

Optimisation

1. Optimisation lisse sans contraintes
2. Optimisation lisse avec contraintes
3. Analyse convexe et dualité Lagrangienne
4. Ouverture vers l'optimisation non-lisse
5. Algorithmes d'optimisation

Optimisation sous contraintes

Problème :

Pour $U \subset \mathbf{R}^n$, un ensemble *ouvert*

$$\text{trouver } \min_{x \in \mathbf{R}^d} f(x) \text{ tel que } \begin{cases} e_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \leq 0, & i \in \mathcal{I} \end{cases}$$

x^* est une solution locale du problème ssi $x^* \in \Omega$ et il existe un ensemble ouvert $V \subset U$ contenant x^* tel que pour tout $x \in V \cap \Omega$, $f(x^*) \leq f(x)$

avec

$$\Omega = \{x \in U \mid \text{pour tout } i \in \mathcal{E}, c_i(x) = 0, \text{ pour tout } j \in \mathcal{I}, c_j(x) \leq 0\}$$

l'ensemble admissible pour ce problème

Condition nécessaire d'optimalité: intuition

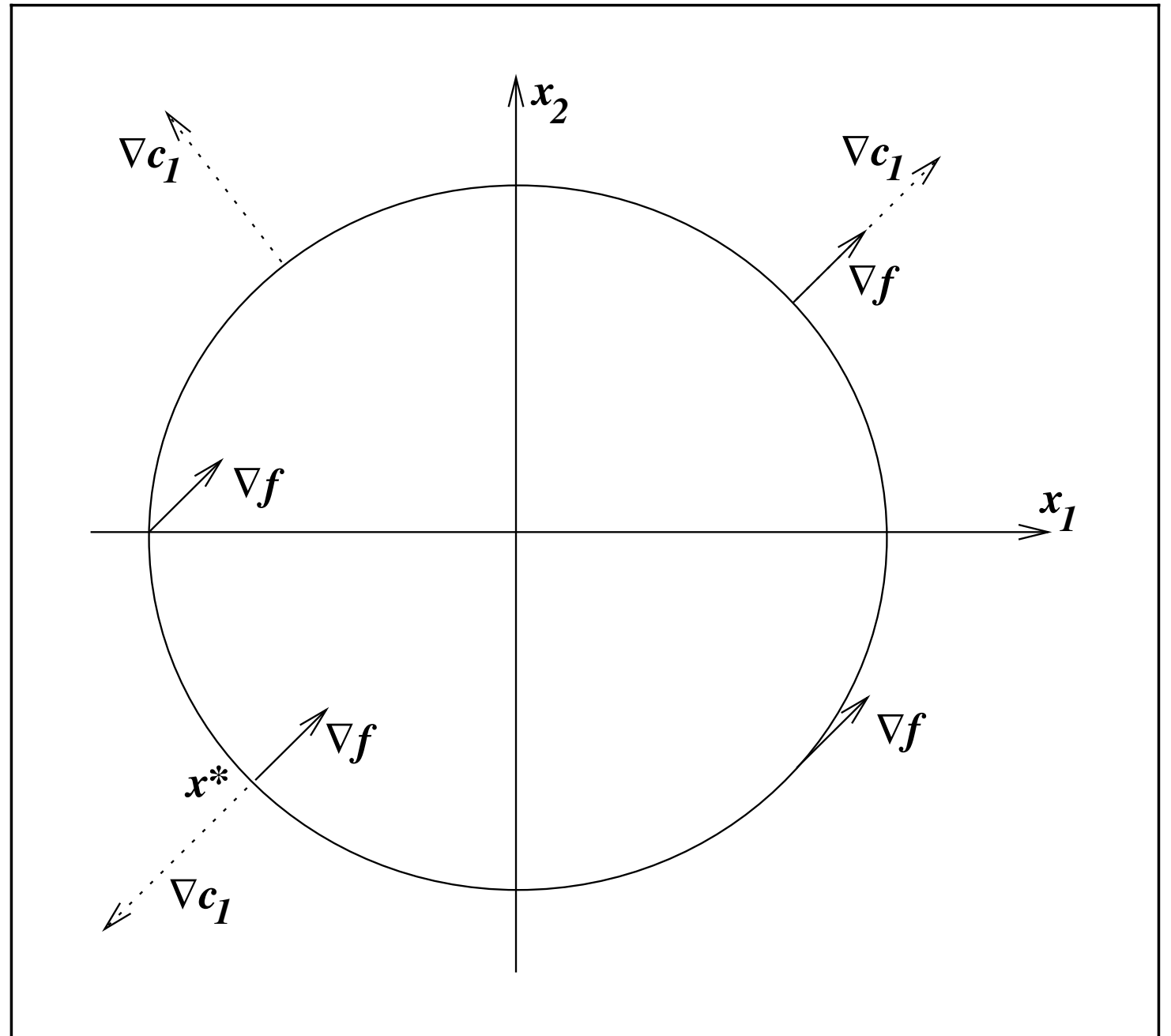
$$f : x_1, x_2 \mapsto x_1 + x_2$$

**Cas d'une seule
contrainte d'égalité**

$$x_1^2 + x_2^2 = 2$$

**Condition nécessaire
d'optimalité**

$$\nabla f(x^*) = \lambda_1^* \nabla c_1(x^*).$$



Condition nécessaire d'optimalité: intuition

$$f : x_1, x_2 \mapsto x_1 + x_2$$

**Cas d'une seule
contrainte d'inégalité**

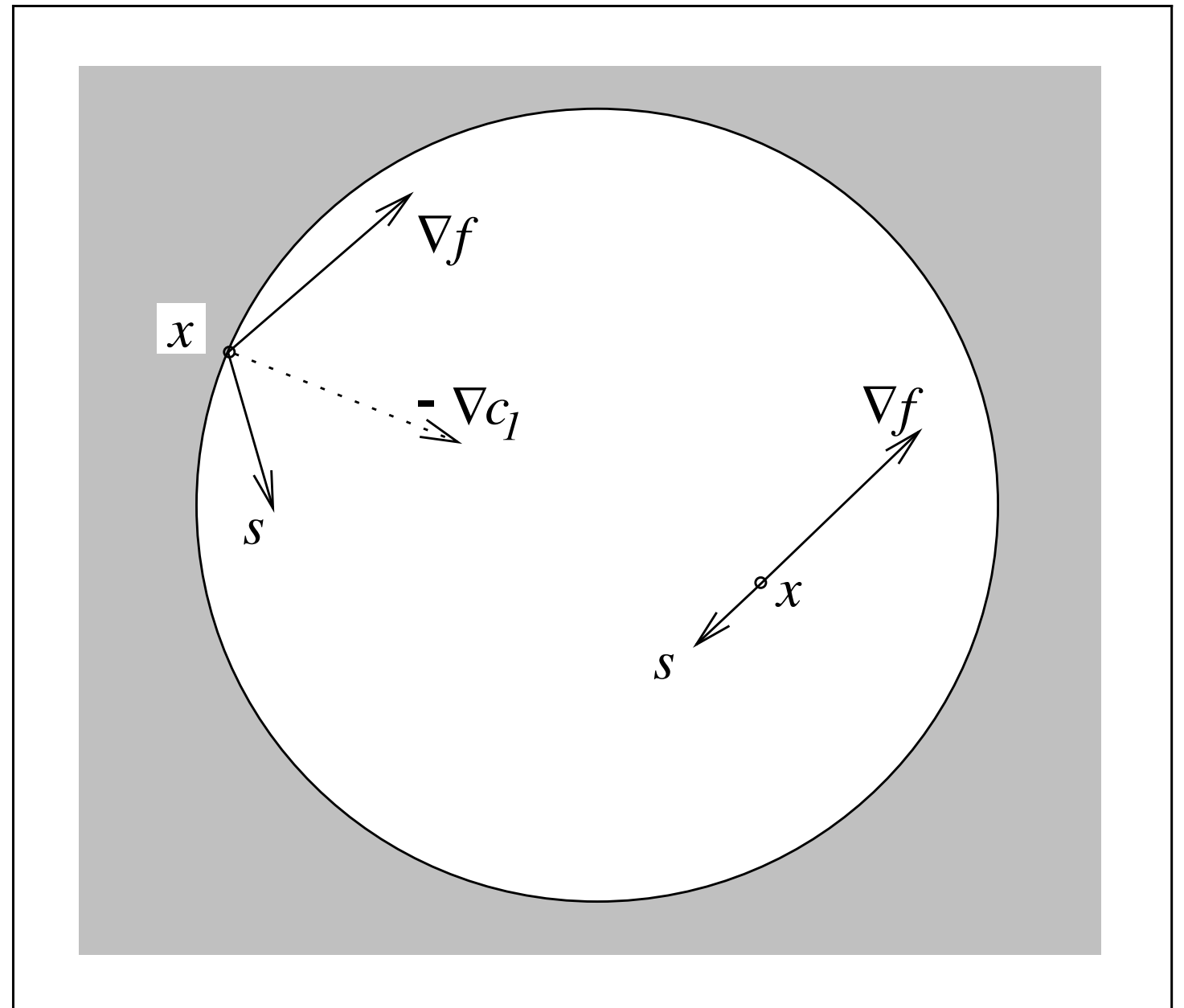
$$x_1^2 + x_2^2 \leq 2$$

**Condition nécessaire
d'optimalité**

$$\nabla f(x^*) = -\lambda_1^* \nabla c_1(x^*)$$

$$\lambda_1^* \geq 0$$

$$\lambda_1^* c_1(x^*) = 0.$$



Cas général : KKT conditions

Soit $U \subset \mathbf{R}^d$, un ensemble *ouvert*

Soit $f, (e_i)_{i=1}^{n_E}$ et $(c_i)_{i=1}^{n_I}$ des fonctions de U vers $\mathbf{R}$ de classe $\mathcal{C}^1$.

on dit que le point $x^* \in U$ vérifie les conditions de Karush-Kuhn-Tucker (KKT) pour le problème :

$$\min_{x \in U} f(x) \text{ tel que } \begin{cases} e_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \leq 0, & i \in \mathcal{I} \end{cases}$$

si et seulement si il existe $\lambda_E^* \in \mathbf{R}^{n_E}$ et $\lambda_I^* \in \mathbf{R}^{n_I}$, tels que :

1. $\nabla_x \mathcal{L}(x^*, \lambda_E^*, \lambda_I^*) = 0$, où $\mathcal{L}(x, \lambda_E, \lambda_I) = f(x) + \sum_{i=1}^{n_E} \lambda_{E,i} e_i(x) + \sum_{i=1}^{n_I} \lambda_{I,i} c_i(x)$
2. $e_i(x^*) = 0$ pour tout i dans $\{1, 2, \dots, n_E\}$
3. $c_i(x^*) \leq 0$ pour tout i dans $\{1, 2, \dots, n_I\}$
4. $\lambda_{I,i} \geq 0$ pour tout i dans $\{1, 2, \dots, n_I\}$
5. $\lambda_{I,i} c_i(x^*) = 0$ pour tout i dans $\{1, 2, \dots, n_I\}$

Condition nécessaire d'optimalité: cas général

Mangasarian-Fromovitz Constraint Qualification (MFCQ)

Les gradients des contraintes d'égalité sont linéairement indépendants au point x^* et il existe un vecteur $d \in \mathbf{R}^n$ tel que $\nabla c_i(x^*)^T d < 0$ pour toutes les contraintes d'inégalité **actives** en x^* et tel que $\nabla e_j(x^*)^T d = 0$ pour toutes les contraintes d'égalité.

Théorème : Si x^* est une solution locale et que la qualification des contraintes de Mangasarian-Fromovitz est respectée en x^* alors les conditions KKT sont vérifiées en x^* .

Autres exemples de qualifications des contraintes

Theorem 1 *Conditions nécessaires de solution locale pour des contraintes linéaires. Si les fonctions e_i et c_j sont affines pour $i = 1 \dots n_E$ et $j = 1 \dots n_I$ alors les conditions KKT sont des conditions nécessaires de solution locale pour le problème considéré.*

Theorem 2 *Conditions nécessaires de solution locale pour des contraintes linéairement indépendantes.*

*Si au point $x \in \mathbf{R}^d$ la matrice formée par les gradients des contraintes d'égalité et des contraintes d'inégalité **actives** est de rang plein (c'est à dire que les n_E vecteurs gradients pour les contraintes d'égalité—qui doivent toujours toutes être actives—et les vecteurs gradients pour les contraintes d'inégalité actives en x —c'est à dire telles que $c_i(x) = 0$ —sont linéairement indépendants), alors les conditions KKT sont des conditions nécessaires de solution locale en x pour le problème considéré.*

Theorem 3 *Conditions nécessaires de solution locale pour un problème convexe.*

Si f est convexe, c_i est convexe pour tout i in $1 \dots n_I$ et e_j est affine pour tout j in $1 \dots n_E$ et si la condition de Slater est vérifiée, c'est à dire s'il existe $x \in \mathbf{R}^d$ tel que $e_j(x) = 0$ pour tout j in $1 \dots n_E$ et $c_i(x) < 0$ pour tout i in $1 \dots n_I$, alors les conditions KKT sont des conditions nécessaires de solution locale pour le problème considéré.

Optimisation

1. Optimisation lisse sans contraintes
2. Optimisation lisse avec contraintes
3. Analyse convexe et dualité Lagrangienne
4. Ouverture vers l'optimisation non-lisse
5. Algorithmes d'optimisation

Optimisation

1. Optimisation lisse sans contraintes
2. Optimisation lisse avec contraintes
3. Analyse convexe et dualité Lagrangienne
4. Ouverture vers l'optimisation non-lisse
5. Algorithmes d'optimisation

Convexité

L'ensemble X est convexe ssi pour tout x, y dans X , le segment $[x, y] := \{tx + (1 - t)y \mid t \in [0, 1]\}$ est inclus dans X

Une fonction $f : X \subset \mathbf{R}^n \rightarrow R$ est convexe ssi X est convexe et pour tout x, y dans X et pour tout $t \in [0, 1]$,

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Si l'inégalité est stricte pour $t \in]0, 1[$, la fonction est dite *strictement* convexe.

Convexité

Reconnaître une fonction convexe

- les fonctions linéaires sont convexes
- les combinaisons linéaires positives de fonctions convexes sont convexes
- le maximum (point par point) d'un ensemble de fonctions convexes est convexe
- ...

Si f est de classe C^2 , f est convexe si et seulement si $\nabla^2 f$ est semi-définie positive sur l'intérieur de X . Si $\nabla^2 f$ est définie positive sur l'intérieur de X , f est strictement convexe. ($X = \text{dom} f$)

Hessienne

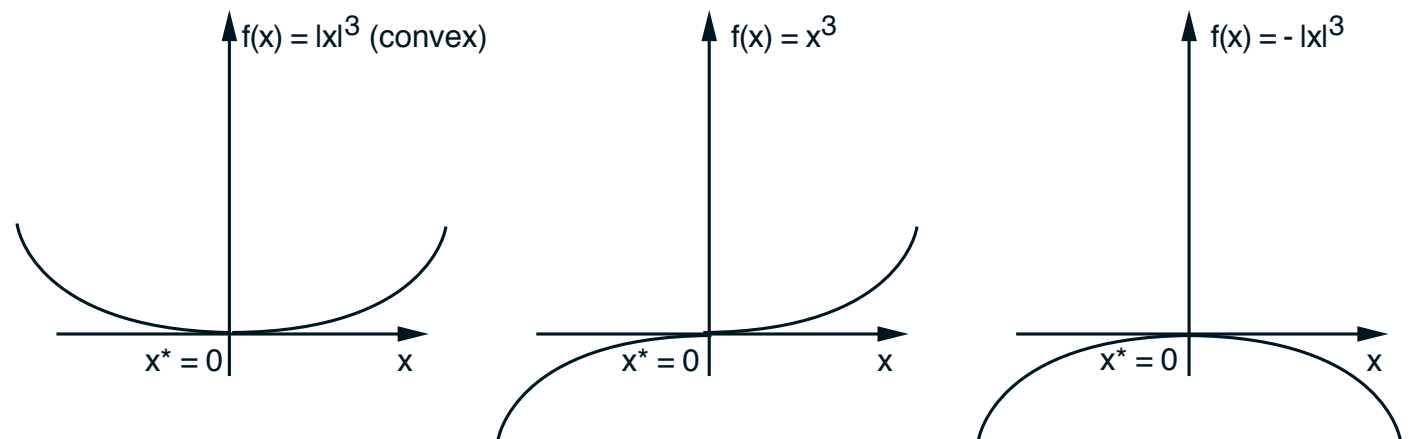
$$f : U \subset \mathbf{R}^n \mapsto \mathbf{R}$$

$$\nabla^2 f = H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Conditions d'optimalité

Conditions nécessaires, premier ordre

Si x^* est un minimum local et f est de classe $\mathcal{C}^1$ sur un ouvert $U \subset \mathbf{R}^n$ contenant x , alors $\nabla f(x^*) = 0$



Conditions nécessaires, second ordre

Si x^* est un minimum local et f est de classe $\mathcal{C}^2$ sur un ouvert $U \subset \mathbf{R}^n$ contenant x , alors $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est semi-définie positive

Conditions suffisantes, second ordre

Si f est de classe $\mathcal{C}^2$ sur un ouvert $U \subset \mathbf{R}^n$ contenant x , si $\nabla f(x^*) = 0$ et si $\nabla^2 f(x^*)$ est définie positive, alors x^* est un minimum local.

Existence et unicité du minimum

Ensemble de sous-niveau α de f : $\{x \in \mathbf{R}^n \text{ tels que } f(x) \leq \alpha\}$

Théorèmes

Si $f : \mathbf{R}^n \mapsto \mathbf{R}$ est convexe et qu'il existe un réel α tel que l'ensemble de sous-niveau α de f est borné et non-vide, alors f admet un minimum global.

Si f est une fonction *strictement* convexe définie sur un ensemble convexe $X \subset \mathbf{R}^n$, alors f admet au plus un minimum global.

Si f est une fonction convexe définie sur un ensemble convexe $X \subset \mathbf{R}^n$, alors tout minimum local de f est aussi un minimum global de f . Si X est ouvert, alors $\nabla f(x^*) = \emptyset$ est équivalent à x^* est un minimum global de f .

Dualité

Problème considéré

$$\min_{x \in X \subset \mathbf{R}^d} f(x) \text{ sous les contraintes } \begin{cases} e_i(x) = 0, & i = 1..r \\ c_i(x) \leq 0, & i = 1..s \end{cases}$$

$$f : X \rightarrow \mathbf{R}$$

$$r \in \mathbf{N}$$

$$s \in \mathbf{N}$$

$$e_i : X \rightarrow \mathbf{R}$$

$$c_i : X \rightarrow \mathbf{R}$$

$$\Omega = \{x \in X \mid e_i(x) = 0, \ c_j(x) \leq 0, \ i = 1..r, \ j = 1..s\}$$

$$f^* = \inf_{x \in \Omega} f(x)$$

$$\text{Hypothèses : } \begin{aligned} \Omega &\neq \emptyset \\ f^* &> -\infty \end{aligned}$$

Dualité

Lagrangien associé au problème

$$L(x, \mu, \lambda) = f(x) + \mu^T e(x) + \lambda^T c(x)$$

avec :

$$x \in X$$

$$c(x) = [c_1(x) \ c_2(x) \ \dots \ c_s(x)]^T$$

$$e(x) = [e_1(x) \ e_2(x) \ \dots \ e_r(x)]^T$$

$$\mu \in \mathbf{R}^r$$

$$\lambda \in \mathbf{R}^s$$

Dualité

Fonction duale $q(\mu, \lambda) = \inf_{x \in X} L(x, \mu, \lambda)$

Domaine $D = \{\mu, \lambda \mid q(\mu, \lambda) > -\infty\}$

Problème dual

$\max_{(\mu, \lambda) \in D} q(\mu, \lambda)$ sous les contraintes $\lambda_i \geq 0, i = 1..s$

Résultats

D est convexe et q est concave sur D

$q^* \leq f^*$ (dualité faible)

avec $q^* = \sup_{\mu \in \mathbf{R}^r, \lambda \in \mathbf{R}^s, \lambda_i \geq 0, i=1..s} q(\mu, \lambda)$