Analyse et Optimisation – L3

Jérôme Bolte
Université Toulouse Capitole / TSE,

9 avril 2014

Chapitre 1

Rappels et compléments sur les espaces de Banach

1.1 Rappels sur les espaces vectoriels normés

1.1.1 Espaces de Banach

Un espace vectoriel réel normé E est un espace vectoriel muni d'une norme $\|\cdot\|$.

Une suite de Cauchy $(x_k)_{k\in\mathbb{N}}$ de E est une suite telle que

$$(\forall\epsilon > 0)\ (\exists N \in \mathbb{N})\ (\forall p, q \geq N)\ (\|x_p - x_q\| \leq \epsilon).$$

Intuitivement cette définition vise à saisir la notion de suite convergente sans faire référence à la limite de la suite. Rappelons en effet qu'une suite $(x_k)_{k\in\mathbb{N}}$ de E est dite convergente si

$$(\exists x \in E)\ (\forall\epsilon > 0)\ (\exists N \in \mathbb{N})\ (\forall p \geq N)\ (\|x_p - x\| \leq \epsilon),$$

et remarquons que dans la définition de suite de Cauchy on fait disparaitre le $\exists x$ ci-dessus pour "l'approcher" par un x_q avec $q \geq N$.

Toutefois, comme le montrerons les considérations à venir, cette intuition a ses limites.

Remarque 1. (a) Les suites convergentes d'un espace vectoriel normé sont de Cauchy, mais la réciproque est fausse en général.
(b) Toute suite de Cauchy est bornée : il existe $M > 0$ tel que $\|x_k\| \leq M$ pour tout k dans $\mathbb{N}$.
(c) Soit $N_E(\cdot)$ une norme sur E équivalente à $\|\cdot\|$. Rappelons que cela signifie qu'il existe des constantes $c_1, c_2 > 0$ telles que

$$c_1 N_E(x) \leq \|x\| \leq c_2 N_E(x),\ \forall x \in E.$$

Alors $(E, \|\cdot\|)$ et $(E, N_E(\cdot))$ ont les même suites de Cauchy ([1]). Autrement dit une suite de Cauchy de $(E, \|\cdot\|)$ est une suite de Cauchy dans $(E, N_E(\cdot))$ et vice-versa.

1. Et, évidemment, les mêmes suites convergentes.

Définition 1 (Espaces de Banach / Espaces complets)**.** *Un espace de Banach est un espace vectoriel normé dans lequel toute suite de Cauchy converge vers un élément de E.*([2])

Exemples d'espaces de Banach :

1. Les espaces vectoriels de dimension finie sont des espaces de Banach.
2. $E = C^0([0,1];\mathbb{R})$ muni de la norme suprémale $\|f\|_\infty = \sup\{|f(x)| : x \in [0,1]\}$ est un espace de Banach.
3. Les espaces $\Big(L^p(0,1;\mathbb{R}), \|\cdot\|_p\Big)$ sont des espaces de Banach (voir cours d'intégration et de théorie de la mesure).
4. Un contre-exemple important : $C^0([0,1];\mathbb{R})$ muni de la norme L^1, à savoir $\|f\|_1 = \int_0^1 |f(s)|ds$ n'est pas complet (suite des "manivelles").

Théorème 1 (Convergence normale des séries dans les Banach)**.** *Soit* $(E, \|\cdot\|)$ *un espace de Banach et* $(x_k)_{k\in\mathbb{N}}$ *une suite dans* E.

Si $\sum_{k=0}^{+\infty} \|x_k\| < +\infty$ *alors la suite des sommes partielles* $S_n = \sum_{k=0}^{n} x_k$ *converge lorsque* n *tend vers l'infini vers un élément de* E *noté* $\sum_{k=0}^{\infty} x_k$*. On a de plus*

$$\|\sum_{k=0}^{\infty} x_k\| \leq \sum_{k=0}^{\infty} \|x_k\|.$$

On dit que la série $\sum x_k$ *converge normalement.*

1.1.2 Espaces produits

Étant donnés un nombre fini d'espaces vectoriels normés $(E_1, \|\cdot\|_{E_1}, \ldots, E_n, \|\cdot\|_{E_n})$, leur produit cartésien $E_1 \times \ldots \times E_n$ est systématiquement muni de l'une des trois normes suivantes :

$$\|X\|_1 = \sum_{i=1}^{n} \|x_i\|_{E_i}$$
$$\|X\|_2 = \sqrt{\sum_{i=1}^{n} \|x_i\|_{E_i}^2}$$
$$\|X\|_\infty = \max_{i=1,\ldots,n} \|x_i\|_{E_i}$$

où $X = (x_1, \ldots, x_n)$ est l'élément courant du produit $E_1 \times \ldots \times E_n$. Ces normes sont équivalentes (savoir calculer les constantes).

On notera et retiendra qu'une suite $\Big((x_1^k, \ldots, x_n^k)\Big)_{k\in\mathbb{N}}$ de $E_1 \times \ldots \times E_n$ converge (resp. est de Cauchy) si et seulement si chacune des suites $(x_1^k)_{k\in\mathbb{N}}, \ldots, (x_n^k)_{k\in\mathbb{N}}$ converge dans son espace respectif (resp. est de Cauchy dans son espace respectif).

Théorème 2. *L'espace vectoriel* $E_1 \times \ldots \times E_n$ *(muni de l'une des 3 normes ci-dessus) est un est espace de Banach si et seulement si chaque* E_i *est un espace Banach.*

2. On dit aussi que E est un espace vectoriel normé complet.

1.1.3 Applications linéaires, applications

Applications linéaires

Révision du cours d'algèbre linéaire élémentaire.

Théorème 3. *Soient E, F des espaces vectoriels normés et $u : E \to F$ une application linéaire.*

Les assertions suivantes sont équivalentes

1. *u est continue*
2. *u est continue en* 0
3. *$\exists M > 0$ telle que $\|u(x)\|_F \leq M\|x\|_E$ pour tout x de E.*

L'espace vectoriel des applications linéaires continues est désigné par $\mathcal{L}(E, F)$.

Deux exemples simples : $E = C^0([0,1], \mathbb{R}), \|\cdot\|_\infty$. L'application $u(f) = \int_0^1 f$ est linéaire et continue. De même, étant donné a dans $[0,1]$, le "Dirac" $\delta_a(f) = f(a)$ est linéaire et continue sur E.

Il faut absolument connaitre et savoir user du résultat fondamental suivant (lié à l'extrême rigidité des applications linéaires).

Théorème 4 (Continuité automatique)**.** *Soit $u : E \to F$ une application linéaire. Si E est de dimension finie alors u est continue.*

L'espaces $\mathcal{L}(E, F)$ a une structure naturelle d'espace vectoriel normé via la *norme d'opérateur*

$$\begin{aligned} \|u\|_{\mathcal{L}(E,F)} &= \sup\left\{\frac{\|u(x)\|_F}{\|x\|_E} : x \in E \setminus \{0\}\right\} \\ &= \sup\{\|u(x)\|_F : x \in E,\ \|x\|_E = 1\} \\ &= \sup\{\|u(x)\|_F : x \in E,\ \|x\|_E \leq 1\} \end{aligned}$$

Théorème 5. *$\mathcal{L}(E, F)$ est un espace de Banach dès que F est un espace de Banach.*

Proposition 1. *Étant donnés E, F, G des espaces vectoriels normés, u dans $\mathcal{L}(E, F)$, v dans $\mathcal{L}(F, G)$, on montre que*

$$v \circ u \in \mathcal{L}(E, G) \text{ et que } \|v \circ u\|_{\mathcal{L}(E,G)} \leq \|u\|_{\mathcal{L}(E,F)} \|v\|_{\mathcal{L}(F,G)}$$

Aspects non-linéaires

Réviser la notion de continuité des fonctions dans les espaces vectoriels normés.

Définition 2 (Applications lipschitziennes et contractantes)**.** *Soient E, F des espaces vectoriels normés et $f : D \subset E \to F$.*

1. *f est dite lipschitzienne s'il existe $L \geq 0$ tel que*

$$\|f(x) - f(y)\|_F \leq L\|x - y\|_E,$$

pour tout x, y dans D.

2. *f est (fermement) contractante si elle est L lipschitzienne avec $0 \leq L < 1$.*

1.2 Théorème du point fixe de Banach

Théorème 6 (Théorème du point fixe de Banach-Picard)**.** *Soit A un fermé* **non vide** *d'un espace de Banach $(E, \|\cdot\|)$ et $f : A \to E$ une contraction ferme telle que $f(A) \subset A$.*

Alors f a un unique point fixe dans A, ie il existe $\bar{x}$ dans A tel que $f(\bar{x}) = \bar{x}$.

Remarque 2. (a) Un des intérêts majeurs du théorème ci-dessus est de permettre la résolution de certaines équations du type

$$f(x) = 0 \text{ ou } g(x) = x \ \ (\text{avec } f, g : E \to E), \tag{1.1}$$

dans un cadre très général.

On notera que l'on passe formellement d'une équation à l'autre sans difficulté : par exemple résoudre $f(x) = 0$ revient à trouver x tel que $x + f(x) = x$, i.e. à trouver un point fixe de $g := \mathrm{Id}_E + f$.

(b) Comprendre que le Théorème de Picard fournit existence et unicité d'une solution avec en sus *une méthode "effective de calcul/d'approximation"*. Cette méthode s'appelle la méthode des *"itérés successifs" :*

$$x_{k+1} = f(x_k), \quad x_0 \in A.$$

(c) L'application linéaire $f(x_1, x_2) = \frac{2}{\sqrt{5}}(\frac{1}{2}(x_1 + x_2), x_2)$ est contractante pour la norme sup alors que $\|f(0,1) - f(0,0)\|_2 = \|f(0,1)\|_2 = 1$. Elle n'est donc pas contractante pour la norme Euclidienne. Une application peut donc être contractante pour une norme mais pas pour une autre. Ainsi en vue de la résolution d'une équation de type (1.1) le choix d'une norme adéquate a son importance.

Approximations rationnelles de $\sqrt{2}$. Par *définition* $\sqrt{2}$ est la solution positive de $x^2 - 2 = 0$. Cette équation peut se voir comme $x^2 - 1 = 1$ et donc $(x - 1)(x + 1) = 1$. Ainsi $\sqrt{2}$ est donc l'unique solution de

$$x = \frac{1}{1 + x} + 1, \quad x \geq 0.$$

On définit $f(x) = \frac{1}{x+1} + 1$ pour $x \geq 0$ et on montre que le théorème de Picard s'applique ([3]). Par la méthode des itérés successifs on obtient une infinité d'approximations rationnelles (fractions continues) de $\sqrt{2}$:

$$x_0 = a,\ x_1 = 1 + \frac{1}{1 + a},\ x_2 = 1 + \frac{1}{2 + \frac{1}{2+a}},\ x_3 = 1 + \frac{1}{2 + \frac{1}{2+\frac{1}{2+a}}}, \ldots$$

3. Prudence..

de sorte que

$$\sqrt{2} = 1 + \cfrac{1}{2 + \frac{1}{2+\frac{1}{2+\cdots}}}.$$

En passant, on devine aussi le “défaut de de complétude" de $\mathbb{Q}$ sur cet exercice.

1.3 Isomorphismes linéaires

Dans un espace vectoriel normé E, on désigne par $B_E(x,r)$ la boule ouverte de centre $x \in E$ et de rayon $r \geq 0$.

Théorème 7 (Application ouverte)**.** *On suppose que E, F sont des espaces de Banach. Soit $u \in \mathcal{L}(E,F)$ une application surjective. Alors il existe $r > 0$ tel que*

$$u(B_E(0,1)) \supset B_F(0,r).$$

La démonstration repose sur le théorème de Baire et donc on ne peut se passer de l'hypothèse de complétude ("espaces de Banach").

Un isomorphisme linéaire de E dans F est une application linéaire continue bijective de E dans F. Isom(E,F) désigne l'ensemble de ces isomorphismes. On a :

1. Si u est dans Isom(E,F) son inverse u^{-1} est dans Isom(F,E).
2. Si $u \in \mathrm{Isom}(E,F)$ et $v \in \mathrm{Isom}(F,G)$, alors $v \circ u \in \mathrm{Isom}(E,G)$ et

$$(v \circ u)^{-1} = u^{-1} \circ v^{-1}.$$

3. Le cas de la dimension finie revient à l'étude du groupe des matrices inversibles.

Théorème 8. *Soit E espace de Banach. Alors Isom(E) contient la boule ouverte de centre Id_E et de rayon 1, autrement dit :*

$$\mathrm{Id}_E - v \text{ est inversible dès que } \|v\|_{\mathcal{L}(E,F)} < 1.$$

Théorème 9. *Soient E, F des espaces de Banach.*

1. $\mathrm{Isom}(E,F)$ *est un ouvert de* $\mathcal{L}(E,F)$.
2. *L'application de “prise d'inverse"* $\mathrm{Isom}(E,F) \ni u \to u^{-1}$ *est continue de* $\mathcal{L}(E,F)$ *vers* $\mathcal{L}(F,E)$ *munis de leurs normes d'opérateur.*

Example 1. Soient n, m des entiers naturels, on considère l'ensemble Isom$(\mathbb{R}^n, \mathbb{R}^m)$. Si $n = m$, Isom$(\mathbb{R}^n, \mathbb{R}^n)$ s'identifie à l'ensemble des matrices inversibles de taille n. Lorsque $m \neq n$, on a bien entendu Isom$(\mathbb{R}^n, \mathbb{R}^m) = \emptyset$.

Chapitre 2

Applications différentiables

2.1 Définition et premières propriétés

Notations de Landau. Soit Ω un ouvert de E contenant 0, un "petit o" de $h \in \Omega$ (à valeurs dans F), désigne *n'importe quelle* fonction de la forme

$$U \ni h \to \|h\|\epsilon(h)$$

où U est un voisinage de 0 dans Ω, $\epsilon : U \to F$ satisfait $\epsilon(0) = 0$ et est continue en 0. Ainsi on a les règle suivantes

$$\begin{aligned}
&o(h) + o(h) = o(h),\\
&\lambda o(h) = o(h) \text{ pour tout } \lambda \text{ réel},\\
&\|o(h)\| \leq o(\|h\|).
\end{aligned}$$

On notera aussi qu'étant donnée une fonction $\psi : U \to F$ définie sur un voisinage U de 0, on a $\Big(\|\psi(h)\| \leq o(\|h\|) \Rightarrow \psi(h) = o(h)\Big)$.

Définition 3. *E, F sont des espaces vectoriels normés. Soit Ω un ouvert de E et $\bar{x}$ un élément de Ω.*

1. *On dit que f est différentiable en $\bar{x}$, si :*
   (a) il existe une application linéaire continue $L : E \to F$
   (b) il existe $r > 0$, une fonction $\epsilon : B(0_E, r) \to F$ continue et nulle en $\bar{x}$
   tels que $B(\bar{x}, r) \subset \Omega$ et
   $$f(\bar{x} + h) = f(\bar{x}) + L(h) + \|h\|\epsilon(h), \quad \forall h \in B(0_E, r),$$
   ou en d'autres termes
   $$f(\bar{x} + h) = f(\bar{x}) + L(h) + o(h), \text{ pour tout } h \text{ dans } \Omega.$$
   Lorsque f est différentiable en $\bar{x}$ on écrit $f \in \Delta^1_{\bar{x}}$.
2. *L'application L est unique elle est notée $f'(\bar{x})$; ainsi $f'(\bar{x}) \in \mathcal{L}(E, F)$*
   $$f(\bar{x} + h) = f(\bar{x}) + f'(\bar{x})(h) + o(h), \forall h \in \Omega.$$

3. *On dit que f est différentiable si f est différentiable en tout point de Ω et on écrit $f \in \Delta^1$. Cela définit*
$$f' : \Omega \to \mathcal{L}(E, F).$$
4. *On dit que f est continûment différentiable en $\bar{x}$ ($f \in C^1_{\bar{x}}$), s'il existe un voisinage U de $\bar{x}$ sur lequel f est différentiable et si l'application $f' : U \to \mathcal{L}(E, F)$ est continue en $\bar{x}$.*
Soit U un ouvert de Ω, on dit que f est $\Delta^1(U)$ (resp. $C^1(U)$) si f est Δ^1 (resp. C^1) en tout point de Ω.

Commentaires.
1. Soient $a < b$ et $f :]a, b[\to \mathbb{R}$. Les notions de dérivabilité est de différentiabilité sont identiques via l'identification usuelle $\mathbb{R} \simeq \mathcal{L}(\mathbb{R}, \mathbb{R})$. Ainsi les deux mots sont utilisés indistinctement dans ce cours.
2. Lorsque f est différentiable en $\bar{x}$, f est continue en $\bar{x}$.
3. La différentiabilité (resp. la différentielle) dépend des normes initiales. Par contre en dimension finie, en raison de l'équivalence des normes, la différentiabilité et la différentielle $f'(\bar{x})$ ne dépendent pas du choix des normes.

Proposition 2. *Soient E, F des espaces vectoriels normés et $f : E \to F$ une application affine continue, ie une application de la forme $f(x) = u(x) + b$ où $u \in \mathcal{L}(E, F)$ et $b \in F$. Alors f est différentiable et*
$$f'(x) = u, \quad \forall x \in E.$$
Exemple. Si $f(x) = \frac{1}{2}\|x\|^2$ pour x dans $(H, \langle \cdot, \cdot \rangle)$ espace de Hilbert, alors f est différentiable et $f'(x)(h) = \langle x, h \rangle$ pour tout x dans E, h dans E.

Pour établir la différentiabilité le protocole qui suit est à absolument maitriser :
1. *On écrit $f(x + h) = ...$ et on essaye de conclure par $... = f(x) + L(h) + R(h)$ où :*
2. *on montre que L est linéaire (c'est souvent évident) et continue (bien distinguer DF et DI),*
3. *on montre que $R(h) = o(h)$,*
4. *et l'on conclut en utilisant la définition.*

Proposition 3 (Applications bilinéaires)**.** *Soient E_1, E_2, F des espaces vectoriels normés et $B : E_1 \times E_2 \to F$ une application bilinéaire.*
1. *B est continue si et seulement s'il existe $M \geq 0$, tel que $B(x, y) \leq M\|x\|_{E_1}\|y\|_{E_2}$ pour tout x, y dans E_1, E_2.*
2. *Si B est continue alors B est C^1. Pour x, y fixés dans E_1, E_2 on a*
$$B'(x, y)(h, k) = B(h, y) + B(x, k), \quad \forall h, k \in E_1, E_2.$$
Proposition 4. *Soient E, F des espaces de Banach. L'application*
$$\mathrm{Inv} \ : \begin{cases} \mathrm{Isom}(E, F) & \to \ \mathrm{Isom}\,(F, E) \\ \qquad u & \to \ u^{-1} \end{cases}$$
est différentiable et pour tout u inversible :
$$\mathrm{Inv}'(u)(v) = -u^{-1} \circ v \circ u^{-1}, \ \forall v \in \mathcal{L}(E, F).$$

2.2 Règles de calculs

Proposition 5 (Linéarité de la différentiation)**.** *Soit E un espace vectoriel normé, Ω un ouvert de E, et $\bar{x}$ un élément de E.*
Soit $\lambda, \mu \in \mathbb{R}$ et $f, g \in \Delta^1_{\bar{x}}$ (resp. $C^1_{\bar{x}}$). On a $\lambda f + \mu g \in \Delta^1_{\bar{x}}$ (resp. $C^1_{\bar{x}}$) et

$$(\lambda f + \mu g)'(\bar{x}) = \lambda f'(\bar{x}) + \mu g'(\bar{x}).$$

Proposition 6 (Différentielle d'une composée : "chain rule")**.** *Soient E, F, G des espaces vectoriels normés, $\Omega \ni \bar{x}$ un ouvert de E, et Γ un ouvert de F.*

Soient $f : \Omega \to F$, $g : \Gamma \to G$ telles que $f \in \Delta^1_{\bar{x}}$, $f(\bar{x}) \in \Gamma$ et $g \in \Delta^1_{f(\bar{x})}$. Alors $g \circ f \in \Delta^1_{\bar{x}}$ et

$$(g \circ f)'(\bar{x}) = g'(f(\bar{x})) \circ f'(\bar{x}).$$

Remarque 3. *On peut remplacer l'hypothèse $\Delta^1_{\bar{x}}$ par $C^1_{\bar{x}}$.*

2.3 Applications à la minimisation des fonctions

2.3.1 Notions de bases

Vocabulaire de base

Dans un espace vectoriel normé, on se donne une fonction $f : E \to \mathbb{R}$ et un sous-ensemble $C \subset E$, généralement fermé dans E et l'on étudie le problème

$$(\mathcal{P}) \quad \inf_{x \in C} f(x).$$

La fonction f est le *critère*, la *fonction objectif*, le *coût*; l'ensemble C est l'ensemble des *contraintes*; la variable x est la variable de *décision*. Lorsque $x \in C$, on dit que x est *réalisable* ou encore *admissible*.
On dit que x^* est *solution de* $(\mathcal{P})$ si $f(x^*) = \inf_{x \in C} f(x)$ et $x^* \in C$ autrement dit si

$$f(x^*) \leq f(x),\ \forall x \in C.$$

On dit que x^* est *une solution locale de* $(\mathcal{P})$ si x^* est dans C et s'il existe $r > 0$ tel que

$$f(x^*) \leq f(x),\ \forall x \in C \cap B(x^*, r).$$

Par la suite, nous écrirons indifféremment

$$\inf_{x \in C} f(x) = \inf\{f(x) : x \in C\} = \inf_C f.$$

Rappels : bornes sup et inf dans $\mathbb{R}$.

Soit A un sous-ensemble non vide de $\mathbb{R}$. Soit $\ell \in (-\infty, +\infty]$. L'identité $\ell = \sup A$ équivaut à :

1. toute suite convergente d'éléments de A a une limite inférieure à ℓ

2. il existe une telle suite qui converge vers ℓ.

On notera que $\sup A = +\infty$ et $\sup A \notin A$ sont deux éventualités à ne pas exclure. Lorsque $\sup A \in A$ on dit que le sup est atteint et l'on écrit :

$$\sup A = \max A.$$

On définit similairement $\inf A \in [-\infty, +\infty)$ et $\min A$.

Remarque 4. (a) Revenons à $(\mathcal{P})$: l'égalité $\ell = \inf_{x \in C} f(x)$ se traduit par l'existence de $(x_k)_{k\in\mathbb{N}}$ dans C telle que $f(x_k) \to \ell$. La suite $(x_k)_{k\in\mathbb{N}}$ est dite *suite minimisante.*
Bien noter que $(x_k)_{k\in\mathbb{N}}$ peut ne pas converger et que $\ell = -\infty$ n'est pas à exclure.
(b) On a (par "convention") $\sup \emptyset = -\infty$ et $\inf \emptyset = +\infty$.

L'identité

$$\sup_C f = -\inf_C(-f)$$

permet d'étudier simultanément les problèmes de maximisation et de minimisation.

Existence de minima/maxima

Théorème 10 (Weierstrass)**.** *Une fonction continue $f : K \to \mathbb{R}$ sur un compact K non vide atteint ses bornes. Formellement*

$$\inf_K f = \min_K f \ \text{ et } \ \sup_K f = \max_K f.$$

2.3.2 Problèmes sans contrainte : condition nécessaire du premier ordre, règle de Fermat

Théorème 11 (Règle de Fermat)**.** *Soit $f : E \to \mathbb{R}$ une fonction différentiable. Si x^* est solution de $\inf_E f$ (resp. $\sup_E f$) alors*

$$f'(x^*) = 0.$$

Remarque 5.
1. Comme on le constate dans la démonstration : si x^* est un min (resp. un max) local de f on a tout autant $f'(x^*) = 0$
2. En toute généralité un point tel que $f'(x^*) = 0$ est dit critique. Ce n'est pas en général un maximum ou un minimum –prendre par exemple $f(x) = x^3$ sur $\mathbb{R}$ et $x^* = 0$.

2.3.3 Conditions nécessaires et suffisantes en optimisation convexe

Définition 4 (Convexes)**.** *Soit E un espace vectoriel et $C \subset E$. On dit que C est convexe si*

$$\forall(\lambda, x, y) \in [0,1] \times C \times C, \lambda x + (1-\lambda)y \in C.$$

Définition 5 (Fonctions convexes)**.** *Soit E un espace vectoriel, C un convexe de E et $f : C \to \mathbb{R}$.*

1. *On dit que f est convexe si* $\forall(\lambda, x, y) \in [0,1] \times C \times C$,
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$
2. *On dit que f est strictement convexe si* $\forall \lambda \in]0,1[, \forall x \neq y \in C$,
$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y).$$

Example 2. On vérifie facilement que les formes affines et les normes sont convexes et qu'elles ne sont jamais strictement convexes.

Le carré d'une norme Hilbertienne est convexe :
$$H \ni x \to \|x\|^2.$$
La preuve peut se faire directement (exercice) ou en usant de Proposition 13

De façon générale le moyen le plus simple d'établir la convexité n'est pas l'utilisation de la définition mais plutôt celui de la caractérisation via la différentielle seconde (voir Proposition 13).

Proposition 7 (Caractérisation de la convexité)**.** *Soit E un espace vectoriel normé. On suppose $f : C \to \mathbb{R}$ différentiable sur un convexe ouvert non vide C de E. Les propriétés suivantes sont équivalentes :*

1. *f est convexe*
2. *f est au dessus de toutes ses tangentes, i.e.*
$$f(y) \geq f(x) + f'(x)(y-x)$$
pour tout x, y dans C.

Une conséquence quasi-immédiate du résultat ci-dessus est :

Théorème 12 (Minima des fonctions convexes)**.** *Soient E un espace vectoriel normé et $f : E \to \mathbb{R}$ une fonction dérivable convexe. Soit x dans E. Alors :*
$$x^* \text{ est un minimum global de } f \iff f'(x^*) = 0.$$

2.4 Rôles des coordonnées et des espaces sous-jacents

2.4.1 Dérivées directionnelles

Définition 6. *Soit E un espace vectoriel normé, Ω un ouvert de E, et $\bar{x}$ un élément de E. Étant donné v dans E, on dit que f admet une dérivée directionnelle en $\bar{x}$ dans la direction v si la quantité/dérivée ci-dessous est bien définie :*
$$f'(\bar{x}; v) := \lim_{t \to 0, t \neq 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}$$

(a) Lorsque f est différentiable en $\bar{x}$ on a $f'(\bar{x}; v) = f'(\bar{x})(v)$.

(b) $E = \mathbb{R}^n$. On désigne par $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. Lorsque f est différentiable en $\bar{x}$ dans la direction e_i, on a $f'(\bar{x}; e_i) = \frac{\partial f}{\partial x_i}(\bar{x})$. Ainsi lorsque f est différentiable en $\bar{x}$, on a la formule bien connue
$$f'(\bar{x})(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\bar{x}) h_i.$$

2.4.2 Différentielles par bloc de coordonnées

E_1, E_2, F sont des espaces vectoriels normés. Soit $f : E_1 \times E_2 \to F$. Sous réserve d'existence, la différentielle partielle de f en $(\bar{x}, \bar{y})$ par rapport à x (le premier bloc de coordonnées) est donnée par la différentielle de $f(\cdot, \bar{y})$ en $\bar{x}$. Elle est notée $\frac{\partial f}{\partial x}(\bar{x}, \bar{y})$ qui est donc un élément de $\mathcal{L}(E_1, F)$. On définit de même $\frac{\partial f}{\partial y}(\bar{x}, \bar{y}) \in \mathcal{L}(E_2, F)$. Lorsque ces deux objets existent, on dit que f *admet des différentielles partielles*. Il est tentant de penser qu'"admettre des différentielles partielles" revient à "être différentiable", ce n'est pas toujours le cas comme nous allons le voir. Par contre :

Proposition 8. *Soit U un ouvert de $E_1 \times E_2$. Si $f : U \subset E_1 \times E_2 \to F$ est différentiable en $(\bar{x}, \bar{y}) \in U$ alors f admet des différentielles partielles en $(\bar{x}, \bar{y})$ et*

$$f'(\bar{x}, \bar{y})(h, k) = \frac{\partial}{\partial x} f(\bar{x}, \bar{y})h + \frac{\partial}{\partial y} f(\bar{x}, \bar{y})k,$$

pour tout (h, k) dans $E_1 \times E_2$.

Comme déjà annoncé la réciproque est fausse. Toutefois, le théorème des accroissements finis (voir Théorème 15) permet d'établir :

Proposition 9. *Même notations. Si les différentielles partielles existent autour d'un point $(\bar{x}, \bar{y})$ et sont continues en ce point alors la fonction f est C^1 en $(\bar{x}, \bar{y})$.*

2.4.3 Fonctions à valeurs dans un produit

Proposition 10. *Soient $E, E_1, \ldots, E_n$ des espaces vectoriels normés et f une application d'un ouvert $\Omega \subset E$ dans $E_1 \times \ldots \times E_n$. On note $f(x) := (f_1(x), ..., f_n(x))$ pour tout x de E.*

Alors f est différentiable en $\bar{x}$ si et seulement si les fonctions $f_1, \ldots, f_n$ sont différentiables en $\bar{x}$. On a alors

$$f'(\bar{x})(h_1, \ldots, h_n) = (f_1'(\bar{x})h_1, \ldots, f_n'(\bar{x})h_n).$$

Représentation matricielle de la différentielle lorsqu'en plus la source est un produit : Jacobien (révisions).

2.4.4 Espaces de Hilbert et notion de gradient

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. On désigne par H^* l'espace des formes linéaires ([1]) continues sur H, ainsi $H^* = \mathcal{L}(H, \mathbb{R})$.

1. Lorsque le but est réel les applications linéaires sont souvent désignées par le terme de *formes linéaires*.

Théorème 13 (Théorème de représentation de Riesz). *L'application*

$$\mathcal{I} : \begin{cases} H & \longrightarrow H^* \\ x & \longrightarrow \begin{cases} H & \longrightarrow \mathbb{R} \\ y & \longrightarrow \langle x, y\rangle \end{cases} \end{cases}$$

est une isométrie de $\big(H, \|\cdot\|\big)$ *vers* H^* *muni de sa norme d'opérateur* $\|\cdot\|_{\mathcal{L}(E,F)}$.

Cela permet donc une représentation simple des formes linéaires continues sur les espaces de Hilbert (et justifie l'identification $H \simeq H^*$).

Dans un premier temps on retiendra surtout que toute forme linéaire $\ell \in \mathcal{L}(H, \mathbb{R})$ s'écrit de manière unique $\ell = \langle z, \cdot\rangle$ avec z dans H.

Soit $\Omega \ni \bar{x}$ un ouvert de H espace de Hilbert et $f : \Omega \to \mathbb{R}$ différentiable en $\bar{x}$. Par définition $f'(\bar{x}) \in \mathcal{L}(H, \mathbb{R}) = H^*$, ainsi, par le théorème de représentation de Riesz, il existe un unique vecteur $z_{\bar{x}}$ tel que

$$f'(\bar{x})(h) = \langle z_{\bar{x}}, h\rangle, \quad \forall h \in H.$$

Le vecteur $z_{\bar{x}}$ est *par définition*, le gradient de f en $\bar{x}$, il est noté $\nabla f(\bar{x})$. Ainsi

$$f'(\bar{x})(h) = \langle \nabla f(\bar{x}), h\rangle, \quad \forall h \in H.$$

$$f(\bar{x} + h) = f(\bar{x}) + \langle \nabla f(\bar{x}), h\rangle + o(h) \text{ pour } h \text{ petit.}$$

On vérifie sans peine que si $H = \mathbb{R}^n$ et $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$, on retrouve

$$\nabla f(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), \ldots, \frac{\partial f}{\partial x_n}(\bar{x})\right).$$

Ainsi pour la métrique euclidienne de $\mathbb{R}^n$ la définition coïncide donc avec celle via les coordonnées.

Interprétation géométrique “gradient et lignes de niveau" : dès que le gradient est non nul, il est orthogonal aux lignes de niveau et pointe dans le sens des valeurs croissantes de f.

2.5 Théorème fondamentaux : l'inégalité des AF, inversion locale et fonctions implicites

2.5.1 Théorème des accroissements finis

Théorème 14 (Rolle). *Soient* $a < b$ *dans* $\mathbb{R}$ *et* $f : [a, b] \to \mathbb{R}$ *continue, dérivable sur* $]a, b[$. *Alors il existe* $c \in]a, b[$ *tel que* $f'(c) = 0$.

Théorème 15 (Accroissements finis)**.** *Soient E et F des espaces vectoriel normés, $f : \Omega \subset E \to F$ une fonction différentiable sur un ouvert Ω de E. Soient $[a, b] \subset \Omega$ et $M \geq 0$ tels que*

$$\|f'(x)\|_{\mathcal{L}(E,F)} \leq M, \ \forall x \in]a, b[. \tag{2.1}$$

Alors

$$\|f(b) - f(a)\|_F \leq M\|b - a\|_E.$$

Remarque 6. (a) Lorsque f est C^1 l'hypothèse (2.1) est automatiquement vérifiée.
(b) Si $\Omega = C$ est convexe, on déduit de ce qui précède le fameux résultat : "Si $f'(x) = 0$ pour tout x de C alors f est constante." Prendre bien garde toutefois au fait que l'hypothèse de convexité ([2]) joue un rôle important.

2.5.2 Inversion locale et fonctions implicites

Dans ce qui suit E, F, G sont des *espaces de Banach* – l'hypothèse est essentielle.

Théorème 16 (Inversion locale)**.** *Soient $\Omega \ni \bar{x}$ un ouvert de E et $f \in C^1(\Omega, F)$ telle que*

$$f'(\bar{x}) \in \mathrm{Isom}(E, F).$$

Alors il existe des voisinages ouverts U de $\bar{x}$, V de $f(\bar{x})$ tels que $f : U \to V$ soit un C^1 difféomorphisme, i.e. $f : U \to V$ est bijective et son inverse $f^{-1} : V \to U$ est C^1.

Théorème 17 (Fonctions implicites)**.** *Soient Ω un ouvert de $E \times F$ contenant $(\bar{x}, \bar{y})$ et $f : \Omega \subset E \times F \to G$ une application C^1 telle que*

$$\begin{cases} f(\bar{x}, \bar{y}) = 0, \\ \dfrac{\partial f}{\partial y}(\bar{x}, \bar{y}) \in \mathit{Isom}(F, G). \end{cases}$$

Alors il existe

$$\begin{cases} \text{un ouvert } U \text{ de } \bar{x} \text{ dans } E, \\ \text{un ouvert } \Gamma \text{ de } (\bar{x}, \bar{y}) \text{ dans } E \times F, \\ \text{une application } \Psi : U \to F \ C^1 \text{ avec } \bar{y} = \Psi(\bar{x}), \end{cases}$$

tels que

$$\begin{cases} (x, y) \in \Gamma \\ \ \ et \\ f(x, y) = 0 \end{cases} \iff \begin{cases} x \in U \\ \ \ et \\ y = \Psi(x). \end{cases}$$

On a de plus

$$\Psi'(\bar{x}) = -\frac{\partial f}{\partial y}(\bar{x}, \bar{y})^{-1} \circ \frac{\partial f}{\partial x}(\bar{x}, \bar{y}).$$

2.6 Différentielles d'ordre 2

Dans ce qui suit E, F sont des espaces vectoriels normés, $\Omega \subset E$ un ouvert de E, $\bar{x}$ un point de Ω et $f : \Omega \to F$ une fonction.

2. Plus généralement de connexité.

2.6.1 Définition

On suppose $f \in \Delta^1(\Omega)$. On a alors $f' : \Omega \to \mathcal{L}(E,F)$. Ainsi si f' est différentiable en $\bar{x}$, on a

$$(f')'(\bar{x}) = f''(\bar{x}) \in \mathcal{L}(E, \mathcal{L}(E,F)).$$

On définit $\mathcal{L}_2(E;F)$ comme étant l'espace des applications bilinéaires continues de $E \times E$ dans F. Il a une structure naturelle d'espace vectoriel normé définie par

$$\|B\|_{\mathcal{L}_2(E;F)} = \sup\{M : \|B(x,y)\|_F \leq M\|x\|_E\|y\|_E, \forall x, y \in E \times E\}$$

pour B dans $\mathcal{L}_2(E;F)$.

Proposition 11. *L'application*

$$J : \begin{cases} \mathcal{L}(E,\mathcal{L}(E,F)) & \longrightarrow & \mathcal{L}_2(E;F) \\ u & \longrightarrow & J(u) := \begin{cases} E \times E & \to & F \\ (x,y) & \to & u(x)(y) \end{cases} \end{cases}$$

est une isométrie linéaire, c'est-à-dire

$$\|J(u)\|_{\mathcal{L}_2(E;F)} = \|u\|_{\mathcal{L}(E,\mathcal{L}(E,F))}$$

pour tout u de $\mathcal{L}(E,\mathcal{L}(E,F))$.

Cela justifie l'identification $\mathcal{L}(E,\mathcal{L}(E,F)) \simeq \mathcal{L}_2(E;F)$ que nous ferons systématiquement par la suite.

Définition 7 (Différentiabilité seconde)**.**
1. On dit que f est deux fois différentiable en $\bar{x}$ (on écrit $f \in \Delta^2_{\bar{x}}$), s'il existe un voisinage ouvert $U \ni \bar{x}$ tel que $f \in \Delta^1(U)$ et si l'application $f' : U \to \mathcal{L}(E,F)$ est différentiable en $\bar{x}$.
La différentielle seconde est notée $f''(\bar{x}) \in \mathcal{L}_2(E;F)$.
2. On dira que f est C^2 en $\bar{x}$, s'il existe un voisinage ouvert $U \ni \bar{x}$ tel que $f \in \Delta^2(U)$ et si l'application $f'' : U \to \mathcal{L}_2(E;F)$ est continue en $\bar{x}$.

2.6.2 Théorème de Clairaut-Schwarz

Proposition 12 (Clairaut-Schwarz)**.** *On suppose que $f \in \Delta^2_{\bar{x}}$. Alors $f''(\bar{x})$ est symétrique, i.e.*

$$f''(\bar{x})(h,k) = f''(\bar{x})(k,h), \quad \forall (h,k) \in E^2.$$

Lorsque $E = \mathbb{R}^n$ et $F = \mathbb{R}$ le théorème ci-dessus revient à dire que

$$\frac{\partial f}{\partial x_i \partial x_j}(\bar{x}) = \frac{\partial f}{\partial x_j \partial x_i}(\bar{x})$$

pour tout i, j dans $\{1, \ldots, n\}$ (simplement remarquer que $f''(\bar{x})(e_i, e_j) = \frac{\partial f}{\partial x_i \partial x_j}(\bar{x})$).

2.6.3 Formule de Taylor-Young

On adoptera la notation suivante : $f''(x)(h,h) = f''(x)(h)^2$ – le carré est une *puissance symbolique*.

Théorème 18 (Formule de Taylor-Young à l'ordre 2)**.** *Soit* $f \in \Delta^2_{\bar{x}}$.
Pour h *dans un voisinage* U *de* 0_E*, on a*

$$f(\bar{x}+h) = f(\bar{x}) + f'(\bar{x})(h) + \frac{1}{2}f''(\bar{x})(h)^2 + \|h\|^2\epsilon(h),$$

où $\epsilon : U \subset E \to F$ *est continue et nulle en* 0_E*.*

Un cas important : $E = \mathbb{R}^n$, $F = \mathbb{R}$. On pose

$$\mathrm{Hess} f(\bar{x}) = [f''(\bar{x})(e_i, e_j)]_{1\leq i,j\leq n}.$$

$\mathrm{Hess} f(\bar{x})$ est la *matrice hessienne* de f en $\bar{x}$.
La formule de Taylor Young s'écrit comme suit

$$\begin{aligned} f(\bar{x}+h) &= f(\bar{x}) + \langle \nabla f(\bar{x}), h\rangle + \frac{1}{2}\langle \text{Hess } f(\bar{x})h, h\rangle + o(\|h\|^2) \\ &= f(\bar{x}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\bar{x})h_i + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \frac{\partial f}{\partial x_i \partial x_j}(\bar{x})h_i h_j + o(\|h\|^2) \end{aligned}$$

2.6.4 Convexité et différentiabilité seconde

On dit que $B \in \mathcal{L}_2(E, \mathbb{R})$ est *positive* si $B(h,h) \geq 0$ pour tout h de E, *définie positive* si $B(h,h) > 0$ pour tout $h \neq 0$ de E.

Proposition 13. $f : C \to \mathbb{R}$ *une fonction deux fois différentiable sur un ouvert convexe* C*.*
La fonction f *est convexe si et seulement si* $f''(x)$ *est positive en tout* x *de* C

2.7 Conditions d'ordre 2 en optimisation

Ces conditions sont éminemment géométriques et sont à comprendre à la lumière de la formule de Taylor-Young à l'ordre 2.

Théorème 19 (Condition nécessaire d'ordre 2)**.** *Soit* $f : \Omega \to \mathbb{R}$ *deux fois différentiable.*
Si x^* *est un minimum local alors* $f''(x^*)$ *est positive.*

Théorème 20 (Condition suffisante d'ordre 2)**.** *Mêmes hypothèses que ci-dessus.*
Si $f'(x^*) = 0$ *et* $f''(x^*)(h)^2 \geq \alpha||h||^2, \forall h \in E$ *avec* $\alpha > 0$ *alors* x^* *est un minimum local.*

Chapitre 3

Initiation à l'optimisation

Par souci de simplicité nous supposerons désormais que E est de dimension finie. On se donne un produit scalaire sur E noté $\langle \cdot, \cdot \rangle$. Les quantités m, n, p désignent des entiers positifs.

3.1 Optimisation avec contraintes d'inégalité

3.1.1 Présentation du problème

$$(P) \quad \inf \{ f(x) : x \in E, f_i(x) \leq 0, \ \forall i = 1 \ldots, m \} .$$

Lemme 1 (Farkas). *Soient $a_1, ..., a_m, b$ des vecteurs de $\mathbb{R}^n$.*

(i) Pour tout x de $\mathbb{R}^n$,

$$\Big(\Big(\langle a_i, x \rangle \leq 0, \ \forall i = 1, \ldots, m \Big) \Rightarrow \langle b, x \rangle \leq 0 \Big)$$

(ii) Il existe des réels $\lambda_1, ..., \lambda_m$ positifs ou nuls tels que

$$b = \lambda_1 a_1 + ... + \lambda_m a_m.$$

3.1.2 Relations de Karush-Kuhn-Tucker

On suppose désormais $f_1, ..., f_m$ différentiables. Soit x un point réalisable du problème (P). On désigne par $I(x)$ l'ensemble des indices actifs, i.e. l'ensemble des indices i tels que $f_i(x) = 0$ – on dit que la contrainte i est active.

Afin de pouvoir écrire des conditions du premier ordre simple pour (P), il est souhaitable que les "normales (extérieures)" à la frontière de l'ensemble des contraintes soient aisément descriptibles par les gradients des contraintes. Les conditions de qualification servent précisément à assurer une représentation simple de ces normales.

Condition de qualification de Arrow-Hurwicz-Uzawa. Un point x réalisable est dit *qualifié* s'il existe $y \in E$ tel que :

$$\langle y, \nabla f_i(x) \rangle < 0 \text{ pour tout } i \text{ dans } I(x).$$

Théorème 21 (Relations “KKT"). *Soient $f, f_1, ..., f_m : E \to \mathbb{R}$ des fonctions différentiables. On suppose le problème qualifié en un point x^* réalisable de (P).*

Si x^* est solution de (P), ou même si x^* est un minimum local pour (P), alors il satisfait les conditions de Karush-Kuhn-Tucker :

(i) $f_i(x^*) \leq 0$ pour tout $i = 1, \dots, m$ (Réalisibilité)

Il existe des réels $\lambda_1, ..., \lambda_m$ tels que

(ii) $\lambda_i \geq 0$ pour tout $i = 1, \dots, m$ (positivité)

(iii) $\lambda_i f_i(x^*) = 0$ pour tout $i = 1, \dots, m$ (Complémentarité)

(iv) $\nabla f(x^*) + \lambda_1 \nabla f_1(x^*) + ... + \lambda_m \nabla f_m(x^*) = 0$ (Conditions de Lagrange)

On dira qu'un point x satisfait les relations KKT s'il satisfait les conditions $(i)-(ii)-(iii)-(iv)$ ci-dessus.

3.1.3 Cas convexe

Théorème 22 (CNS d'ordre 1). *Soient $f, f_1, ..., f_m : E \to \mathbb{R}$ des fonctions différentiables convexes.*

On suppose qu'il existe x_0 tel que $f_i(x_0) < 0$ pour tout $i = 1, \dots, m$ (Condition de Slater). Soit $x^* \in E$.
Alors :

x^* est solution de (P) si et seulement si x^* satisfait les conditions KKT.

3.2 Contraintes d'égalités

On se donne $g_1, \dots, g_p$ des fonctions C^1 et l'on considère le problème ([1])

$$(P) \quad \inf \{f(x) : \; g_i(x) = 0, \, \forall i = 1, \dots, p\}.$$

Il s'agit du premier problème considéré historiquement par Lagrange.

On dira que le problème (P) est qualifié en un point (réalisable) x si

$$\{\nabla g_i(x)\}_{i=1,\dots,p} \text{ forme un système libre.}$$

Cette condition nous assure que l'ensemble des contraintes peut localement se décrire comme un graphe de fonction C^1 dont l'espace normal (ie l'espace orthogonal à l'espace tangent) est généré en tout point par $\{\nabla g_i(x)\}_{i=1,\dots,p}$.

On en déduit le second théorème fondateur de l'optimisation, après celui de Fermat :

Théorème 23 (Lagrange). *Soient f différentiable et $g_1, ..., g_p : E \to \mathbb{R}$ des fonctions C^1. On suppose le problème (P) qualifié en un point x^* réalisable.*

Si x^* est solution de (P), ou même si x^* est un minimum local pour (P), alors il satisfait les conditions de Lagrange :

1. Bien remarquer que l'on exige ici plus de régularité

(i) $g_i(x^*) = 0$ *pour tout* $i = 1, \ldots, p$ *(Réalisibilité)*

(ii) *Il existe des réels* $\mu_1, ..., \mu_p$ *tels que :*
$\nabla f(x^*) + \mu_1 \nabla g_1(x^*) + ... + \mu_p \nabla g_p(x^*) = 0$ *(Conditions de Lagrange)*

A moins que les fonctions $g_1, \ldots, g_p$ ne soient affines, l'ensemble des contraintes n'est pas en général un convexe de E.

3.3 Contraintes d'égalité et d'inégalité

On considère ici un problème général de la forme

$$(P) \quad \inf \{ f(x) : f_i(x) \leq 0, \ g_j(x) = 0, \ \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, p \} .$$

On suppose $f, f_1, ..., f_m : E \to \mathbb{R}$ différentiables et $g_1, \ldots, g_p : E \to \mathbb{R}$ C^1. Soit x un point réalisable du problème (P).

Comme précédemment on désigne par $I(x)$ l'ensemble des indices actifs, i.e. l'ensemble des indices $i = 1, \ldots, m$ tels que $f_i(x) = 0$. Bien noter que $I(x)$ n'inclut pas les indices correspondants aux contraintes d'égalité.

Les conditions de qualification considérées ici combinent les précédentes sous la forme suivante :

Condition de qualification de Mangasarian-Fromovitz. Un point x réalisable est dit *qualifié* s'il existe $y \in E$ tel que :

$$\begin{cases} \langle y, \nabla f_i(x) \rangle < 0 \text{ pour tout } i \text{ dans } I(x), \\ \{\nabla g_j(x)\}_{j=1,\ldots,p} \text{ forme un système libre.} \end{cases}$$

Le lecteur soucieux d'harmonie pourra remarquer que la première hypothèse équivaut à

$$\sum_{i=1}^{m} \nu_i \nabla f_i(x) = 0 \text{ et } \nu_i \geq 0, \ \forall i = 1, \ldots, m \Longrightarrow \nu_i = 0, \ \forall i = 1, \ldots, m,$$

alors que la deuxième revient à

$$\sum_{j=1}^{p} \nu_i \nabla g_j(x) = 0 \Longrightarrow \nu_j = 0, \ \forall j = 1, \ldots, p.$$

Le théorème suivant résume les précédents en un unique résultat :

Théorème 24 (Relations "KKT"). *Soient* $f, f_1, ..., f_m : E \to \mathbb{R}$ *des fonctions différentiables et* $g_1, \ldots, g_p : E \to \mathbb{R}$ *des fonctions* C^1*. On suppose le problème qualifié en un point* x^* *réalisable de* (P)*.*

Si x^* *est solution de* (P) *(ou même si* x^* *est un minimum local pour* (P)*) alors il satisfait les conditions :*

(i) $f_i(x^*) \leq 0$ *pour tout* $i = 1, \ldots, m$,
$g_j(x^*) = 0$ *pour tout* $j = 1, \ldots, p$, *(Réalisibilité)*

Il existe des réels $\lambda_1, ..., \lambda_m, \mu_1, \ldots, \mu_p$ *tels que* *(ii)* $\lambda_i \geq 0$ *pour tout* $i = 1, \ldots, m$ *(Positivité)*

(iii) $\lambda_i f_i(x^*) = 0$ *pour tout* $i = 1, \ldots, m$ *(Complémentarité)*

(iv) $\nabla f(x^*) + \lambda_1 \nabla f_1(x^*) + ... + \lambda_m \nabla f_m(x^*) + \mu_1 \nabla g_1(x^*) + \ldots + \mu_p \nabla g_p(x^*) = 0$ *(Conditions de Lagrange)*

Comme précédemment, et sans ambiguïté, on dira qu'un point x satisfait les relations KKT s'il satisfait les conditions $(i) - (ii) - (iii) - (iv)$ ci-dessus.

3.3.1 Problèmes convexes généraux

Théorème 25 (CNS d'ordre 1)**.** *Soient* $f, f_1, ..., f_m, g_1, \ldots, g_p : E \to \mathbb{R}$ *des fonctions différentiables convexes.*

On suppose qu'il existe x_0 *tel que* $f_i(x_0) < 0$ *pour tout* $i = 1, \ldots, m$ *(Condition de Slater). Soit* x^* *dans* E*.*
Alors :

x^* *est solution de* (P) *si et seulement si* x^* *satisfait les conditions KKT.*