COURS D'OPTIMISATION AVEC CONTRAINTES

Polytch'Clermont – GMM3

JONAS KOKO

ISIMA

Plan du cours

1. Généralités
2. Méthodes de Newton
3. Méthodes de points intérieurs
4. Pénalités et multiplicateurs
5. Programmation quadratique séquentielle

Références Bibliographiques et codes

- D. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley (1984)
- J.F. Bonnans et al., Optimisation Numérique, Springer (1997)
- M. Minoux, Programmation Mathématique, Lavoisier TEC & DOC (2008)
- M. Bierliaire, Introduction à l'Optimisation Différentiable, Presses Poly. Univ. Romandes (2006)

Quelques codes libres

- DONLP2
- IPFilter
- IPOPT

Problème d'optimisation avec contraintes

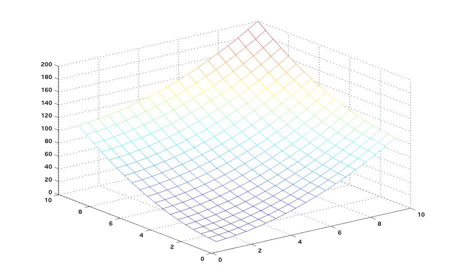
$$\min f(x)$$
$$g_i(x) \leq 0, \quad i = 1, \ldots, p$$
$$h_j(x) = 0, \quad j = 1, \ldots, q$$
$$x \in \mathbb{R}^n.$$

- $f : \mathbb{R}^n \to \mathbb{R}$ fonction coût ou objectif du problème
- $g : \mathbb{R}^n \to \mathbb{R}^p$ contrainte d'inégalité' du problème
- $h : \mathbb{R}^n \to \mathbb{R}^q$ contrainte d'égalité 'du problème

Un exemple

$$\min f(x) = x_1^2 + x_2^2$$
$$x_1 \geq 1,\ \ x_2 \geq 1$$
$$x \in \mathbb{R}^2.$$

$K = \{x \in \mathbb{R}^2 \mid x_1 \geq 1,\ \ x_2 \geq 1\}$
On sait que $\nabla f(0) = 0$ mais $0 \notin K$.
$x^* = (1,1)$ minimise f mas $\nabla f(x^*) \neq 0$



Point et direction admissibles

On pose : $K = \{x \in \mathbb{R}^n \mid g(x) \leq 0 \text{ et } h(x) = 0\}$

Problème avec contraintes

$$(P) \qquad \min_{x \in K} f(x)$$

Définition (Point (ou solution) admissible)

Si g et h son continues, alors K est un sous-ensemble fermé de $\mathbb{R}^n$ appelé ensemble des points (ou solutions) admissibles ou réalisables.

Définition (Direction admissible)

$x \in \mathbb{R}^n$ un point admissible. Une direction d est admissible en x s'il existe $\eta > 0$ tel que $x + td \in K$, pour tout $0 < t \leq \eta$.

Questions d'existence dans $\mathbb{R}^n$

Théorème (Existence de solutions)

Le problème admet une solution dans les deux cas suivants

- f *continue et* K *compact (i.e. fermé et borné)*
- K *fermé et* f *coercive (i.e.* $f(x) \to +\infty$ *quand* $\| x \| \to +\infty$*)*

Théorème (Unicité)

Le problème admet une solution unique si f *est strictement convexe et* K *convexe non vide.*

Définition (Minimum local, global)

- $x^* \in K$ *est minimum local si* $\exists V$ *voisinage de* x^* *tel que* $f(x^*) \leq f(x), \forall x \in V \cap K$
- $x^* \in K$ *est minimum global si* $f(x^*) \leq f(x), \forall x \in K$

Conditions d'optimalité

Définition (Cône)

$C \subset \mathbb{R}^n$ est un cône si $\forall \lambda > 0$, $\lambda x \in C$, $\forall x \in C$.

Définition (Cône tangent)

$a \in K$

$$\begin{aligned} C_a &= \{d \in \mathbb{R}^n \mid \exists \{d_k\} \subset \mathbb{R}^n,\ d_k \to d; \\ &\quad \exists \{t_k\} \subset \mathbb{R}_+,\ t_k \to 0;\ x_k = a + t_k d_k \in K\ \forall k.\} \end{aligned}$$

Théorème (Conditions d'optimalités)

On suppose f de classe $\mathscr{C}^1$ dans un voisinage de $x^* \in K$

1. *Si x^* minimum local pour f alors $\nabla f(x^*)^T d \geq 0$, $\forall d \in C_{x^*}$*
2. *Si K convexe et f convexe, si $\nabla f(x^*)^T d \geq 0$, $\forall d \in C_{x^*}$ alors f admet un minimum global en x^* sur K.*

Cône tangent linéarisé

f et g de classe $\mathscr{C}^1$ dans un voisinage de $a \in K$
Au point a dans la direction d

$$
\begin{array}{rcl}
g_i(a+td) & \approx & g_i(a) + t\nabla g_i(a)^T d \le 0 \ \text{ il faut que } \nabla g_i(a)^T d \le 0 \\
h_j(a+td) & \approx & h_j(a) + t\nabla h_j(a)^T d = 0 \ \text{ il faut que } \nabla h_j(a)^T d = 0.
\end{array}
$$

Définition (Cône tangent linéarisé)

$$
\begin{array}{rcl}
I(a) & = & \{i \in \{1,2,\ldots,p\} \mid g_i(a) = 0\} \ \textit{contraintes saturées (actives) en } a \\
L_a & = & \left\{ d \in \mathbb{R}^n \mid \nabla g_i(a)^T d \le 0,\ \forall i \in I(a);\ \nabla h_j(a)^T d = 0,\ \forall j \right\}
\end{array}
$$

En principe $C_a \neq L_a$.
Conditions pour que $C_a = L_a \equiv$ Conditions de qualification des contraintes.

Qualification des contraintes

on a $C_a = L_a$ dans les cas suivants :

K Polyèdre convexe

Si g et h sont affines alors $C_a = L_a$

Conditions de Slater

- g_i convexes, h_j affines
- $\exists x_0$ tel que $g_i(x_0) < 0$ et $h_j(x_0) = 0$, $\forall i, j$

Conditions de Mangasarian-Fromowitz

- $\nabla h_j(a)$ sont linéairement indépendants
- $\exists \widetilde{d}$ tel que $\nabla g_i(a)\widetilde{d} < 0 \ \forall i \in I(a)$ et $\nabla h_j(a) = 0$, $\forall j$

Contraintes d'égalité

$$\min f(x)$$
$$h_j(x) = 0, \quad j = 1, \ldots, q$$
$$x \in \mathbb{R}^n.$$

Définition (Contraintes régulières)

h_j sont dites régulières en a si les $\nabla h_j(a)$ sont linéairement indépendants

$$\nabla h(x) = \begin{bmatrix} \nabla h_1(x)^T \\ \vdots \\ \nabla h_q(x)^T \end{bmatrix}$$

x régulier :

$$C_x = \{d \in \mathbb{R}^n \mid \nabla h(x) d = 0\} = \left\{d \in \mathbb{R}^n \mid \nabla h_j(x)^T d = 0, \ \forall j\right\}$$

Conditions d'optimalité

Théorème (Conditions (nécessaires) d'optimalité du 1er ordre)

Si x^* est un minimum local et point régulier pour les h_j, alors $\exists \lambda^* \in \mathbb{R}^q$:

$$\nabla f(x^*) + \nabla h(x^*)\lambda^* = \nabla f(x^*) + \sum_{j=1}^{q} \lambda_j^* \nabla h_j(x^*)^T = 0$$

λ_j multiplicateur de Lagrange

On pose : $L = \nabla^2 f(x^*) + \sum_{j=1}^{q} \lambda_j^* \nabla^2 h_j(x^*)$

Théorème (Conditions (nécessaires) d'optimalité du 2ème ordre)

Si x^* est un minimum local et point régulier pour les h_j, alors L est semi-définie positive sur C_{x^*} :

$$d^T L d \geq 0, \quad \forall d \in C_{x^*}$$

Equations d'optimalité

Si x^* est un minimul local :

$$\begin{array}{rcl} \nabla f(x^*) + \nabla h(x^*)\lambda^* & = & 0 \\ h_j(x^*) & = & 0,\ j = 1, \ldots, q \end{array}$$

Système de $n+q$ équations à $n+q$ inconnues

Fonction de Lagrange $\mathscr{L} : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$

$$\mathscr{L}(x, \lambda) = f(x) + h(x)^T \lambda$$

Equations d'optimalité $\Longleftrightarrow \nabla \mathscr{L}(x^*, \lambda^*) = 0$

$\mathscr{L}$ fonction de Lagrange ou Lagrangien

Un exemple

$$\min f(x) = x_1^2 + x_2^2$$
$$x_1 + x_2 = 1$$

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \quad \nabla h(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Le système

$$\begin{aligned} 2x_1 + \lambda &= 0 \\ 2x_2 + \lambda &= 0 \\ x_1 + x_2 &= 1 \end{aligned}$$

La solution $x^* = (1/2, 1/2)$, $\lambda^* = -1$

Contrainte d'inégalité

$$(PI) \qquad \begin{array}{l} \min f(x) \\ g_i(x) \leq 0, \quad i = 1, \dots, p; \;\; x \in \mathbb{R}^n. \end{array}$$

Définition (Contrainte active ou saturée)

$I(x) = \{i \in \{1, 2, \dots, p\} \mid g_i(x) = 0\}$

Proposition

Si x^* solution optimale de (PI), alors x^* est solution optimale de

$$\begin{array}{l} \min f(x) \\ g_i(x) = 0, \quad i \in I(x^*) \end{array}$$

Définition (Cône de directions admissibles)

$$C_{x^*} = \left\{ d \in \mathbb{R}^n \mid \nabla g_i(x^*)^T d \leq 0 \right\}$$

Conditions d'optimalité

Théorème (Karush-Kuhn-Tucker (KKT))

Soit x^* *tel que* $\nabla g_i(x^*)$, $i \in I(x^*)$, *linéairement indépendants*
x^* *minimum local pour* (PI), *alors* $\exists \mu^* \in \mathbb{R}^p$, $\mu_i^* \geq 0$ *:*

$$\begin{aligned} \nabla f(x^*) + \sum_{i \in I(x^*)} \mu_i^* \nabla g_i(x^*) &= 0 \\ \mu_i^* g_i(x^*) &= 0 \end{aligned}$$

μ_i^* *multiplicateur de Kuhn-Tucker*

Équation de complémentarité

$$\mu_i^* g_i(x^*) = 0 \Longleftrightarrow g_i(x^*) \leq 0 \text{ ou } \mu_i^* \geq 0.$$

Fonction de Lagrange (Lagrangien)

Fonction de Lagrange $\mathscr{L} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$

$$\mathscr{L}(x, \mu) = f(x) + \mu^T g(x)$$

Gradient du Lagrangien

$$\nabla \mathscr{L}(x, \lambda) = \begin{pmatrix} \nabla f(x) + \sum \mu_i \nabla g_i(x) \\ g_i(x) \end{pmatrix}$$

$\mu_i = 0$ si $g_i(x^*) < 0 \implies$ contraintes non actives ne comptent pas

Dérivées secondes :

$$L = \nabla^2 f(x^*) + \sum_{i \in I(x^*)} \mu_i^* \nabla g_i(x^*)$$

Théorème (Condition d'optimalité du 2nd ordre)

x^* minimum local pour (PI), alors $d^T L d \geq 0$, $\forall d \in C_{x^*}$.

Problème primal - problème dual

Contraintes :
$X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \; i = 1, \ldots, p \; h_j(x) = 0, j = 1 \ldots, q\}$

Problème primal

$$(P) \qquad m = \inf_{x \in X} f(x)$$

Fonction de Lagrange (Lagrangien) :
$\mathscr{L} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R} \quad \mathscr{L}(x, \mu, \lambda) = f(x) + \mu^T g(x) + \lambda^T h(x)$

Problème dual

$$(P') \qquad \alpha = \sup_{\mu \geq 0, \lambda} \mathscr{L}(x, \mu, \lambda)$$

Point-selle

Définition (Point-selle)

(x^*, μ^*, λ^*) *est un point selle si*

1. x^* *solution optimale de* (P)
2. (μ^*, λ^*) *solution optimale de* (P')
3. $m = \alpha = L(x^*, \mu^*, \lambda^*)$

Si (x^*, μ^*, λ^*) point-selle :

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

ou

$$\sup_{\mu \geq 0, \lambda} \inf_{x \in X} \mathscr{L}(x, \mu, \lambda) = \inf_{x \in X} \sup_{\mu \geq 0, \lambda} \mathscr{L}(x, \mu, \lambda)$$

En général

$$\sup_{\mu \geq 0, \lambda} \inf_{x \in X} \mathscr{L}(x, \mu, \lambda) \leq \inf_{x \in X} \sup_{\mu \geq 0, \lambda} \mathscr{L}(x, \mu, \lambda)$$

Existence d'un point-selle

Théorème (Point-selle)

- *f convexe et $\mathscr{C}^1$*
- *g_i convexes et $\mathscr{C}^1$, h_j affines*
- *Conditions de qualification de Slater satisfaites*

Si (P) admet une solution optimale alors un point selle de $\mathscr{L}$ existe.

Méthodes de résolution

- Méthodes basées sur x : Méthodes primales
- Méthodes basées sur (μ, λ) : Méthodes duales
- Méthodes basées sur x et (μ, λ) : Méthodes primales-duales