OPTIMISATION ET CONTRÔLE

Grégoire Allaire
Département de mathématiques appliquées
Ecole Polytechnique

1. Existence de minima en dimension infinie
2. Conditions d'optimalité
   1. Minimisation sur un convexe: inégalité d'Euler
   2. Minimisation avec contraintes d'égalité: multiplicateurs de Lagrange

Quelques informations

Equipe enseignante: Grégoire Allaire, Beniamin Bogosel, Ugo Boscain, Alexandre Ern, Jean-Frédéric Gerbeau, Michael Goldman, Mazyar Mirrahimi, Flore Nabet.

Délégué des élèves: Thomas Li

Rappel:

1. Faire le QCM sur moodle avant la PC.
2. Devoir obligatoire disponible le 4 avril à rendre sur moodle avant le 18 avril.

Rappels du premier cours.

Définition. On dit qu'une fonction J définie sur un ensemble convexe K est fortement convexe ou α-convexe s'il existe $\alpha > 0$ tel que, pour tout $u, v \in K$ et pour tout $0 \leq \theta \leq 1$,

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) - \frac{\alpha}{2}\theta(1-\theta)\|u-v\|^2$$

Proposition. Si J est fortement convexe continue sur un convexe fermé non vide $K \subset V$, alors il existe deux constantes $\gamma > 0$ et $\eta \in \mathbb{R}$ telles que

$$J(v) \geq \gamma\|v\|^2 + \eta \quad \forall\, v \in K\ .$$

Existence de minima (1)

Théorème. Soit K un convexe fermé non vide d'un Hilbert V et J une fonction α-convexe continue sur K. Alors, il existe un unique minimum u de J sur K et toute suite minimisante de J dans K converge vers u.

Preuve. Soit (u^n) une suite minimisante de J sur K. On applique la définition de la forte convexité pour $\theta = 1/2$

$$\frac{\alpha}{8}\|u^n - u^m\|^2 + J\left(\frac{u^n + u^m}{2}\right) \leq \frac{J(u^n) + J(u^m)}{2} .$$

On soustrait de part et d'autre la valeur $\inf_{v \in K} J(v)$ (qui est finie car J est minorée par une fonction quadratique) et comme

$$0 \leq J\left(\frac{u^n + u^m}{2}\right) - \inf_{v \in K} J(v) ,$$

on en déduit

$$\frac{\alpha}{8}\|u^n - u^m\|^2 \leq \frac{1}{2}\left(J(u^n) - \inf_{v \in K} J(v)\right) + \frac{1}{2}\left(J(u^m) - \inf_{v \in K} J(v)\right) .$$

Existence de minima (2)

On a obtenu

$$\frac{\alpha}{8}\|u^n - u^m\|^2 \leq \frac{1}{2}\left(J(u^n) - \inf_{v\in K} J(v)\right) + \frac{1}{2}\left(J(u^m) - \inf_{v\in K} J(v)\right) .$$

Donc la suite (u^n) est de Cauchy (dans un Hilbert), et ainsi elle converge vers une limite u, qui est nécessairement un minimum de J sur K puisque J est continue et K fermé.

L'unicité du point de minimum a déjà été montrée pour J strictement convexe.

Existence de minima (3)

On admet la généralisation suivante.

Théorème. Soit K un convexe fermé non vide d'un Hilbert V et J une fonction convexe continue sur K qui est "infinie à l'infini"

$$\forall (u^n)_{n\geq 0} \text{ suite dans } K\ ,\quad \lim_{n\to+\infty} \|u^n\| = +\infty \Longrightarrow \lim_{n\to+\infty} J(u^n) = +\infty\ .$$

Alors, il existe un minimum u de J sur K.

Ingrédients de la preuve (hors programme):

1. notion de convergence faible dans un espace de Hilbert,
2. une fonction convexe est le max de ses hyperplans tangents

$$J(u) = \max_{v\in V} \Big(J(v) + \langle J'(v), u - v\rangle \Big).$$

Remarques

1. Les contre-exemples du premier cours n'étaient pas des fonctions convexes.
2. Il existe d'autres cadres que la convexité pour démontrer l'existence de minima en dimension infinie...

II - Conditions d'optimalité

Pour caractériser ou calculer les solutions de

$$\inf_{v \in K \subset V} J(v)$$

on écrit les conditions d'optimalité qui sont des conditions sur les dérivées de J aux éventuels points de minimum.

- Très utile d'un point de vue théorique.
- Très utile aussi pour les algorithmes numériques d'optimisation.

Motivation: exemple en 1-d

Soit une fonction $f : [a, b] \to \mathbb{R}$ de classe C^1, avec $a < b$.
On veut résoudre

$$\inf_{x \in [a,b]} f(x)$$

Il est bien connu que les points de minimum de f se trouvent

- soit en un point x tel que $f'(x) = 0$,
- soit en a si $f'(a) \geq 0$,
- soit en b si $f'(b) \leq 0$.

(condition nécessaire mais pas suffisante !)

Remarque. Le signe de $f''(x)$ permet aussi de distinguer entre minimum et maximum.

Différentiabilité

Soit V un espace de Hilbert muni d'un produit scalaire $\langle u, v \rangle$ et d'une norme associée $\|u\|$.

Définition. On dit que la fonction J, définie sur un voisinage de $u \in V$ à valeurs dans $\mathbb{R}$, est différentiable au sens de Fréchet en u s'il existe une forme linéaire continue $L : \ V \to \mathbb{R}$ telle que

$$J(u+w) = J(u) + L(w) + o(w) \quad \text{avec} \quad \lim_{w \to 0} \frac{|o(w)|}{\|w\|} = 0 \, .$$

On appelle L la différentielle (ou la dérivée, ou le gradient) de J en u et on note $L = J'(u)$.

- Une forme linéaire L sur V est continue s'il existe $C > 0$ telle que $|L(w)| \leq C\|w\|$ pour tout $w \in V$.
- En dimension finie (par exemple, si $V = \mathbb{R}^N$) toute forme linéaire est continue, donc il n'y a rien à vérifier en ce qui concerne la continuité !

Remarques

- Dans un espace de Hilbert, on sait, grâce au théorème de représentation de Riesz, que pour toute forme linéaire continue L il existe un unique $p \in V$ tel que $\langle p, w \rangle = L(w)$. Par conséquent, on peut simplifier la définition en

$$J(u+w) = J(u) + \langle p, w \rangle + o(w) \quad \text{avec} \quad \lim_{w \to 0} \frac{|o(w)|}{\|w\|} = 0 \, .$$

- Pour calculer la dérivée (mais pas pour prouver la différentiabilité Fréchet !) on peut calculer la dérivée directionnelle. Pour $t \in \mathbb{R}$ et $u, v \in V$, on pose

$$j(t) = J(u + tv)$$

et on dérive $t \to j(t)$

$$j'(0) = J'(u)(v) = \langle J'(u), v \rangle.$$

Convexité

Proposition. Soit J une application différentiable de V dans $\mathbb{R}$. Les définitions suivantes de la convexité sont équivalentes

$$(1) \quad J(\theta u+(1-\theta)v) \leq \theta J(u)+(1-\theta)J(v) \quad \forall\, u,v\in K\,,\ \forall\,\theta\in[0,1]$$

$$(2) \qquad J(u)\geq J(v)+\langle J'(v),u-v\rangle \quad \forall\, u,v\in V$$

$$(3) \qquad \langle J'(u)-J'(v),u-v\rangle\geq 0 \quad \forall\, u,v\in V$$

Remarque. Même type de résultat pour les fonctions α-convexes.

Preuve. On réécrit (1) sous la forme

$$J(v)+\theta\langle J'(v),u-v\rangle+o(\theta)=J(v+\theta(u-v))\leq J(v)+\theta(J(u)-J(v))$$

On élimine $J(v)$, on divise par θ qui tend vers 0 pour obtenir (2).

Fin de la preuve

Pour obtenir (3) on additionne (2) à lui-même avec u et v échangés.
Pour montrer (3) $\Rightarrow$ (1) on introduit $\varphi(t) = J(u + t(v - u))$ qui vérifie $\varphi'(t) = \langle J'(u + t(v - u)), v - u\rangle$. Donc (3) implique

$$\varphi'(t) - \varphi'(s) \geq 0 \quad \text{si} \quad t \geq s\,.$$

Soit $\theta \in]0, 1[$. En intégrant cette inégalité de $t = \theta$ à $t = 1$ et de $s = 0$ à $s = \theta$, on obtient

$$\theta(\varphi(1) - \varphi(\theta)) - (1 - \theta)(\varphi(\theta) - \varphi(0)) \geq 0\,,$$

qui est équivalent à

$$\theta\varphi(1) + (1 - \theta)\varphi(0) - \varphi(\theta) \geq 0\,, \quad \text{CQFD.}$$

Dérivée seconde

Définition. Soit J une fonction de V dans $\mathbb{R}$. On dit que J est deux fois dérivable en $u \in V$ si J est dérivable dans un voisinage de u et si sa dérivée $J'(u)$ est dérivable en u. On note $J''(u)$ la dérivée seconde de J en u qui vérifie

$$J'(u+w) = J'(u) + J''(u)w + o(w) \,, \quad \text{avec} \quad \lim_{w \to 0} \frac{\|o(w)\|_{V'}}{\|w\|_V} = 0 \,.$$

Lemme. Si J est deux fois dérivable de V dans $\mathbb{R}$, elle vérifie

$$J(u+w) = J(u) + J'(u)w + \frac{1}{2} J''(u)(w,w) + o(\|w\|^2),$$

où $J''(u)$ est identifiée à une forme bilinéaire continue sur $V \times V$.

Convexité (bis)

Proposition. Soit J deux fois dérivable de V dans $\mathbb{R}$. Les définitions suivantes de la convexité sont équivalentes

$$(1)\quad J(\theta u+(1-\theta)v) \leq \theta J(u)+(1-\theta)J(v) \quad \forall\, u,v \in V\,,\ \forall\, \theta \in [0,1]$$

$$(2)\qquad J''(v)(w,w) \geq 0 \quad \forall\, v,w \in V\,.$$

Preuve. (1) est équivalent à

$$\langle J'(u)-J'(v), u-v\rangle \geq 0$$

où l'on choisit $u = v + \delta w$, avec $\delta > 0$, et on utilise $J'(v+\delta w) = J'(v) + \delta J''(v)w + o(\delta)$ pour obtenir (2). Réciproquement, (2) pour un certain $\tilde{v}$ conduit à

$$\langle J'(u)-J'(v), u-v\rangle = J''(\tilde{v})(u-v,u-v) \geq 0$$

Théorème (Inéquation d'Euler). Soit un convexe K. Soit J différentiable en $u \in K$. Si u est un point de minimum local de J sur K, alors

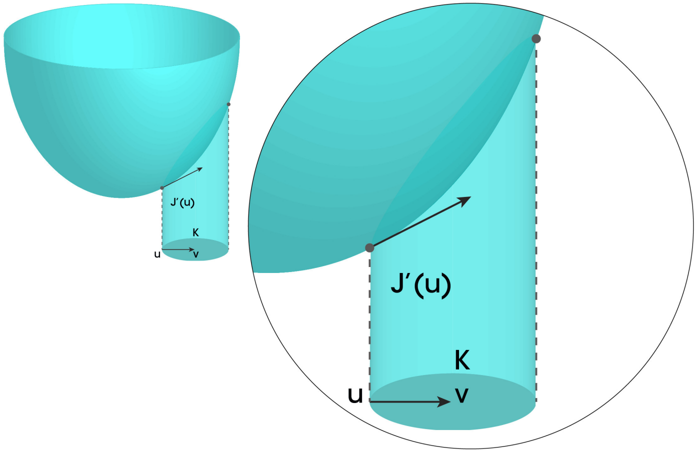
$$\langle J'(u), v - u \rangle \geq 0 \quad \forall\, v \in K\ .$$

Réciproquement, si $u \in K$ vérifie cette inéquation et si J est convexe, alors u est un minimum global de J sur K.

Remarques.

- Si u est intérieur à K, on obtient l'équation d'Euler $J'(u) = 0$.
- L'inéquation d'Euler est seulement nécessaire en général. Pour les fonctions convexes, elle est nécessaire et suffisante.

Inéquation d'Euler



Interprétation. La dérivée directionnelle de J dans la direction rentrante $(v - u)$ est positive.

Démonstration

Pour $v \in K$ et $h \in]0, 1]$, $u + h(v - u) \in K$, car K est convexe. Donc, pour h petit, comme u est un minimum local

$$\frac{J(u + h(v - u)) - J(u)}{h} \geq 0 \ .$$

On utilise la différentiabilité pour obtenir

$$\langle J'(u), v - u \rangle + o(1) \geq 0$$

On en déduit le résultat en faisant tendre h vers 0.

Réciproquement, si J est convexe, la condition d'optimalité pour u implique que u est un minimum global sur K

$$J(v) \geq J(u) + \langle J'(u), v - u \rangle \geq J(u) \quad \forall v \in K \ .$$

Exemple 1

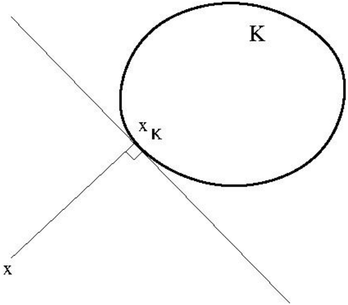
Projection sur un convexe fermé K. Pour $x \in V$, on minimise

$$\|x - x_K\| = \min_{y \in K} \|x - y\|.$$

La condition nécessaire et suffisante d'optimalité est

$$x_K \in K, \quad \langle x_K - x, y - x_K \rangle \geq 0 \quad \forall y \in K,$$

ce qui veut dire que x_K est la projection orthogonale de x sur K.



Exemple 1 bis

Projection sur le cône positif $K = (\mathbb{R}^+)^n$.

Pour $x \in \mathbb{R}^n$, on cherche la projection orthogonale $x_K \in K$

$$\|x - x_K\| = \min_{y \in K} \|x - y\|.$$

La condition nécessaire et suffisante d'optimalité est

$$x_K \in K, \quad \langle x_K - x, y - x_K \rangle \geq 0 \quad \forall y \in K.$$

On choisit y tel que $y_i = (x_K)_i$ pour tout $i \neq j$ et on en déduit

$$(x_K)_j = x_j \text{ si } x_j > 0, \qquad (x_K)_j = 0 \text{ si } x_j \leq 0$$

Autrement dit, $(x_K)_j = x_j^+ = \max(0, x_j)$.

Exemple 2

Soit A une matrice carrée d'ordre n, symétrique définie positive.
Soit B une matrice rectangulaire de taille $m \times n$ avec $m \leq n$.
Soit b un vecteur de $\mathbb{R}^n$. On considère

$$\inf_{x \in \ker B} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\}.$$

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

$$Ax^* - b = B^* p \;\text{ avec } p \in \mathbb{R}^m.$$

avec p une solution de $BA^{-1}B^*p = -BA^{-1}b$.
En effet, comme $K = \ker B$ est un sous-espace vectoriel, l'inéquation d'Euler implique $J'(x^*) \perp \ker B$. Le résultat se déduit de $J'(x^*) = Ax^* - b$ et $(\ker B)^\perp = \mathrm{Im} B^*$.

Condition d'optimalité du 2ème ordre

Proposition. On suppose que $K = V$ et que J est deux fois dérivable en u. Si u est un point de minimum local de J, alors

$$(1) \qquad J'(u) = 0 \quad \text{et} \quad J''(u)(w,w) \geq 0 \quad \forall\, w \in V\,.$$

Réciproquement, si, pour tout v dans un voisinage de u,

$$(2) \qquad J'(u) = 0 \quad \text{et} \quad J''(v)(w,w) \geq 0 \quad \forall\, w \in V\,,$$

alors u est un minimum local de J.

Preuve. Si u est minimum, on sait déjà que $J'(u) = 0$ et on a

$$J(u+w) = J(u) + J'(u)w + \frac{1}{2}J''(u)(w,w) + o(\|w\|^2) \geq J(u)$$

d'où l'on déduit (1). Si (2) est vrai, alors une formule de Taylor avec reste exact implique que u est un minimum local.

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 $\lambda w \in K(v)$.

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. Soit une direction admissible $w \in K(u)$. Par définition, il existe une suite

$$\lim_{n \to +\infty} v^n = u \ , \quad \lim_{n \to +\infty} \varepsilon^n = 0 \ , \quad \lim_{n \to +\infty} \frac{v^n - u}{\varepsilon^n} = w$$

Pour n suffisamment grand,

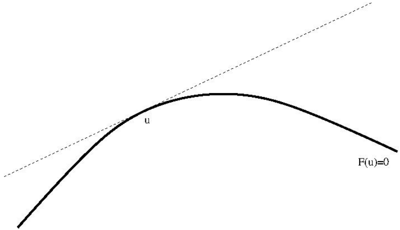
$$J(u) \leq J(v^n) = J\big(u + \varepsilon^n w + o(\varepsilon^n)\big) = J(u) + \varepsilon^n \langle J'(u), w \rangle + o(\varepsilon^n).$$

En soustrayant $J(u)$ et en divisant par ε^n on obtient le résultat.

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

  Il suffit de prendre $v^n = u + \frac{1}{n}w$ pour tout $w \in V$ et $n \in \mathbb{N}$ grand.
- Si K est convexe, alors $K(u) = \{w = v - u \text{ avec } v \in K\}$.

  Il suffit de prendre $v^n = u + \frac{1}{n}(v - u)$ pour tout $v \in K$ et $n \in \mathbb{N}$.
- Soit $F : V \to \mathbb{R}$ et $K = \{v \in V \text{ tel que } F(v) = 0\}$. Si $F'(u) \neq 0$, alors $K(u) = [F'(u)]^\perp$ (hyperplan tangent en u à la surface K).



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 $(F_i)_{1 \leq i \leq M}$ sont de classe C^1 dans un voisinage de u. On suppose de plus que les vecteurs $\left(F_i'(u)\right)_{1 \leq i \leq M}$ 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 sur l'hypothèse

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

Contre-exemple: avec $M = 1$

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

Le seul point admissible est $v = 0$, c'est donc un point de minimum !
On a $F'(0) = 0$. Par contre, il n'est pas toujours vrai que $J'(0) = 0$... Prendre, par exemple, $J(v) = \|v - v_0\|^2$ avec $v_0 \neq 0$.

Preuve

L'inéquation d'Euler (cas général) pour un minimum local u est

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

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

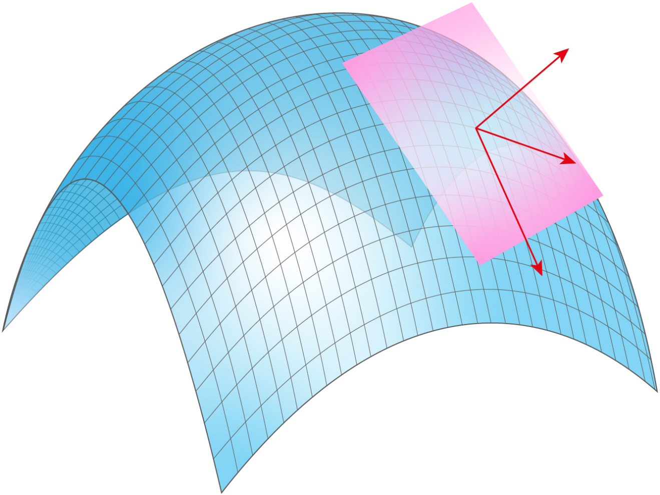
$$K(u) = \left\{ \begin{array}{l} w \in V \text{ tel que } \exists\, (v^n)_{n\geq 0} \in K\,,\ \exists\, (\varepsilon^n)_{n\geq 0} \in \mathbb{R}_+^* \text{ vérifiant} \\ \lim\limits_{n\to+\infty} v^n = u\,,\ \lim\limits_{n\to+\infty} \varepsilon^n = 0\,,\ \lim\limits_{n\to+\infty} \dfrac{v^n - u}{\varepsilon^n}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

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

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

c'est-à-dire que $K(u)$ est l'hyperplan tangent à la variété K en u.

Hyperplan tangent $K(u)$



En bleu $K = \{F(v) = 0\}$ et en rose le plan tangent $[F'(u)]^{\perp}$.

Preuve (suite)

Soit $w \in K(u)$: il existe $v^n \in V$ et $\varepsilon^n > 0$ tel que $F(v^n) = 0$,

$$\lim_{n\to+\infty} v^n = u, \quad \lim_{n\to+\infty} \varepsilon^n = 0, \quad \lim_{n\to+\infty} \frac{v^n - u}{\varepsilon^n} = w.$$

Un développement de Taylor conduit à

$$0 = F_i(v^n) = F_i(u) + \varepsilon^n \langle F_i'(u), w\rangle + o(\varepsilon^n)$$

et comme $F_i(u) = 0$, divisant par ε^n, on obtient

$$0 = \langle F_i'(u), w\rangle + o(1),$$

c'est-à-dire, à la limite $n \to +\infty$, que $w \in \bigcap_{i=1}^{M} [F_i'(u)]^\perp$.

L'inclusion inverse découle du Lemme 2.5.7 du polycopié (théorème de l'application surjective) qu'on ne détaille pas ici.

Preuve (fin)

Comme $K(u)$ est un espace vectoriel, on peut prendre successivement w et $-w$ dans l'inéquation d'Euler

$$\langle J'(u), w\rangle = 0 \quad \forall\, w \in K(u) = \bigcap_{i=1}^{M} \left[F_i'(u)\right]^{\perp},$$

c'est-à-dire que $J'(u) \in \left(\bigcap_{i=1}^{M} [F_i'(u)]^{\perp}\right)^{\perp} = \oplus_{i=1}^{M} [F_i'(u)]$.
Autrement dit, $J'(u)$ est engendré par les $\left(F_i'(u)\right)_{1\leq i\leq M}$.
(Notons que les multiplicateurs de Lagrange sont définis de manière unique.)

Exemple

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

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

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

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

On applique le théorème précédent avec $F_i(x) = (Bx - c)_i$ dont l'hypothèse est vérifiée (les lignes de B sont libres car $rg(B) = m$). On calcule p avec la contrainte $Bx^* = c$.

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

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

Preuve. Si $F(v) = 0$ on a évidemment $J(v) = \mathcal{L}(v,\mu)$ pour tout $\mu \in \mathbb{R}^M$, tandis que, si $F(v) \neq 0$, alors $\sup_{\mu \in \mathbb{R}^M} \mathcal{L}(v,\mu) = +\infty$, d'où l'on déduit le résultat.

Stationnarité du Lagrangien

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

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 $\left(F_i'(u)\right)_{1 \le i \le M}$ 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

Pour $\epsilon \in \mathbb{R}^M$, J et F de classe C^1, on considère le problème d'optimisation

$$J_{\min}(\epsilon) = \inf_{v \in V,\ F(v)=\epsilon} J(v),$$

où le niveau des contraintes est paramétré par ϵ.

Lemme. On suppose que, pour ϵ petit, le problème ci-dessus admet une solution unique, notée u_ϵ, supposée différentiable par rapport à ϵ en 0. Pour $\epsilon = 0$, on suppose que les vecteurs $\left(F_i'(u_0)\right)_{1\le i\le M}$ sont libres et on note $\lambda \in \mathbb{R}^M$ le multiplicateur de Lagrange associé. Alors, pour $1 \le i \le M$, on a

$$\lambda_i = -\frac{\partial J_{\min}}{\partial \epsilon_i}(0).$$

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

Une interprétation des multiplicateurs de Lagrange (fin)

Preuve. Pour $(v,\mu,\epsilon) \in V \times \mathbb{R}^M \times \mathbb{R}^M$, on définit le Lagrangien

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

Comme la solution u_ϵ vérifie les contraintes, on a

$$\mathcal{L}(u_\epsilon,\mu,\epsilon) = J(u_\epsilon) = J_{\min}(\epsilon).$$

En dérivant cette relation en $\epsilon = 0$, on obtient

$$\frac{\partial J_{\min}}{\partial \epsilon_i}(0) = \langle \frac{\partial \mathcal{L}}{\partial v}(u_0,\mu,0), \frac{\partial u_\epsilon}{\partial \epsilon_i}(0)\rangle + \frac{\partial \mathcal{L}}{\partial \epsilon_i}(0)$$

$$= \langle J'(u_0) + \mu \cdot F'(u_0), \frac{\partial u_\epsilon}{\partial \epsilon_i}(0)\rangle - \mu_i.$$

On remplace alors μ par λ et on utilise la condition d'optimalité pour u_0 afin de conclure.