APPROXIMATION NUMERIQUE ET OPTIMISATION

G. ALLAIRE

23 octobre 2018

CHAPITRE IV (suite)

☞ Cas général: cône des directions admissibles.

☞ Contraintes d'égalité.

☞ Contraintes d'inégalité.

Conditions d'optimalité, cas général: K non convexe

Exemples de K non convexe:

$$K = \{v \in V \text{ tel que } F(v) = 0\} \quad \text{ou} \quad K = \{v \in V \text{ tel que } F(v) \leq 0\}$$

Définition. Pour $v \in K$, on appelle cône des directions admissibles au point v

$$K(v) = \left\{ \begin{array}{l} w \in V \text{ tel que } \exists\, (v^n)_{n \geq 0} \in K \,,\, \exists\, (\varepsilon^n)_{n \geq 0} \in \mathbb{R}_+^* \text{ vérifiant} \\ \\ \lim_{n \to +\infty} v^n = v \,,\, \lim_{n \to +\infty} \varepsilon^n = 0 \,,\, \lim_{n \to +\infty} \frac{v^n - v}{\varepsilon^n} = w \end{array} \right\}$$

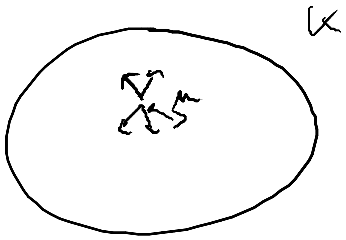
Remarque. $0 \in K(v)$ et $K(v)$ est un cône car si $w \in K(v)$ et $\lambda \geq 0$, alors $\lambda w \in K(v)$.

$u(t) : t \in [0,1] \longrightarrow K$

$u(0) = u$

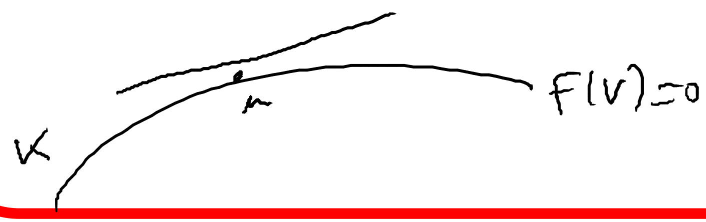
Exemples de cône $K(u)$

☞ Si u est intérieur à K, alors $K(u) = V$.



☞ Si K est convexe, alors $K(u) = \{w = v - u \text{ avec } v \in K\}$.

☞ Soit $F : V \to \mathbb{R}$ fonction régulière. Si $K = \{v \in V \text{ tel que } F(v) = 0\}$ et si $F'(u) \neq 0$, alors $K(u) = [F'(u)]^{\perp}$ (hyperplan tangent en u à la surface K).



Inéquation d'Euler (cas général)

Proposition. Soit u un minimum local de J sur K. Si J est différentiable en u, on a

$$\langle J'(u), w \rangle \geq 0 \quad \forall\, w \in K(u)\ .$$

Remarque. Tout le problème est d'identifier $K(u)$!

Preuve. $w \in K(u)$ $\exists$ $w_n = u + \varepsilon_n w \in K$

$J(w_n) \geq J(u)$ si ε_n petit

$$\cancel{J(u)} + \varepsilon_n \langle J'(u), w \rangle + o(\varepsilon_n) \geq \cancel{J(u)}$$

on divise par $\varepsilon_n > 0$

$$\Longrightarrow \quad \langle J'(u), w \rangle + \cancel{o(1)} \geq 0$$

$\varepsilon_n \longrightarrow 0$

Minimisation avec contraintes d'égalité

Soit $F(v) = (F_1(v), ..., F_M(v))$ une application de V dans $\mathbb{R}^M$. On considère

$$\inf_{v\in K} J(v) \quad \text{avec} \quad K = \{v \in V,\ F(v) = 0\}$$

Théorème. Soit $u \in V$ tel que $F(u) = 0$. On suppose que J est dérivable en u et que les fonctions $(F_i)_{1\le i\le M}$ sont de classe C^1 dans un voisinage de u. On suppose de plus que les vecteurs $\left(F_i'(u)\right)_{1\le i\le M}$ sont libres. Alors, si u est un minimum local de J sur K, il existe $\lambda_1, \ldots, \lambda_M \in \mathbb{R}$, appelés **multiplicateurs de Lagrange**, tels que

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

Remarque. Pour calculer les M multiplicateurs de Lagrange λ_i on peut utiliser les M contraintes $F_i(u) = 0$.

Remarque. Il faut absolument une hypothèse sur les $\left(F_i'(u)\right)_{1\leq i\leq M}$.

Contre-exemple: avec $M = 1$

$$\inf_{F(v)=\|v\|^2=0} J(v)$$

Si le théorème était vrai dans ce cas $J'(u) + \lambda F'(u) = 0$

$u = 0$ qui est le min $F'(u) = u = 0$

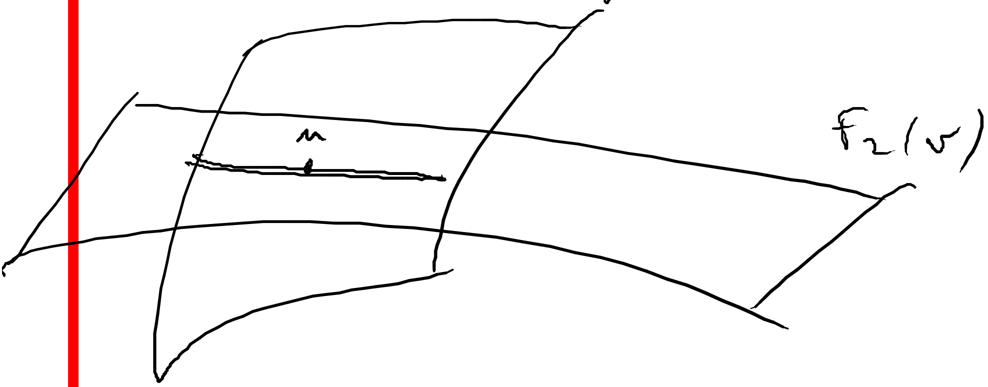
$\Longrightarrow J'(u) = 0$

impossible par exemple avec

$J(v) = \|v - e\|^2$

$e \neq 0$

$F_1(v) = 0$



Preuve

Soit $K(u)$ le cône des directions admissibles au point u

$$K(u) = \left\{ \begin{array}{l} w \in V \text{ tel que } \exists\,(v^n)_{n\geq 0} \in K\,,\ \exists\,(\varepsilon^n)_{n\geq 0} \in \mathbb{R}_+^* \text{ vérifiant} \\ \lim\limits_{n\to+\infty} v^n = u\,,\ \lim\limits_{n\to+\infty} \varepsilon^n = 0\,,\ \lim\limits_{n\to+\infty} \dfrac{v^n - u}{\varepsilon^n} = w \end{array} \right\}$$

Rappel (Proposition 4.2.12): si u est un minimum local, alors

$$\langle J'(u), w\rangle \geq 0 \quad \forall\, w \in K(u)\,.$$

Il faut donc caractériser $K(u)$! Montrons que

$$K(u) = \{w \in V \text{ tel que } \langle F_i'(u), w\rangle = 0,\ 1 \leq i \leq M\} = \bigcap_{i=1}^{M} [F_i'(u)]^{\perp}$$

si c'est vrai alors $K(u)$ est un sev. Donc $w \in K(u) \Rightarrow -w \in K(u)$

$\Rightarrow \langle J'(u), w\rangle \leq 0 \quad \forall w \in K(u) \Leftrightarrow \quad J'(u) \perp K(u)$

$J'(u) \in [F_1'(u) \dots F_p'(u)]$

Preuve (suite)

Soit $w \in K(u)$ $\quad \exists\, w_n = u + \varepsilon_n w \in K$

$\Rightarrow F_i(w_n) = 0 \Rightarrow \underbrace{F_i(u)}_{=0} + \varepsilon_n \langle F_i'(u), w\rangle + o(\varepsilon_n) = 0$

on divise par ε_n et on fait $\varepsilon_n \to 0 \Rightarrow \langle F_i'(u), w\rangle = 0$

$$\Rightarrow K \subset \bigcap_{i=1}^{n} \left[F_i'(u)\right]^{\perp}$$

Réciproque par le théorème des fonctions implicites. Hypothèse $(F_i'(u))_{1 \le i \le n}$ libre

Exemple

Soit A une matrice carrée d'ordre n, symétrique définie positive.

Soit B une matrice rectangulaire de taille $m \times n$ avec $m \leq n$ et $rg(B) = m$.

Soit $b \in \mathbb{R}^n$ et $c \in \mathbb{R}^m$. On considère

$$\inf_{x \in \mathbb{R}^n \text{ tel que } Bx = c} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\}.$$

Lemme. Il existe un unique point de minimum $x^* \in \mathbb{R}^n$ et un unique multiplicateur de Lagrange $p \in \mathbb{R}^m$ qui vérifient

$$Ax^* - b = B^* p \text{ avec } p = \left(BA^{-1}B^*\right)^{-1}(c - BA^{-1}b).$$

Preuve contraintes $Bx = c$ 1 contrainte $\Leftrightarrow$ 1 ligne de $B = b_i$

$(b_i)_{1 \leq i \leq m}$ libre? oui car $rg B = m$ b_i = colonne de B^*

Thm $\Rightarrow$ $\exists\, p_i \in \mathbb{R}$ tel que $J'(x^*) = Ax^* - b = -\sum_{i=1}^{n} p_i b_i$

$\Leftrightarrow$ $\boxed{Ax^* - b = -B^* p}$ on calcule p grâce aux contraintes $Bx^* = c$

Lagrangien

Définition. On appelle **Lagrangien** du problème la fonction

$$\mathcal{L}(v,\mu) = J(v) + \sum_{i=1}^{M} \mu_i F_i(v) = J(v) + \mu \cdot F(v) \qquad \forall (v,\mu) \in V \times \mathbb{R}^M .$$

La nouvelle variable $\mu \in \mathbb{R}^M$ est appelée **multiplicateur de Lagrange** pour la contrainte $F(v) = 0$.

Lemme. Le problème de minimisation est équivalent à

$$\inf_{v \in V,\ F(v)=0} J(v) = \inf_{v \in V} \sup_{\mu \in \mathbb{R}^M} \mathcal{L}(v,\mu).$$

<u>Preuve</u>

$$\sup_{\mu \in \mathbb{R}^M} J(v) + \mu . F(v) = \begin{cases} +\infty & \text{si } F(v) \neq 0 \\ J(v) & \text{si } F(v) = 0 \end{cases}$$

Stationnarité du Lagrangien

Définition du Lagrangien: $\mathcal{L}(v,\mu) = J(v) + \mu \cdot F(v) \qquad \forall (v,\mu) \in V \times \mathbb{R}^M .$

On peut réécrire le précédent théorème sous la forme suivante.

Théorème. Soit $u \in V$ tel que $F(u) = 0$. Si u est un minimum local sur K, et si les vecteurs $\left(F_i'(u)\right)_{1\leq i\leq M}$ sont libres, alors il existe des multiplicateurs de Lagrange $\lambda_1, \ldots, \lambda_M \in \mathbb{R}$, tels que

$$\frac{\partial \mathcal{L}}{\partial v}(u,\lambda) = J'(u) + \lambda \cdot F'(u) = 0 \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \mu}(u,\lambda) = F(u) = 0 \, .$$

Les deux dérivées partielles du Lagrangien s'annulent !

Une interprétation des multiplicateurs de Lagrange

Soit le Lagrangien pour la minimisation de $J(v)$ sous la contrainte $F(v) = c$

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

On étudie la sensibilité du minimum à la variation de c.

On note $u(c)$ et $\lambda(c)$ le point de minimum et le multiplicateur de Lagrange correspondant. On suppose qu'ils sont dérivables par rapport à c. Alors

$$\nabla_c \Big(J(u(c)) \Big) = -\lambda(c).$$

λ donne la dérivée (sans la calculer) du minimum par rapport à c !

En effet

$$\nabla_c \Big(J(u(c)) \Big) = \nabla_c \Big(\mathcal{L}(u(c), \lambda(c), c) \Big) = \frac{\partial \mathcal{L}}{\partial c}(u(c), \lambda(c), c) = -\lambda(c)$$

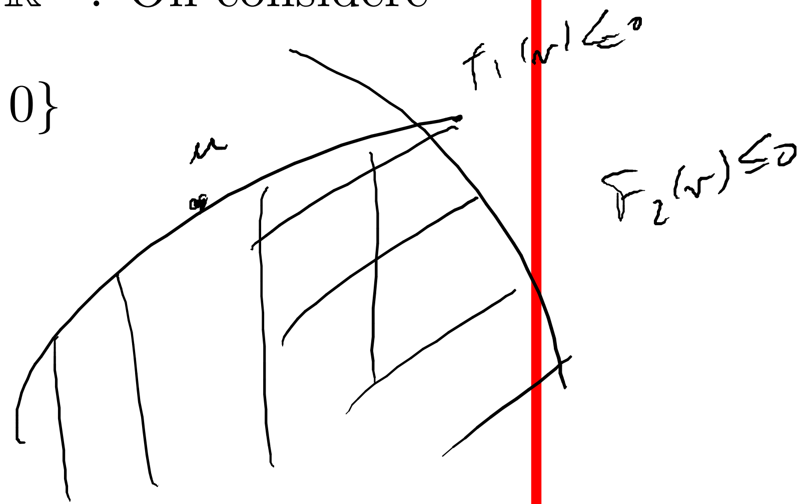
car

$$\frac{\partial \mathcal{L}}{\partial v}(u(c), \lambda(c), c) = 0 \quad , \quad \frac{\partial \mathcal{L}}{\partial \mu}(u(c), \lambda(c), c) = 0 \, .$$

Minimisation avec contraintes d'inégalité

Soit $F(v) = (F_1(v), ..., F_M(v))$ une application de V dans $\mathbb{R}^M$. On considère

$$\inf_{v \in K} J(v) \quad \text{avec} \quad K = \{v \in V,\ F(v) \leq 0\}$$



où $F(v) \leq 0$ signifie que $F_i(v) \leq 0$ pour $1 \leq i \leq M$.

Définitions. Soit u tel que $F(u) \leq 0$.

1. On appelle ensemble des contraintes actives en u
   $I(u) = \{i \in \{1, \ldots, M\}, F_i(u) = 0\}$.

2. On dit que les contraintes d'inégalité sont qualifiées en $u \in K$ si la famille $(F_i'(u))_{i \in I(u)}$ est libre.

Remarque. Il existe d'autres conditions de qualification.

Contraintes et variations

- ☞ Les contraintes inactives $F_i(u) < 0$ ne jouent aucun rôle.
- ☞ Pour les contraintes actives $F_i(u) = 0$ on peut faire des variations dans certaines directions (mais pas dans d'autres).

Théorème. Soit $u \in V$ tel que $F(u) \leq 0$. On suppose que J et les contraintes $(F_i)_{1 \leq i \leq M}$ sont dérivables en u. On suppose de plus que les contraintes sont qualifiées en u. Alors, si u est un minimum local de J sur K, il existe $\lambda_1, \ldots, \lambda_M \in \mathbb{R}^+$, appelés **multiplicateurs de Lagrange**, tels que

$$J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0, \quad \lambda_i \geq 0, \quad F_i(u) \leq 0,$$
$$\lambda_i = 0 \ \text{ si } \ F_i(u) < 0.$$

Remarque. Les conditions sur les λ_i peuvent se réécrire

$$\lambda \cdot F(u) = 0, \quad \lambda \geq 0 \ \text{ et } \ F(u) \leq 0.$$

et s'appellent conditions de complémentarité.

Remarque. Pour calculer les multiplicateurs de Lagrange λ_i des contraintes actives on peut utiliser les égalités $F_i(u) = 0$ si $i \in I(u)$.

Preuve

Soit $K(u)$ le cône des directions admissibles au point u

$$K(u) = \left\{ \begin{array}{l} w \in V \text{ tel que } \exists\, (v^n)_{n\geq 0} \in K \,,\, \exists\, (\varepsilon^n)_{n\geq 0} \in \mathbb{R}_+^* \text{ vérifiant} \\ \lim\limits_{n\to+\infty} v^n = u \,,\ \lim\limits_{n\to+\infty} \varepsilon^n = 0 \,,\ \lim\limits_{n\to+\infty} \dfrac{v^n - u}{\varepsilon^n} = w \end{array} \right\}$$

Rappel (Proposition 4.2.12): si u est un minimum local, alors

$$\langle J'(u), w\rangle \geq 0 \quad \forall\, w \in K(u) \,.$$

Il faut donc caractériser $K(u)$!

Lemme de Farkas. Soient $a_1, \ldots, a_M$ une famille libre de V. On considère les ensembles

$$\mathcal{K} = \Big\{ w \in V \,, \quad \langle a_i, w \rangle \leq 0 \ \text{ pour } \ 1 \leq i \leq M \Big\} \,,$$

et

$$\hat{\mathcal{K}} = \Big\{ q \in V \,, \quad \exists \, \lambda_1, \ldots, \lambda_M \geq 0 \,, \; q = - \sum_{i=1}^{M} \lambda_i a_i \Big\} \,.$$

Alors pour tout $p \in V$, on a l'implication

$$\langle p, w \rangle \geq 0 \;\forall \, w \in \mathcal{K} \implies p \in \hat{\mathcal{K}} \,.$$

Remarque. La réciproque étant évidente, il s'agit en fait d'une équivalence.

Preuve du lemme de Farkas

Lagrangien

Définition. On appelle Lagrangien du problème la fonction

$$\mathcal{L}(v,\mu) = J(v) + \sum_{i=1}^{M} \mu_i F_i(v) = J(v) + \mu \cdot F(v) \qquad \forall (v,\mu) \in V \times (\mathbb{R}^+)^M.$$

La nouvelle variable **positive** $\mu \in (\mathbb{R}^+)^M$ est appelée multiplicateur de Lagrange pour la contrainte $F(v) \leq 0$.

Lemme. Le problème de minimisation est équivalent à

$$\inf_{v \in V,\ F(v) \leq 0} J(v) = \inf_{v \in V} \sup_{\mu \in (\mathbb{R}^+)^M} \mathcal{L}(v,\mu).$$

Chic
au COMPROMP

Stationnarité du Lagrangien

Définition du Lagrangien: $\mathcal{L}(v,\mu) = J(v) + \mu \cdot F(v) \qquad \forall (v,\mu) \in V \times (\mathbb{R}^+)^M$.

On peut réécrire le précédent théorème sous la forme suivante.

Théorème. On suppose que les contraintes sont qualifiées en u tel que $F(u) \leq 0$. Alors, si u est un minimum local, il existe $\lambda_1, \ldots, \lambda_M \geq 0$, appelés multiplicateurs de Lagrange, tels que

$$\frac{\partial \mathcal{L}}{\partial v}(u,\lambda) = 0 \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \mu}(u,\lambda) \cdot (\mu - \lambda) \leq 0 \quad \forall \mu \in (\mathbb{R}^+)^M$$

Remarque. La deuxième condition est l'inéquation d'Euler pour la **maximisation** du Lagrangien par rapport à μ dans le convexe fermé $(\mathbb{R}^+)^M$.

Preuve

Les conditions d'optimalité sont

$$J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0\,, \quad \lambda_i \geq 0\,, \quad F_i(u) \leq 0\,, \quad \lambda_i = 0 \text{ si } F_i(u) < 0\,.$$

Cette condition est bien la stationnarité du Lagrangien puisque

$$\frac{\partial \mathcal{L}}{\partial v}(u,\lambda) = J'(u) + \lambda \cdot F'(u) = 0,$$

et que l'inéquation d'Euler

$$\frac{\partial \mathcal{L}}{\partial \mu}(u,\lambda) \cdot (\mu - \lambda) = F(u) \cdot (\mu - \lambda) \leq 0 \quad \forall \mu \in (\mathbb{R}^+)^M$$

implique que $\lambda_i = 0$ si $F_i(u) < 0$.

Exemple: régularisation d'un signal

Soit A une matrice carrée d'ordre n, symétrique définie positive.

Soit $b \in \mathbb{R}^n$ et $\epsilon > 0$. On considère

$$\min_{x \in \mathbb{R}^n \text{ tel que } \|x-b\|^2 \leq \epsilon^2} Ax \cdot x.$$

Il existe une unique solution $x^* \in \mathbb{R}^n$ qui vérifie

$$\text{soit } x^* = 0 \quad \text{ et } \|b\| < \epsilon,$$

$$\text{soit } x^* = (\,\mathrm{Id} + \lambda^{-1} A)^{-1} b \text{ avec } \lambda > 0 \text{ tel que } \|x^* - b\| = \epsilon.$$