabla\phi_\rho(\mu) := Ax_\rho(\mu) - b$$

MGDm: $\mu^{k+1} = \mu^k + \alpha(Ax_\rho(\mu^k) - b)$

$$\min_x \ \|x - y\|_2^2 \ \text{ s.l. } \ Ax = b$$

$$\begin{aligned}
\mathcal{L}(x,\mu) &:= f(x) + \mu^T(Ax - b) \\
\mathcal{L}_\rho(x,\mu) &:= f(x) + \mu^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|^2 \\
\phi_\rho(\mu) &:= \min_x f(x) + \mu^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|^2 \\
x_\rho(\mu) &:= \arg\min_x f(x) + \mu^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|^2
\end{aligned}$$

Se arată că:

$$\nabla \phi_\rho(\mu) := Ax_\rho(\mu) - b$$

MGDm: $\mu^{k+1} = \mu^k + \alpha(Ax_\rho(\mu^k) - b)$

- B. Polyak, Introduction to Optimization, Optimization Software Inc., New York, 1987
- Y. Nesterov, Introductory Lectures on Convex Optimization, Kluwer, 2004.
- Boyd, Stephen P., and Lieven Vandenberghe. Convex Optimization. Cambridge university press, 2004.