Mathématiques Appliquées

Révisions d'Optimisation Convexe

1 Rappel sur les fonctions convexes

On considère ici $\mathcal{X}$ un espace de Hilbert et $f : \mathcal{X} \to \mathbb{R}$. On appelle *épigraphe* de f, noté $\mathbb{U}(f)$ le sous ensemble :

$$\mathbb{U}(f) \overset{\text{def}}{=} \{(x, \xi) \,|\, x \in \mathcal{X},\ \xi \in \mathbb{R} : \xi \geq f(x)\}$$

Définition 1 (Convexité)**.** *Une fonction $f : \mathcal{X} \to \mathbb{R}$ est convexe si, de façon équivalente :*

- *L'épigraphe $\mathbb{U}(f)$ de f est un ensemble convexe*
- *f vérifie, $\forall \alpha \in [0,1]$, $\forall x, y \in \mathcal{X}$:*

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) \tag{1}$$

Définition 2 (Stricte et forte convexité)**.** • *On dit qu'une fonction f est strictement convexe si elle vérifie strictement l'inégalité 1.*
• *On dit qu'une fonction est fortement convexe si $\exists b > 0$, tel que $\forall x, y \in \mathcal{X}, \forall \alpha \in [0,1]$:*

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) - b\alpha(1-\alpha)\frac{||x-y||_{\mathcal{X}}^2}{2} \tag{2}$$

2 Différentiabilité au sens de Gateau

Dans un espace de Hilbert, il existe diverses notions de différentiabilité. Nous nous intéresserons ici à la différentiabilité au sens de Gateau.

Définition 3 (Dérivé directionnelle)**.** *On appelle* dérivée directionnelle *de $f : \mathcal{X} \to \mathbb{R}$ au point $x \in \mathcal{X}$ dans la direction $d \in \mathcal{X}$, notée $Df(x; d)$, la quantité (lorsqu'elle existe) :*

$$Df(x; d) = \lim_{\varepsilon \to 0_+} \frac{f(x + \varepsilon d) - f(x)}{\varepsilon}$$

Définition 4 (Dérivée de Gateau)**.** *Si f admet en x des dérivées directionnelles pour toutes les directions d et si $Df(x; d)$ est linéaire continue en d, alors f est dite différentiable au sens de Gateau au point x.*

Ainsi, si f est différentiable au sens de Gateau au point x, on aura que $\forall y \in \mathcal{X}$:

$$f(x + \varepsilon y) = f(x) + \varepsilon(y \cdot D(x)) + o_0(\varepsilon) \tag{3}$$

où l'on a écrit la forme linéaire $D(x; d)$ comme un produit scalaire entre d et $D(x)$ (cf. théorème de représentation de Riesz), que l'on peut également considéré comme la définition de $f'(x)$.

Proposition 1. *Soit f une application différentiable de $\mathcal{X}$ dans $\mathbb{R}$. Les assertions suivantes sont équivalentes :*

$$J \text{ est convexe sur } \mathcal{X}$$
$$J(v) \geq J(u) + (J'(u) \,|\, v - u),\ \forall u, v \in \mathcal{X}$$
$$(J'(u) - J'(v) \,|\, v - u) \geq 0,\ \forall u, v \in \mathcal{X}$$

3 Optimisation convexe sous contrainte

On cherche à résoudre le problème suivant :

$$\min_{u \in U^{ad}} J(u) \tag{4}$$

où J est une fonction convexe (critère) et U^{ad} un sous-ensemble convexe fermé d'un espace de Hilbert $\mathcal{U}$. On peut sans trop de problème se ramener à un problème d'optimisation sans contrainte, en posant :

$$I_{U^{ad}}(u) = \begin{cases} 0 & \text{si } u \in U^{ad} \\ +\infty & \text{sinon} \end{cases}$$

et en résolvant :

$$\min_{u \in \mathcal{U}} \left[J(u) + I_{U^{ad}}(u) \right] \tag{5}$$

Cependant, cette technique n'a qu'un intérêt pratique, peu théorique.

3.1 Condition d'existence et d'unicité

Lorsque J est convexe, il est clair que tout minimum local est un minimum global (l'ensemble des minima locaux est en fait un convexe). On dispose du théorème suivant :

Théorème 1. *Si J est une fonction convexe et coercive (ie* infini à l'infini*) sur U^{ad}, si U^{ad} est convexe et fermé, alors il existe au moins une solution au problème 4 . L'ensemble des solutions est un* ***sous-ensemble convexe fermé****. Il est réduit à un point (la solution est donc unique) si J est strictement convexe.*

3.2 Inéquation d'Euler

Théorème 2 (Inéquation d'Euler)**.** *Soit $u \in U^{ad}$ (que l'on a supposé convexe). Alors si J est différentiable en u, et u est un point de minimum global de J sur U^{ad}, alors :*

$$(J'(u) \,|\, v - u) \geq 0 \;\forall v \in U^{ad} \tag{6}$$

4 Techniques Lagrangienne

4.1 Description des contraintes

On se donne la forme explicite suivante pour les contraintes. L'espace admissible U^{ad} est décrit par =

$$u \in U^{ad} \Leftrightarrow \left| \begin{array}{lcl} \Theta_i(u) & \leq & 0 \;\; i = 1, \ldots, m \\ \Omega_j(u) & = & 0 \;\; j = 1, \ldots, p \end{array} \right. \tag{7}$$

Pour que cet ensemble soit bien convexe, il suffit que chaque composante soit définie par une fonction convexe. Ceci est assuré par :

- Les fonctions $\Theta_i(u) : \mathcal{U} \to \mathbb{R}$ sont convexes.
- Les fonctions $\Omega_j(u) : \mathcal{U} \to \mathbb{R}$ sont affines (réunion de deux contraintes convexes et concaves)

4.2 Contraintes d'égalité

On a ici seulement des contraintes d'égalité, c'est-à-dire que U^{ad} est simplement donné par :

$$u \in U^{ad} \Leftrightarrow F_j(u) = 0 \;\; j = 1, \ldots, p$$

et on note $U^{ad} = \{v \in \mathcal{U} \,|\, F(v) = 0\}$.

Théorème 3. *Soit $u \in U^{ad}$. On suppose que :*

(i) J est dérivable en u

(ii) Les fonctions F_j sont continûment dérivables dans un voisinage de u.

(iii) La famille $(F_j'(u))_{j\in\{1,\dots,p\}}$ est libre.

Alors si u est un minimum local de J sur U^{ad}, *alors il existe $\lambda_1,\dots,\lambda_p \in \mathbb{R}$ appelés* ***multiplicateurs de Lagrange*** *tels que :*

$$J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0 \tag{8}$$

On peut introduire le Lagrangien du problème de minimisation de J sur U^{ad}, définie sur $\mathcal{U} \times \mathbb{R}^p$ par :

$$\mathcal{L}(v,\mu) = J(v) + \mu \cdot F(v) \tag{9}$$

Si $u \in U^{ad}$ est un minimum local de J sur U^{ad}, le théorème précédent nous assure que dans le cas régulier, $\exists \lambda \in \mathbb{R}^p$ tel que :

$$\left[\begin{array}{lcl} \dfrac{\partial L}{\partial v}(u,\lambda) & = & J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0 \\ \\ \dfrac{\partial L}{\partial \mu}(u,\lambda) & = & F(u) = 0 \ (u \in U^{ad}) \end{array}\right. \tag{10}$$

On peut également noter que le Lagrangien permet d'éliminer les contraintes, au prix du rajout d'une variable car

$$\inf_{v\in U^{ad}} J(v) = \inf_{v\in\mathcal{U}} \sup_{\mu\in\mathbb{R}^p} \mathcal{L}(v,\mu) \tag{11}$$

4.3 Contraintes d'inégalités

On a maintenant que U^{ad} est donné par :

$$U^{ad} = \{v \in \mathcal{U} \,|\, F(v) \leq 0\}$$

.

Définition 5. *Soit $u \in U^{ad}$. L'ensemble $I(u) = \{i \in \{1,\dots,p\}\,,\, F_i(u) = 0\}$ est appelé l'ensemble des contraintes* ***actives*** *en u.*

Définition 6. *On dit que les contraintes sont* ***qualifiées*** *en u si la famille $(F_i'(u))_{i\in I(u)}$ est libre.*

Le théorème se ré-écrit comme suit :

Théorème 4. *Soit $u \in U^{ad}$. On suppose que :*

(i) J est dérivable en u

(ii) Les fonctions F_j sont continûment dérivables dans un voisinage de u.

(iii) Les contraintes sont qualifiées en u

Alors si u est un minimum local de J sur U^{ad}, *alors il existe $\lambda_1,\dots,\lambda_p \in \mathbb{R}^+$ appelés* ***multiplicateurs de Lagrange*** *tels que :*

$$\begin{array}{l} J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0 \\ \lambda_i = 0 \text{ si } F_i(u) < 0\,,\, \forall i \in \{1,\dots,p\} \end{array} \tag{12}$$

5 Théorème de Kuhn-Tucker

Définition 7. *On dit que* $(u,p)) \in U \times P$ *est un point-selle de* $\mathcal{L}$ *sur* $U \times P$ *si :*

$$\forall q \in P \quad \mathcal{L}(u,p) \leq \mathcal{L}(u,p) \leq \mathcal{L}(v,p) \quad \forall v \in U$$

Théorème 5 (Kuhn-Tucker, CNS d'optimalité). *On suppose les fonctions* J*,* $F_1, \dots, F_p$ *convexes continues sur* $\mathcal{U}$ *et dérivables sur* U^{ad}*. On introduit le Lagrangien associé :*

$$\mathcal{L}(v,q) = J(v) + q \cdot F(v) \; \forall (v,q) \in \mathcal{U} \times (\mathbb{R}_+)^p$$

Soit $u \in U^{ad}$ *où les contraintes sont qualifiées. Alors* ***u est un minimum global de*** J ***sur*** U^{ad} ***si et seulement si il existe*** $p \in (\mathbb{R}_+)^p$ *tel que* (u,p) *soit un* ***point-selle*** *du lagrangien* $\mathcal{L}$ *sur* $\mathcal{U} \times (\mathbb{R}_+)^p$*, ou, de manière équivalente :*

$$\begin{aligned} F(u) &\leq 0 \\ p &\geq 0 \\ p \cdot F(u) &= 0 \\ J('u) + \sum_{i=1}^{p} p_i F_i'(u) &= 0 \end{aligned}$$