1 Rappels de calcul différentiel

On considère ici E et F deux *espaces normés* munie des normes $\|\cdot\|_E$ et $\|\cdot\|_F$. On désigne par f une application définie sur une partie Ω de E à valeurs dans F. On supposera que Ω est ouvert.

1.1 Dérivées directionnelle et au sens de Gâteaux

Définition :
Dérivées directionnelle On dit que f a une dérivée directionnelle en $x \in \Omega$ (ici Ω n'est pas forcément ouvert) dans la direction $h \in E$ si, pour $t > 0$ suffisamment petit, $x + th \in \Omega$ et si la limite

$$f'(x;h) = \lim_{\substack{t\to 0\\ t>0}} \frac{1}{t}\left(f(x+th) - f(x)\right)$$

existe.

Remarque : Pour les fonctions à valeurs dans $\mathbb{R}$, il arrivera que l'on admette des dérivées directionnelles valant $\pm\infty$, c'est-à-dire de prendre la limite dans $\overline{\mathbb{R}}$.

Définition :
Gâteaux différentiabilité On dit qu'une fonction f est Gâteaux-*différentiable* en $x \in \Omega$ si elle admet une dérivée directionnelle en x suivant toutes les directions $h \in E$ et si l'application

$$h \in E \longmapsto f'(x;h) \in F$$

est linéaire continue. On note $f'(x)$ cet opérateur. On a alors

$$\forall h \in E,\ f'(x)\cdot h = f'(x;h)$$

1.2 Dérivée au sens de Fréchet

Définition :
Fréchet différentiabilité On dit que f est Fréchet-*différentiable* en $x \in \Omega$ s'il existe un opérateur linéaire continu L de E dans F tel que

$$\lim_{\substack{\|h\|_E\to 0\\ \|h\|_E>0}} \frac{1}{\|h\|_E}\left(f(x+h) - f(x) - Lh\right) = 0.$$

L'opérateur L est appelé *dérivée* de f en x.

Remarque : La notion de Fréchet-différentiabilité est plus forte que celle de Gâteaux-différentiabilité.

Propriété :
Si $f : \Omega \longrightarrow F$ est F-différentiable en $x \in \Omega$ avec une dérivée L, alors f est G-différentiable en x et $L = f'(x)$.

2 Définition d'un problème d'optimisation

Soit X un ensemble et f une fonction définie sur X à valeur dans $\overline{\mathbb{R}}$. On cherche à resoudre le problème

$$(P_X) = \left\{ \begin{array}{l} \inf f(x) \\ x \in X \end{array} \right.$$

Si $X \neq \emptyset$, on dit que (P_X) est réalisable. On adopte les conventions suivantes :

$$\inf_{x\in\emptyset} f(x) = +\infty \quad \sup_{x\in\emptyset} f(x) = -\infty.$$

2.1 Existence de solution

Définition : Fonction semie-continue inférieurement
Une fonction f est semi-continue inférieurement (s.c.i.) sur X si $\forall x \in X$, pour toute suite $(x_k)_k \longmapsto x$,

$$f(x) \leqslant \lim_{k\to\infty} \inf_{i\geqslant k} f(x_i)$$

Remarque : On dit aussi que f est fermée, son épigraph est fermé

$$\operatorname{epi} f = \{(x,\alpha) \,/\, f(x) \leqslant \alpha\}.$$

Théorème : Weierstrass
Si f est fermée sur X et si X est compact et non vide, alors (P_X) a (au moins une solution.

Remarque : En dimension finie, X compact est équivalalent à X fermé borné. On peut alors (en dimension finie) remplacer l'hypothèse X compact par X fermé et $\lim_{\substack{x\in X \\ \|x\|_E\to\infty}} f(x) = +\infty$.

2.2 Unicité de la solution

Théorème : Unicité de la solution
Si X est *convexe* et si f est *strictement-convexe* sur X, alors $P_X)$ a au plus une solution.

3 Analyse convexe

On considère dans cette section X un convexe inclu dans E et $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ différentiable.

3.1 Caractérisation de la convexité

Théorème :
f est convexe si et seulement si

❶
$$\forall x \neq y \in X, \quad f(x) - f(y) \geqslant f'(x) \cdot (y - x)$$

❷
$$\forall x \neq y \in X, \quad (f'(y) - f'(x)) \cdot (y - x) \geqslant 0$$

❸ si f est deux fois dérivable (X ouvert),
$$\forall x \in X,\ \forall h \in E, \quad f''(x) \cdot h^2 = f''(x) \cdot (h,h) \geqslant 0$$

Pour la *strict convexité*, les inégalités ci-dessus sont strictes.

3.1.1 Projection sur un ensemble convexe

On appelle *projection* de x sur une partie C de E, toute solution éventuelle du problème

$$\begin{cases} \min \frac{1}{2} \left\| y - x \right\|^2 \\ y \in C \end{cases}$$

Propriété : Caractérisation de la projection
Soit C un convexe non vide de E. Un point $\overline{x} \in C$ est une projection de $x \in E$ sur C si et seulement si l'une des conditions équivalentes suivantes est vérifiée :

❶
$$\forall y \in C, \quad \langle y - \overline{x}, \overline{x} - x \rangle \geqslant 0,$$

❷
$$\forall y \in C, \quad \langle y - \overline{x}, y - x \rangle \geqslant 0.$$

3.1.2 Séparation de convexes

On suppose que E est de *dimension finie*, muni du produit scalaire noté $\langle \cdot, \cdot \rangle$. La séparation des deux convexes se fait géométriquement dans E en utilisant un *hyperplan affine* H

$$H := \{x \in E \,/\, \langle \xi, x \rangle = t\},$$

où $\xi \in E$ non nul. On dit que cet hyperplan *sépare* deux convexes C_1 et C_2 si l'on a

$$\forall x_1 \in C_1,\ \forall c_2 \in C_2, \quad \langle \xi, x_1 \rangle \leqslant t \leqslant \langle \xi, x_2 \rangle.$$

Théorème : Séparation stricte de convexes
On peut séparer strictement deux convexes fermés non vides disjoins C_1 et C_2 d'un espace euclidien E dans chacun des cas suivants :

❶ $C_1 - C_2$ est fermé ;

❷ C_1 ou C_2 est compact.

3.2 Cône

Définition : Cône

Une partie K d'un espace vectoriel E est un cône si $tK \subset K$ avec $t \geqslant 0$.

Définition : Cône dual

Soient E un espace vectoriel euclidien, dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$, et P une partie de E. On appelle *cône dual* de P l'ensemble P^+ défini par

$$P^+ := \{x \in E \,/\, \forall y \in P,\ \langle x, y \rangle \geqslant 0\}.$$

Propriété : Cône

1. P^+ est un convexe fermé non vide.
2. P est un convexe fermé non vide si et seulement si $P^{++} = P$.

Théorème : Lemme de Farkas

Soient E et F deux espaces euclidiens, $A : E \longrightarrow F$ une application linéaire et K un cône convexe fermé non vide de E. On note A^* l'application adjointe de A et K^+ le cône dual de K. Alors

$$\left\{y \in F \,\middle/\, A^* y \in K^+\right\}^+ = \overline{\{Ax \,/\, x \in K\}}$$

4 Conditions d'optimalité

Définition : Vecteur tangent

Soient $X \subset E$ et $x \in E$. On dit que le vecteur $d \in E$ est *tangent* à X en x s'il existe une suite $\{d_k\} \subset E$ et une suite $\{t_k\} \subset \mathbb{R}_{++}$ telles que

$$d_k \to d,\ t_k \to 0,\ t > 0, \quad x + t_k d_k \in X$$

On note $T_x X$ l'ensemble des vecteurs tangents à X en x et on l'appelle le *cône tangent*.
Une définition plus pratique est de dire que $d \in E$ est *tangent* à X en x s'il existe une suite $\{x_k\} \subset E$ et une suite $\{t_k\} \subset \mathbb{R}_{++}$ telles que

$$\{x_k\} \subset X,\ t_k \to 0,\ t > 0, \quad \frac{x_k - x}{t_k} \to d$$

Propriété :

$T_x X$ est un cône fermé.

Théorème : CN1

Si x_* est un minimum local de (P_X) et si f est dérivable en x_*, on a

$$\forall d \in T_x X, \quad f'(x_*) \cdot d \geqslant 0.$$

Ceci s'écrit aussi

$$\nabla f(x_*) \in (T_{x_*} X)^+$$

où ∇ est le gradient pour le produit scalaire de E et $(\cdot)^+$ désigne le dual pour ce même produit scalaire.

Propriété : CN1 et CS1 en présence de convexité

Supposons que X soit convexe et que f ait des dérivées directionnelles en un point $x_* \in X$. Si x_* est un minimu local de (P_x), on a

$$\forall x \in X, \quad f'(x_*; x - x_*) \geqslant 0.$$

Inversement, si f est convexe sur le convexe X et si l'équation ci-dessus est vérifiée, alors x_* est un *minimum global* de (P_X).

4.1 Problème sans contrainte

On considère le problème sans contrainte

$$\min_{x \in E} f(x)$$

avec, ici $X = E$ convexe. La condition CN1 et CS1 devient alors :

Propriété : Corrolaire CN1, CS1, sans contrainte

Supposons $f : E \longrightarrow \mathbb{R}$ soit dérivable en $x_* \in E$. Si f a un minimum local en x_* alors

$$f'(x_*) = 0 \quad \text{ou} \quad \nabla f(x_*) = 0.$$

Inversement, si f est convexe et si $f'(x_*) = 0$ alors x_* est un minimum global de f sur E.

Propriété : CN2–CS2

❶ Supposons que x_* soit un minimum local de f sur E et que f soit dérivable dans un voisinage de x_* et deux fois dérivable en x_*. Alors

$$\nabla f(x_*) = 0 \quad \text{et} \quad \nabla^2 f(x_*) \text{ est semie-définie positive.}$$

❷ Si f est dérivable dans un voisinage de x_* et deux fois dérivable en x_* et si

$$\nabla f(x_*) = 0 \quad \text{et} \quad \nabla^2 f(x_*) \text{ est définie positive,}$$

alors x_* est un minimum local strict de f.

4.2 Problèmes avec contraintes d'égalité

On se place dans cette section dans le cas d'un problème d'optimisation dans lequel l'ensemble admissible n'est pas E tout entier mais une partie de celui-ci définie par :

$$X := \{x \in E \,/\, c(x) = 0\}.$$

c est une fonction de E dans F respectivement de dimension n et m. Le problème d'optimisation est alors noté

$$(P_E) = \left\{ \begin{array}{l} \min f(x) \\ c(x) = 0 \end{array} \right. ,$$

où f de E dans $\mathbb{R}$ en est le critère.

4.2.1 Conditions de Lagrange

Propriété : Conditions de Lagrange

Supposons que $c : E \longrightarrow F$ soit dérivable en $x \in X$. Alors

$$T_x X \subset \ker(c'(x)).$$

Si, de plus, $c'(x)$ est surjective et si c est $\mathscr{C}^1$ dans un voisinage de x, alors il y a égalité des deux ensembles.

Définition : Contrainte qualifiée

On dit que la contrainte de (P_E) est *qualifiée* en $x \in X$ si c est dérivable en x et si

$$T_x X = \ker(c'(x)).$$

Propriété : CN1

Soit x_* une solution locale de (P_E). Supposons que f et c soient dérivables en x_* et que la contrainte soit qualifiée en x_*. Alors, il existe un vecteur $\lambda_* \in F$ tel que

$$\nabla f(x_*) + c'(x_*)^* \lambda_* = 0,$$

où $\nabla f(x_*)$ est le gradient de f en x_* et $c'(x_*)^* : F \longrightarrow E$ est l'opérateur adjoint de la jacobienne $c'(x_*)$ pour les produits scalaires donnés sur E et sur F.
Le vecteur λ_* est unique si $c'(x_*)$ est *surjective*. Ce vecteur s'appelle le *multiplicateur de* Lagrange.

On appelle *point stationnaire* du problème (P_E) un point x_* vérifiant ses conditions nécessaires d'optimalité du premier ordre :

$$\left\{ \begin{array}{l} \nabla f(x_*) + c'(x_*)^* \lambda_* = 0 \\ c(x_*) = 0, \end{array} \right.$$

pour un certain multiplicateur $\lambda_* \in F$.

Définition : Lagrangien

On introduit le *lagrangien* qui est la fonction

$$\ell : E \times F \longrightarrow \mathbb{R}$$

définie par

$$\ell(x, \lambda) = f(x) + \langle \lambda, c(x) \rangle.$$

Propriété : CS1

Supposons que le problème (P_E) soit convexe et que $x_* \in E$ vérifie la contrainte de (P_E), que f soit dérivable en x_* et qu'il existe un multiplicateur $\lambda_* \in F$ tel que (x_*, λ_*) vérifie

$$\begin{cases} \nabla_x \ell(x_*, \lambda_*) = 0 \\ c(x_*) = 0 \end{cases}.$$

Alors x_* est un minimum global de (P_E).

4.2.2 Conditions du second ordre

Lemme :

Soit $x_* \in E$ un minimum local de (P_E) et supposons que f et c soient dérivables dans un voisinage de x_* et deux fois dérivable en x_*. On note $L_* = \nabla^2_{xx}\ell(x_*, \lambda_*)$. S'il existe un multiplicateur $\lambda_* \in F$ tel que

$$\nabla_x \ell(x_*, \lambda_*) = 0,$$

alors

$$\forall d \in T_{x_*}X, \quad \langle L_* d, d \rangle \geqslant 0.$$

Propriété : CN2

Soit x_* un minimum local de (P_E) et supposons que f et c soient dérivables dans un voisinage de x_* et deux fois dérivable en x_* et que la contrainte c soit qualifiée en x_*. Alors, il *existe* un multiplicateur $\lambda_* \in F$ tel que l'on ait

$$\nabla_x \ell(x_*, \lambda_*) = 0.$$

De plus L_* est semie-définie positive sur $\ker(c'(x_*))$, c'est-à-dire

$$\forall d \in \ker(c'(x_*)), \quad \langle L_* d, d \rangle \geqslant 0.$$

Propriété : CS2

Supposons que f et c soient dérivables dans un voisinage de x_* et deux fois dérivable en x_*. Supposons que $c(x_*) = 0$ et qu'il existe $\lambda_* \in F$ tel que l'on ait

$$\nabla_x \ell(x_*, \lambda_*) = 0$$

et

$$\forall d \in T_{x_*}X \backslash \{0\}, \quad \langle L_* d, d \rangle > 0.$$

Alors x_* est un minimum local strict de (P_E).