# Optimisation différentiable : Théorie et Algorithmes

Séance 2 : Résultats d'existence et conditions d'optimalité

Abdelilah HAKIM

30/04/2024

Ecole Centrale Casablanca

# Outline

# Introduction

- Les résultats d'existence de solutions aux problèmes d'optimisation
  - cas de la dimension finie : Simple
  - cas de la dimension infinie ; Délicat
- Les conditions d'optimalité ; Sans contraites et avec contraintes.
- Caractérisation des éventuelles solutions à l'aide de la notion de differentiabilité
- Conditions nécessaires d'optimalité
  - L'ensemble des contraintes est convexe ; Inéquation d'Euler
  - Contraintes égalités ou inégalités ; Multiplicateurs de Lagrange
- Le théorème de Kuhn et Tucker

# Définitions et Notations

# Position du problème

- $V$ : Espace de Hilbert, muni de $\|.\|$

- $V = \mathbb{R}^N$ : Cas d'un espace Euclidien

- $K \subset V$ : Ensemble des contraintes ou bien ensemble des solutions admissibles

- $J$ : Fonction définie sur $K$ à valeurs dans $\mathbb{R}$

- $J$ : Fonction coût ou bien fonction objective

- Position du problème :

$$(P)\begin{cases} \text{Trouver} \quad \bar{x} \in K \quad \textit{tel} \quad \textit{que} : \\ J(\bar{x}) = \inf_{x \in K} J(x) \end{cases} \tag{1}$$

## Minimum local/global

**Définition 1.**

- Minimum local On dit que u est un minimum (ou un point de minimum) local de $J$ sur $K$ si et seulement si

$$u \in K \quad \text{et } \exists \delta > 0, \forall v \in K, \|v - u\| < \delta \Longrightarrow J(v) \geq J(u).$$

- Minimum global On dit que u est un minimum (ou un point de minimum) global de $J$ sur $K$ si et seulement si

$$u \in K \quad \text{et} \quad J(v) \geq J(u) \quad \forall v \in K.$$

**Remarque 1.**

Si $J$ est convexe, alors un minimum local est un minimum global.

# Suite minimisante

**Note 1.**

- Caractérisation de *l'inf*

$$m = \inf F \quad \Leftrightarrow \quad \begin{cases} \forall x \in F, m \leq x \\ \forall \varepsilon > 0, \exists x \in F, x - m < \varepsilon \end{cases}$$

- Caractérisation séquentielle :

$$m = \inf F \quad \Leftrightarrow \quad \exists \left(u_n\right)_{n \in \mathbf{N}} \in \mathbf{F}^{\mathbf{N}}, \lim_{n \to +\infty} u_n = m$$

**Définition 2.**

Suite minimisante : Une suite minimisante de $J$ dans $K$ est une suite $\left(u^n\right)_{n\in\mathbb{N}}$ telle que

$$u^n \in K \quad \text{et} \quad \lim_{n\to+\infty} J\left(u^n\right) = \inf_{v\in K} J(v).$$

Par la définition même de l'infimum de $J$ sur $K$ il existe toujours des suites minimisantes.

# Fonction coercive

**Définition 3.**

Soit $C$ une partie non bornée d'un espace vectoriel normé $E$ et $f : C \to \mathbb{R}$. On dit que $f$ est coercive si

$$\lim_{\|x\| \to +\infty, x \in C} f(x) = +\infty.$$

Autrement dit, $f$ est coercive si

$$\forall A > 0, \exists M > 0, \forall x \in C, \|x\| \geq M \Longrightarrow f(x) \geq A$$

**Exemple 1.**

- Une forme bilinéaire $a : X \times X \to \mathbb{R}$ est dite coercive si elle vérifie :

$$\exists \alpha > 0, \quad \forall x \in X : \quad a(x, x) \geqslant \alpha \|x\|^2$$

- Opérateur d'un espace de Hilbert dans lui-même : Un opérateur $A$ d'un espace de Hilbert $H$ dans lui-même est dit coercif ssi

$$\lim_{\|x\| \to +\infty} \frac{\langle A(x), x \rangle}{\|x\|} = +\infty$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire de $H$ et $\| \cdot \|$ la norme associée.

# Résultat d'existence en dimension finie

# Résultat d'existence en dimension finie

- $V = \mathbb{R}^N$ muni du produit scalaire usuel $u \cdot v = \sum_{i=1}^{N} u_i v_i$ et de la norme euclidienne $\|u\| = \sqrt{u \cdot u}$.
- Conditions suffisantes d'existence d'un minimum
- Théorème le plus classique en dimension finie

**Théorème 1.**

Soit $K \subset \mathbb{R}^N$, soit $J$ une fonctionnelle continue sur $\Omega$ contenant $K$, et $K$ fermé.
Si $K$ est compact, ou si $J$ est coercive, alors $J$ a au moins un minimum sur $K$.
On peut extraire de toute suite minimisante sur $K$ une sous-suite convergeant vers un point de minimum sur $K$.

# Preuve

**Preuve:**

1. **K compact :**
   - Soit $I = \inf_{v \in K \subset V} J(v)$
   - Suite minimisante $u_n \in K$ telle que $J(u_n)$ tend vers $I$
   - Sous-suite extraite $u_{n'}$ convergente vers $a \in K$
   - $J$ continue, $J(u_{n'})$ tend vers $J(a)$, et donc $J(a) = I$
2. **J est coercive :**
   - Si K n est pas compacte, on montre par l'absurde que la suite minimisante est bornée.
   - Dans le cas contraire, il existerai une sous-suite extraite $u_{n'}$ telle que $|u_{n'}| \to +\infty$
   - $J$ est coercive, $J(u_{n'})$ ne converge pas vers $I$
   - Soit $B$ une boule fermée contenant tous les termes de la suite. Alors $u_n \in K \cap B$ est une suite dans un compact.

**Remarque 2.**

- La condition à l'infini garanti que la suite minimisante est bornée et puisque on est en dimension finie on conclut grace au théorème de Bolzano-Weirstrass.
- Dans un espace de dimension infinie, un fermé borné n'est pas necessairement compact.

# Existence d'un minimum en dimension infinie

# Cas de dimension infinie

**Difficulté : Les fermés bornés ne sont pas compacts !**

Contre exemple :

- Soit l'espace de Hilbert (de dimension infinie) des suites de carré sommable dans $\mathbb{R}$

$$\ell_2(\mathbb{R}) = \left\{ x = (x_i)_{i\geq 1} \text{ tel que } \sum_{i=1}^{+\infty} x_i^2 < +\infty \right\},$$

- Produit scalaire $\langle x, y\rangle = \sum_{i=1}^{+\infty} x_i y_i$
- $J$ définie sur $\ell_2(\mathbb{R})$ par :

$$J(x) = \left(\|x\|^2 - 1\right)^2 + \sum_{i=1}^{+\infty} \frac{x_i^2}{i}.$$

- Problème : $K = \ell_2(\mathbb{R})$,

$$\inf_{x\in\ell_2(\mathbb{R})} J(x),$$

- Vérifier que

$$\left(\inf_{x\in\ell_2(\mathbb{R})} J(x)\right) = 0$$

- Soit $x^n$ dans $\ell_2(\mathbb{R})$ définie par $x_i^n = \delta_{in}$ pour tout $i \geq 1$, où $\delta_{in}$ est le symbole de Kronecker qui vaut $1$ si $i = n$ et $0$ sinon.
- vérifier que

$$J(x^n) = \frac{1}{n} \to 0 \text{ quand } n \to +\infty$$

- $J$ est positive, on en déduit que $x^n$ est une suite minimisante et que la valeur du minimum est nulle.
- Même si $x^n$ est bornée, il est évident qu'il n'existe aucun $\bar{x} \in \ell_2(\mathbb{R})$ tel que $J(\bar{x}) = 0$. Par conséquent, il n'existe pas de point de minimum.

# Optimisation convexe

- Pour remédier à ce défaut de compacité en dimension infinie nous nous intéressons à une classe particulière de problèmes, très importants en théorie comme en pratique : les problèmes de minimisation convexe.

# Un résultat de la convexité

**Proposition 1.**

1. Si J est une fonction convexe sur un ensemble convexe $K$, tout point de minimum local de $J$ sur $K$ est un minimum global
2. L'ensemble des points de minimum est un ensemble convexe (éventuellement vide).
3. Si de plus $J$ est strictement convexe, alors il existe au plus un point de minimum.

**Preuve:**

1. Si $u$ est un minimum local de $J$ sur $K$, alors :
   $\exists \delta > 0, \forall w \in K, \|w-u\| < \delta \Longrightarrow J(w) \geq J(u)$.
   Soit $v \in K$, pour $\theta \in ]0,1[$ suffisamment petit, $w_\theta = \theta v + (1-\theta) u$ vérifie $\|w_\theta - u\| < \delta$ et $w_\theta \in K$ puisque $K$ est convexe. Donc, $J(w_\theta) \geq J(u)$
   D'après la convexité de $J$ on a $J(u) \leq J(w_\theta) \leq \theta J(v) + (1-\theta) J(u)$, ce qui montre bien que : $J(u) \leq J(v)$. C'est-à-dire que $u$ est un minimum global sur $K$.
2. L'ensemble de solution est un convexe : Si $u_1$ et $u_2$ sont deux minima et si $\theta \in [0,1]$, alors $w = \theta u_1 + (1-\theta) u_2$ est un minimum puisque $w \in K$ et que

$$\inf_{v \in K} J(v) \leq J(w) \leq \theta J(u_1) + (1-\theta) J(u_2) = \inf_{v \in K} J(v).$$

3. Le même raisonnement avec $\theta \in ]0,1[$ montre que, si $J$ est strictement convexe, alors nécessairement $u_1 = u_2$.

# Résultat d'existence en dimension infinie : Cas fortement convexe

**Théorème 2.**

Soit $K$ un convexe fermé non vide d'un Hilbert $V$ et $J$ une fonction $\alpha$-convexe continue sur $K$. Alors, il existe un unique minimum u de $J$ sur $K$ et on a :

$$\|v-u\|^2 \leq \frac{4}{\alpha}(J(v)-J(u)) \quad \forall v \in K.$$

En particulier, toute suite minimisante de $J$ sur l'ensemble $K$ converge vers u.

**Preuve:**

- Soit $(u^n)$ une suite minimisante de $J$ sur $K$.
  $J(v) \geq \delta$ pour tout $v \in K$,car $J$ est $\alpha$-convexe. D'où $J$ est minorée sur $K$, donc $\inf_{v\in K} J(v)$ est une valeur finie.
- Pour $n, m \in \mathbb{N}$ la propriété de la forte convexité entraîne que

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

- la suite $(u^n)$ est de Cauchy, donc 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 été déja montrée pour les fonctions strictement convexes.
- Enfin, si $v \in K, (u+v)/2 \in K$ car $K$ est convexe, d'où,

$$\frac{\alpha}{8}\|u-v\|^2 \leq \frac{J(u)}{2} + \frac{J(v)}{2} - J\left(\frac{u+v}{2}\right) \leq \frac{J(v)-J(u)}{2},$$

  car $J\left(\frac{u+v}{2}\right) \geq J(u)$

# Existence d'un minimum, cas convexe

**Remarque 3.**

On peut donner une version analogue du théorème précédent dans le cas convexe mais la preuve nécessite la notion de "convergence faible" qui est "hors-programme".

**Théorème 3.**

Soit $K$ un convexe fermé non vide d'un espace de Hilbert $V$, et $J$ une fonction convexe continue sur $K$, qui est "infinie à l'infini" dans $K$, c'est-à-dire qui vérifie la condition :

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

Alors il existe un minimum de $J$ sur $K$.

**Remarque 4.**

- Le théorème du cas convexe donne l'existence d'un minimum comme le précédent théorème , mais ne dit rien sur l'unicité ni sur l'estimation d'erreur qui est très utile pour l'étude des algorithmes numériques de minimisation puisqu'elle fournit une estimation de la vitesse de convergence d'une suite minimisante $(u^n)$ vers le point de minimum $u$.
- Le théorème reste vrai si l'on suppose simplement que $V$ est un espace de Banach réflexif (i.e. que le dual de $V'$ est $V$ ).

# Conditions d'optimalité

# Conditions d'optimalité (sans contraintes)

**Théorème 4.**

Condition nécessaire d'optimalité : Si u est minimum local de J dans $V$, alors

- Si J est différentiable en $u$ alors $J'(u) = 0$
- Si J est deux fois différentiable en $u$, on a de plus
$$\forall w \in V, J''(u)(w,w) \geqslant 0$$
.

**Preuve:**

Soit $v \in V$, on définit la fonction d'une variable réelle $\varphi(t) = J(u+tv)$ :

- Pour $t$ réel assez petit, $u+tv$ est dans un voisinage de $u$ et puisque $u$ est minimum local, $J(u+tv) \geqslant J(u)$ ce qui se traduit par $\varphi(t) \geqslant \varphi(0)$.
Par définition $\varphi$ est dérivable en 0 et $\varphi'(0) = J'(u) \cdot v$. Écrivons :
$$0 \leq \lim_{t \to 0_+} \frac{\varphi(t) - \varphi(0)}{t} = \varphi'(0) = \lim_{t \to 0_-} \frac{\varphi(t) - \varphi(0)}{t} \leq 0$$
ce qui prouve que $\varphi'(0) = 0$.
On a donc pour tout $v$, $J'(u) \cdot v = 0$, et donc $J'(u) = 0$.
- Dans le deuxième cas
  - on applique la formule de Taylor-Young à l'ordre 2 à la fonction $\varphi$ :
$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\left(\varphi''(0) + \varepsilon(t)\right), \text{ avec } \lim_{t \to 0} \varepsilon(t) = 0$$
  - Si $\varphi''(0)$ est strictement négatif alors $\frac{1}{2}\varphi''(0) + \varepsilon(t)$ l'est aussi.
  - donc $\varphi(t) < \varphi(0)$ ce qui contredit le fait que $u$ est minimum local.
  - Donc $\varphi''(0) \geqslant 0$.
  - Il est facile de calculer que c'est égal à $J''(u)(v,v)$.

# Conditions d'optimalité (sans contraintes)

**Théorème 5.**

condition suffisante : Soit $J$ une fonction différentiable dans $V$ et u un point de V tel que $J'(u) = 0$.

- Si J est deux fois differentiable dans un voisinage de u et s'il existe un voisinage $\Omega$ de u tel que $\forall v \in \Omega, \forall w \in V, J''(v)(w,w) \geqslant 0$, alors u est minimum local de $J$
- Si J est deux fois differentiable, et s'il existe $\alpha > 0$ tel que

$$\forall w \in V, J''(u)(w,w) \geqslant \alpha \|w\|^2,$$

alors u est minimum local strict pour $J$.

**Preuve:**

- Puisque $\Omega$ est un voisinage, il contient une boule ouverte de centre $u$, donc convexe.
- Soit $v$ dans cette boule. Alors tout le segment$]u, v[$ est contenu dans cette boule,
- d'après la formule de Taylor Mac-Laurin à l'ordre 2,

$$J(v) = J(u) + \underbrace{J'(u) \cdot (v-u)}_{0} + \frac{1}{2} \underbrace{J''(u + \theta(u-u))(v-u, v-u)}_{\geqslant 0}$$

- $J(v) \geqslant J(u)$ d'où $u$ est un minimum local.
- Utilisons la formule de Taylor-Young et l'hypothèse :

$$J(v) \geqslant J(u) + \alpha \|(v-u)\|^2 + \varepsilon(v-u)\|v-u\|^2, \quad \lim_{w \to 0} \varepsilon(w) = 0$$

- Pour $v$ suffisamment proche de $u$

$$\alpha + \varepsilon(v-u) > 0$$

- Donc $J(v) > J(u)$

# Condition d'optimalité (Inéquation d'Euler) : cas convexe

- Minimisation sur un ensemble de contrainte $K$ convexe

$$u \in K \text{ tel que } \quad J(u) = \inf_{v \in K} J(v) \tag{2}$$

.

- $K$ est fermé non vide
- $J$ est différentiable sur un ouvert contenant $K$.
- Conditions d'optimalité : on peut tester l'optimalité de $u$ dans la "direction admissible " $(v - u)$ car $u + t(v - u) \in K$ si $t \in [0, 1]$.

**Théorème 6.**

Inéquation d'Euler : cas convexe

- Condition nécessaire : Si u est solution optimale, on a l'inéquation d'Euler suivante :

$$\begin{cases} u \in K \\ \forall v \in K, J'(u) \cdot (v - u) \geqslant 0. \end{cases} \tag{3}$$

- Condition suffisante :Réciproquement si on a l'inéquation d'Euler en $u$ et si de plus J est convexe, alors u est solution optimale.

# Condition d'optimalité (Inéquation d'Euler) : cas convexe

**Preuve:**

1. Condition nécessaire :
   - Soit $v \in K$. Puisque $K$ est convexe, le segment $[u, v]$ est contenu dans $K$
   - $\varphi(t) = J(u + t(v - u))$ pour $t \in [0, 1]$
   - $\varphi$ est dérivable puisque $J$ est différentiable
   - On a : $\varphi(t) = \varphi(0) + t\varphi'(0) + t\varepsilon(t), \quad \lim_{t\to 0} \varepsilon(t) = 0$
   - En terme de $J$ :

     $$J\left(u + t(v - u\right) = J(u) + t\left(J'(u)(v - u) + \varepsilon(t)\right) \geqslant J(u),$$

     car $u$ est solution optimale, d'où $J'(u)(v - u) + \varepsilon(t) \geqslant 0$.
   - tendre $t$ vers 0, d'où $J'(u)(v - u) \geqslant 0$
2. Condition suffisante :
   - On suppose que $J$ est en plus convexe
   - $J$ satisfait l'inégalité d'Euler $J'(u)(v - u) \geqslant 0$
   - caractérisation de la convexité. Pour tout $v$ dans $K$,

     $$J(v) \geqslant J(u) + J'(u) \cdot (v - u) \geqslant J(u).$$

**Remarque 5.**

- Cette condition exprime que la dérivée directionnelle de $J$ au point $u$ dans toutes les directions $(v - u)$, qui sont rentrantes dans $K$, est positive.
- Il faut aussi remarquer que si $u$ est intérieur à $K$ , cette condition se réduit simplement à l'équation d'Euler $J'(u) = 0$

# Conditions d'optimalité : Cas particuliers de $K$

- $K = V$ : Sans contrainte

**Théorème 7.**

- Si J est convexe différentiable, alors u réalise le minimum de J sur $V$ si et seulement si $J'(u) = 0$

- si $J$ est $\alpha$-convexe, il existe une unique solution optimale, caractérisée par $J'(u) = 0$.

# Conditions d'optimalité : Cas particuliers de $K$

**$K$ sous-espace affine engendré par le sous-espace vectoriel fermé $E$ de $V$, i.e. $K = \{u_0 + v, v \in E\}$**

- La condition d'optimalité :

$$(3) \iff \begin{cases} u \in K \\ \forall w \in E, J'(u) \cdot w = 0 \end{cases} \tag{4}$$

- Ce qui se traduit par : $J'(u) \in E^{\perp}$.
- Cas particulier de F : Si $E$ est défini par $m$ contraintes

$$E = \{v \in V, (a_i, v) = 0, 1 \leqslant i \leqslant m\}$$

alors l'orthogonal de $E$ est l'espace vectoriel engendré par les $a_i$, et donc

$$(4) \iff \begin{cases} u \in K \\ \exists \lambda_{1,\dots,} \lambda_m \in \mathbb{R}, \quad J'(u) + \sum_{i=1}^{m} \lambda_i a_i = 0. \end{cases} \tag{5}$$

**Remarque 6.**

Si on définit les fonctions affines

$$F_i(w) = (w - u_0, a_i)$$

alors

$$K = \{w \in V, F_i(w) = 0\}$$

, (5) se réecrit

$$(5) \iff \begin{cases} u \in K \\ \exists \lambda_1, \dots, \lambda_m, \nabla J(u) + \sum_{i=1}^{m} \lambda_i F_i'(u) = 0. \end{cases} \tag{6}$$

# Conditions d'optimalité : Cas particuliers de $K$

**$K$ cône convexe fermé de sommet $u_0$**

**Définition 4.**

$K_0$ est un cone si $w \in K_0 \Longrightarrow tw \in K_0$ pour tout $t \geqslant 0$

**Proposition 2.**

- Soit $K_0$ le cône de sommet $0$
- La condition d'optimalité devient :

$$(3) \Longleftrightarrow \begin{cases} u \in K \\ J'(u) \cdot (u_0 - u) = 0 \\ \forall w \in K_0, J'(u) \cdot w \geqslant 0. \end{cases} \tag{7}$$

**Preuve:**

- Soit $v \in K, u_0 + t\,(v - u_0) \in K$
- En particulier pour $v = u$ l'inéquation d'Euler s'écrit :

$$\forall t \geqslant 0, \quad J'(u) \cdot (u_0 + t\,(u - u_0) - u) \geqslant 0,$$

$$\forall t \geqslant 0, \quad (1-t)J'(u) \cdot (u_0 - u) \geqslant 0.$$

- $1 - t$ peut prendre des valeurs positives ou négatives, ceci implique que $J'(u) \cdot (u_0 - u) = 0$
- Soit maintenant $v = u_0 + w, w \in K_0$
- On a $J'(u) \cdot (u_0 + w - u) \geqslant 0$
- Or $J'(u) \cdot (u_0 - u) = 0$
- $\forall w \in K_0, J'(u).w \geqslant 0$

# Conditions d'optimalité : Cas particuliers de $K$

**Définition 5.**

Soit $M$ un cône convexe fermé de sommet $0$, on définit le cône dual de $M$ par :

$$M^\star = \{c \in V, \forall v \in M, (c, v) \geqslant 0\}. \tag{8}$$

Si $M$ est engendré par un nombre fini de vecteurs, alors on peut décrire $M^\star$ par le Lemme de Farkas :

**Lemme 1.**

i $M = \{v \in V, \forall i \in \{1, \dots, m\}, (v, a_i) \leqslant 0\}$, alors $c \in M^\star$ si et seulement si $-c$ appartient au cone convexe engendré par les $a_i$, i.e. il existe $\{\lambda_1, \dots, \lambda_m\}$ tous $\geqslant 0$ tels que $c = -\sum_{i=1}^{m} \lambda_i a_i$

**Preuve:**

- Si $c = -\sum_{i=1}^{m} \lambda_i a_i$ avec tous les $\lambda_i$ positifs ou nuls, alors pour tout $v$ dans $M$,

$$(v, c) = -\sum_{i=1}^{m} \lambda_i (a_i, v) \geqslant 0$$

  donc $c \in M^*$
- La réciproque repose sur le théorème de Hahn-Banach qui est hors programme.

# Conditions d'optimalité : Cas particuliers de $K$

**$K_0$ cône convexe fermé engendré par un nombre fini de vecteurs**

- Intéressons nous maintenant au cas où $K_0$ est défini par $m$ contraintes,

$$K_0 = \{w \in V, (a_i, w) \leq 0, 1 \leqslant i \leqslant m\}$$

.

- La troisième ligne dans (7) exprime que $-J'(u)$ est dans $K_0^\star$
- L'équation (7) se réécrit

$$(3) \Longleftrightarrow \begin{cases} u \in K \\ J'(u) \cdot (u_0 - u) = 0 \\ \exists (\lambda_1, \cdots, \lambda_m) \geqslant 0, \nabla J(u) + \sum_{i=1}^{m} \lambda_i a_i = 0 \end{cases} \tag{9}$$

**Remarque 7.**

Comme dans le cas précédent, $K$ se définit ici par

$$K = \{w \in V, \;\; F_i(w) \leq 0, \;\; 1 \leq i \leq m\}$$

et (9) s'écrit

$$(3) \Longleftrightarrow \begin{cases} u \in K \\ J'(u) \cdot (u_0 - u) = 0 \\ \exists (\lambda_1, \cdots, \lambda_m) \geqslant 0, \nabla J(u) + \sum_{i=1}^{m} \lambda_i F_i'(u) = 0 \end{cases} \tag{10}$$

# Récapitulatif :

**Résumé:**

Caractérisation des fonctions convexes dans le cas régulier :

1. Si $J : V \to \mathbb{R}$ est différentiable, on a les équivalences entre : (i) $J$ est convexe. (ii) $J(y) \geq J(x) + \langle \nabla J(x), y - x\rangle, \forall (x, y) \in V^2$. (iii) $\langle \nabla J(y) - \nabla J(x), y - x\rangle \geq 0, \forall (x, y) \in V^2$.
2. On a équivalence entre convexité stricte et les inégalités (ii) et (iii) précédentes rendues strictes, pour $x \neq y$.
3. Si $J : V \longrightarrow \mathbb{R}$ est deux fois différentiable, on a les équivalences entre (i) $J$ est convexe, (ii) pour tout $h \in V, \langle J"(x_0)h, h\rangle \geq 0$.

Caractérisation des fonctions $\alpha - convexes$ dans le cas régulier :

1. Si $J : V \longrightarrow \mathbb{R}$ est différentiable, on a les équivalences : (i) $J$ est $\alpha$-elliptique ($\alpha - convexe$) ; (ii) $J(y) \geq J(x) + \langle \nabla J(x), y - x\rangle + \frac{\alpha}{2}\|y - x\|^2, \forall (x, y) \in V^2$ ; (iii) $\langle \nabla J(y) - \nabla J(x), y - x\rangle \geq \alpha\|y - x\|^2, \forall (x, y) \in V^2$.
2. Si $J : V \longrightarrow \mathbb{R}$ est deux fois différentiable, on a les équivalences/ (i) $J$ est $\alpha$-elliptique ; (ii) $\langle \text{Hess}\, J(x)h, h\rangle \geq \alpha\|h\|^2, \forall x \in V, \forall h \in V$.
3. (lemme) Soit $J$, une fonction $\alpha$-convexe sur $K$. Alors, il existe deux constantes $(\alpha_1, \alpha_2) \in \mathbb{R}_+ \times \mathbb{R}$ telles que $J(x) \geq \alpha_1\|x\|^2 + \alpha_2$

# Récapitulatif :

**Résumé:**

Existence et unicité :

1. Cas de dimension finie : $J : \Omega$**(ouvert)** $\subset K$**(fermé)** $\subset \mathbb{R}^N \to \mathbb{R}$ une fonctionnelle continue, alors :
   - $K$ est compact $\Rightarrow J$ a au moins un minimum sur $K$.
   - $J$ coercive $\Rightarrow J$ a au moins un minimum sur $K$. Pour assurer l'unicité dans les deux cas, il faut que $J$ soit strictement convexe.
2. Cas de dimension infinie : $J$ une fonction $\alpha$-convexe continue sur $K$ convexe fermé non vide d'un Hilbert. Alors, il existe un unique minimum $u$ de $J$ sur $K$ et on a :

$$\|v-u\|^2 \leq \frac{4}{\alpha}(J(v)-J(u)) \quad \forall v \in K.$$

Conditions d'optimalité :

1. Cas sans contraintes $K = V$ : Le problème à résoudre : $\begin{cases} \min J(x) \\ x \in V \end{cases}$
   - Condition nécessaire (CN) : $u$ min local $\Longrightarrow J'(u) = 0$.
   - Condition suffisante (CS) pour un minimum local (resp.strict) :
     $\begin{cases} J'(u) = 0 \\ J''(u) \geq 0 \quad (resp. J''(u)(w,w) \geq \alpha \|w\|^2) \end{cases} \Longrightarrow$ $u$ min local (resp.strict).

   Cas avec contrainte (K convexe) :Inéquation d'Euler
   - Condition nécessaire : Si u est solution optimale, on a l'inéquation d'Euler suivante :
     $\begin{cases} u \in K \\ \forall v \in K, J'(u)\cdot(v-u) \geqslant 0. \end{cases}$
   - Condition suffisante :Réciproquement si on a l'inéquation d'Euler en $u$ et si de plus J est convexe, alors u est solution optimale.

# Multiplicateurs de Lagrange, cas général

## Le cone de direction admissible

- Le lemme de Farkas va nous permettre de trouver des conditions nécessaires d'optimalité dans le cas général.

- Définition du cone de direction admissible :

**Définition**
Soit $K$ un fermé non vide, pour tout $v$ dans $K$, on définit le cone des directions admissibles par

$$K(v) = \{0\} \cup \{w \in V,$$

$$\exists \{v_k\}_{k\in\mathbb{N}} \subset K \lim_{k\to+\infty} v_k = v, v_k \neq v \text{ pour tout } k, \lim_{k\to+\infty} \frac{v_k - v}{\|v_k - v\|} = \frac{w}{\|w\|} \quad (11)$$

**Théorème**
*Pour tout v dans $K$, $K(v)$ est un cone fermé de sommet $0$.*

**Preuve**

- vérifier que $0 \in K(v)$ : Prendre la suite constante $v^n = v$
- Verifier que l'ensemble $K(v)$ est un cône, c'est-à-dire que $\lambda w \in K(v)$ pour tout $w \in K(v)$ et tout $\lambda \geq 0$.
- On montre que $K(v) = V$ si $v$ est dans l'intérieur de $K$.

# Inéquation d'Euler,cas général

- Le cone des directions admissibles va permettre de donner une condition nécessaire d'optimalité.

**Théorème (Inéquation d'Euler, Cas général)**
*Si $J$ a un minimum local en $u \in K$ et si $J$ est différentiable en u, alors $J'(u) \in K(u)^\star$*

**Preuve**

- Soit $w \in K(u)$, et $u_k$ une suite telle que

$$\lim_{k \to +\infty} u_k = u, \text{ et } \lim_{k \to +\infty} \frac{u_k - u}{\|u_k - u\|} = \frac{w}{\|w\|}.$$

- Puisque $u$ est un point de minimum local, pour $k$ suffisamment grand, $J(u_k) \geqslant J(u)$

- Formule de Taylor-Young à l'ordre 1 pour $u_k$ et $u$.

$$J(u_k) - J(u) = J'(u) \cdot (u_k - u) + \epsilon(u_k - u)\|u_k - u\|, \quad \lim_{w \to 0} \epsilon(w) = 0$$

- On divise par $\|u_k - u\|$ :

$$0 \leq \frac{J(u_k) - J(u)}{\|u_k - u\|} = J'(u) \cdot \frac{u_k - u}{\|u_k - u\|} + \epsilon(u_k - u).$$

- Passage à la limite sur $k$ :

$$0 \leq J'(u) \cdot \frac{w}{\|w\|}.$$

- Ceci prouve que $J'(u) \cdot w \geqslant 0$ pour tout $w \in K(u)$, c.a.d $J'(u) \in K(u)^*$.

## Cas particulier de $K(u)$

- 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

$$K(u) = \{w = v - u \;\; avec \;\; v \in K\}$$

  . 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 $K = \{v \in V$ tel que $F(v) = 0\}$.
  Si $F'(u) \neq 0$, alors $K(u) = [F'(u)]^{\perp}$ (hyperplan tangent en $u$ à la surface $K$ ).

# Contraintes égalités

- On suppose que l'ensemble $K$ est donné par la forme suivante :

$$K = \{v \in V, F_1(v) = F_2(v) = \cdots = F_m(v) = 0\} \qquad (12)$$

- Les fonctions $F_1, .., F_m$ sont $\mathcal{C}^1 : V \to \mathbb{R}$.

- $F$ l'application de $V$ dans $\mathbb{R}^m$ définie par les $F_i$

- soit $F(v) = (F_1(v), F_2(v), \ldots, F_m(v)) \in \mathbb{R}^m$

**Définition**
Les contraintes sont régulières en $u \in K$ si les $F_i'(u)$ sont linéairement indépendantes.
On dit alors que u est un point régulier.

## caractérisation du cone des directions admissibles

**Lemme**
Si les contraintes sont regulières en $u \in K$, alors

$$K(u) = \{w \in V, F_i'(u) \cdot w = 0, 1 \leq i \leq m\} \tag{13}$$

Ici en fait le cône est un espace vectoriel : l'orthogonal de l'espace engendré par les $F_i'(u)$.

**Preuve**

- On note par :

$$\tilde{K}(u) = \{w \in V, F_i'(u), w = 0, 1 \leq i \leq m\}$$

.

- On montre que$K(u) \subset \tilde{K}(u)$
  - Soit $w \in K(u), u_k$ la suite associée
  - On applique la formule de Taylor-Lagrange à l'ordre 1 à chaque $F_i$ :

$$F_i\left(u_k\right) = F_i(u) + F_i'\left(u + \theta\left(u_k - u\right)\right) \cdot \left(u_k - u\right)$$

  - $u$ et $u_k$ sont dans $K, F_i(u) = F_i\left(u_k\right) = 0$
  - On a alors

$$F_i'\left(u + \theta\left(u_k - u\right)\right) \cdot \left(u_k - u\right) = 0 = F_i'\left(u + \theta\left(u_k - u\right)\right) \cdot \frac{u_k - u}{\left\|u_k - u\right\|}.$$

  - Il suffit maintenant de passer à la limite en $k$ pour obtenir $F_i'(u) \cdot w = 0$ et donc $w \in \tilde{K}(u)$.

# Les multiplicateurs de Lagrange

**Théorème**
*Si $u \in K$, u régulier, est minimum local pour $J$, il existe m réels $p_1, \dots, p_m$ tels que*

$$J'(u) + \sum_{i=1}^{m} p_i F_i'(u) = 0. \tag{14}$$

**Preuve**

- $J'(u) \in K(u)^*$ c.a.d pour tout $w$ dans $K(u)$, $\quad J'(u), w \geqslant 0$
- $K(u)$ est un espace vectoriel : si $w \in K(u), -w \in K(u)$.
- Donc pour tout $w$ dans $K(u)$, $\quad J'(u), w = 0$ :
- $J'(u)$ est dans l'othogonal de $K(u)$ qui est lui meme l'orthogonal de vect $(F_1'(u), \dots, F_m'(u))$

**Remarque**
Si K et J sont convexes, alors c'est une condition nécessaire et suffusante.

**Théorème**
*Supposons que J est convexe différentiable et les fonctions* $F_i$ *sont convexes. Alors* $u \in K, u$ *régulier, est minimum local pour* $J$, *si et seulement si il existe m réels* $p_1, \ldots, p_m$ *tels que*

$$J'(u) + \sum_{i=1}^{m} p_i F_i'(u) = 0. \tag{15}$$

- il faut montrer que : si $J'(u) \in \text{vec}\,(F_1'(u), \ldots, F_m'(u))$ alors $u$ est minimum local.

- Puisque les $F_i$ sont convexes, $K$ est convexe : si $v$ et $w$ sont dans $K$ cela signifie que pour tout $i, F_i(v) = F_i(w) = 0$.
- Alors

$$F_i(\theta v + (1-\theta)w) = \theta F_i(v) + (1_8)\, F_i(w) = 0,$$

- Soit $v \in K$, alors $w = v - u \in K(u)$.
D'après le lemme $w$ est orthogonal aux vecteurs $F_i'(u)$, et donc à $J'(u)$.
On a donc,

$$\forall v \in K, J'(u) \cdot (v-u) = 0,$$

Ce qui est plus fort que l'inéquation d'Euler dans le théorème précédent.
D'où le résultat d'après la caractérisation de la convexité de $J$

# Le Lagrangien

**Définition**
On définit le lagrangien sur $V \times \mathbb{R}^m$ à valeurs dans $\mathbb{R}$ par

$$\mathcal{L}(v,q) := J(v) + \sum_{i=1}^{m} q_i F_i(v) \tag{16}$$

On a alors

$$\begin{aligned} D_v\mathcal{L}(v,q) &= J'(v) + \sum_{i=1}^{m} q_i F_i'(v) \\ D_q\mathcal{L}(v,q) &= F(v) \end{aligned} \tag{17}$$

et

$$\begin{aligned} u \in K \quad &\Longleftrightarrow \forall q \in \mathbb{R}^m, D_v\mathcal{L}(u,q) = 0 \\ u \text{ minimum local} \quad &\Longleftrightarrow \exists p \in \mathbb{R}^m, D_q\mathcal{L}(u,p) = 0. \end{aligned} \tag{18}$$

$p$ s'appelle la variable duale de $u$.

## Contraintes inégalitées

Soit $F(v) = (F_1(v), \dots, F_M(v)) : V \to \mathbb{R}^M$. On considère

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

avec

$$K = \{v \in V, F(v) \leq 0\}$$

où

$F(v) \leq 0$ signifie que $F_i(v) \leq 0$ pour $1 \leq i \leq m$.
On suppose que $F$ est une fonction $C^1$ de $V$ dans $\mathbb{R}^m$.s

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

- On appelle ensemble des contraintes actives en $u$
  $I(u) = \{i \in \{1, \dots, m\}, F_i(u) = 0\}$.
- Les contraintes sont dites qualifiées en u si

**Remarque**
Si toutes les contraintes sont affines, alors $\bar{w} = 0$ convient : les contraintes sont qualifiées en tout point.

On peut encore caractériser le cône des directions admissibles.

**Lemme**
Si les contraintes sont qualifiées en $u \in K$, alors

$$K(u) = \left\{w \in V, \ \ \forall i \in I(u), F_i'(u), w \leq 0\right\} \tag{20}$$

**Preuve**
On note par

$$\tilde{K}(u) = \{w \in V, \forall i \in I(u), F_i'(u).w \leq 0\}$$

.

- Montrons que $K(u) \subset \tilde{K}(u)$

  Soit $w \in K(u), u_n$ une suite d'éléments de $K$ convergeant vers $u$

  Soit $i \in I(u)$. Alors $F_i(u) = 0$ et $F_i\left(u_n\right) \leq 0$.

  D'après la formule de Taylor -Young à l'ordre 1 pour $F_i$ au voisinage de $u$ :

$$F_i\left(u_n\right) = \underbrace{F_i(u)}_{0} + F_i'(u) \cdot \left(u_n - u\right) + \varepsilon\left(u_n - u\right)\left\|u_n - u\right\| \leq 0$$

On divise maintenant par $\|u_n - u\|$ :

$$F_i'(u) \cdot \frac{u_n - u}{\|u_n - u\|} + \varepsilon(u_n - u) \leq 0$$

on passe à la limite, ce qui donne $F_i'(u) \cdot w \leq 0$. Donc $w \in \tilde{K}(u)$.

- Montrons que $\tilde{K}(u) \subset K(u)$

  Soit $w \in \tilde{K}(u)$.

  Nous allons montrer que pour tout $\delta \geqslant 0, w + \delta \bar{w} \in K(u)$.
  *) Si $w + \delta \bar{w} = 0$, puisque $0 \in K(u)$, c'est vrai.

*) Si $w + \delta \bar{w} \neq 0$,

pour $\varepsilon_n$ une suite décroissante tendant vers 0 , posons

$$u_n = u + \varepsilon_n (w + \delta \bar{w}).$$

Il est clair que $u_n \to u$.

Montrons que $u_n \in K$.

- Soit $i \in I(u)$, donc $F_i(u) = 0$.
  - Si $F_i$ est affine, alors la formule de Taylor à l'ordre 1 est exacte

$$F_i\,(u_n) = F_i(u) + \varepsilon_n F_i'(u)(w + \delta \bar{w}) = \varepsilon_n F_i'(u)(w + \delta \bar{w}) \leq 0.$$

- Si $F_i$ n'est pas affine,

$$F_i\left(u_n\right) = F_i(u) + \varepsilon_n\left(F_i'(u)(w + \delta\bar{w}) + \varepsilon\left(\varepsilon_n\right)\|w + \delta\bar{w}\|\right)$$

  Ici $F_i'(u)(w + \delta\bar{w}) < 0$,
  et donc pour $n$ assez grand le terme en facteur de $\varepsilon_n$ est $\leq 0$.

  Donc pour $i \in I(u), F_i\left(u_n\right) \leq 0$ pour $n$ suffisamment grand.

- Soit $i \notin I(u)$,

  alors $F_i(u) < 0$.

  Par la formule de Taylor ci-dessus, $F_i\left(u_n\right) \leq 0$ pour $n$ suffisamment grand.

On a donc montré que $u_n \in K$ pour $n$ assez grand.

Maintenant

$$\frac{u_n - u}{\|u_n - u\|} = \frac{w + \delta \bar{w}}{\|w + \delta \bar{w}\|}$$

Tous ces éléments montrent que $w + \delta \bar{w} \in K(u)$.

En faisant tendre $\delta$ vers 0 , puisque $K(u)$ est fermé, On obtient que $w \in K(u)$.

Le lemme de Farkas permet alors d'établir l'existence de multiplicateurs de Lagrange :

**Théorème**
*Si $u \in K$, où les contraintes sont qualifiées, est minimum local pour J, il existe m réels $p_1, \ldots, p_m \geqslant 0$ tels que*

$$\begin{aligned} J'(u) + \sum_{i=1}^{m} p_i F_i'(u) = 0 \\ \sum_{i=1}^{m} p_i F_i(u) = 0 \textit{ ou encore } F_i(u) < 0 \Longrightarrow p_i = 0. \end{aligned} \tag{21}$$

**Preuve**
D'après le théorème précédent, pour tout $v$ dans $K, J(u) \in K(u)^*$.

D'après le lemme de Farkas,

$$J'(u) = -\sum_{i \in I(u)} \lambda_i F_i'(u), \quad \lambda_i \geqslant 0.$$

ce qui est équivalent à (21).

**Remarque**
Le lagrangien est maintenant défini sur $V \times \mathbb{R}_+^m$, et l'on peut écrire

$$\begin{gathered} u \in K \text{ solution optimale } \Rightarrow \exists p \in \mathbb{R}_+^m \\ \mathcal{L}_v'(u,p) = \mathcal{L}_q'(u,p) \cdot p = 0. \end{gathered} \tag{22}$$

Attention, contrairement au cas des contraintes égalités, on n'a qu'une condition nécessaire.
Le développement d'une condition nécessaire et suffisante est l'objet du chapitre suivant.