Chapitre 16 :

FONCTIONS DE PLUSIEURS VARIABLES : OPTIMISATION

Dans tout le chapitre, n désigne un entier non nul, $\mathcal{A}$ une partie non vide de $\mathbb{R}^n$, $\mathcal{U}$ un ouvert non vide de $\mathbb{R}^n$ et $f : \mathcal{A} \longrightarrow \mathbb{R}$ une fonction. Dans les situations théoriques où f est définie sur une partie ouverte, on note plutôt $\mathcal{U}$ son domaine de définition.

La théorie de l'*optimisation* s'intéresse aux extremums des fonctions d'une ou plusieurs variables.

Définition 0.1 *On dit que f présente en $A \in \mathcal{A}$:*
*(i) un **maximum global** (ou **absolu**) si :*

$$\forall X \in \mathcal{A}, \quad f(X) \leqslant f(A) ;$$

*(ii) un **maximum local** (ou **relatif**) s'il existe $r > 0$ tel que :*

$$\forall X \in \mathcal{A} \cap \mathcal{B}(A, r), \quad f(X) \leqslant f(A).$$

Remarques 0.2 • On dispose bien entendu des définitions correspondantes de minimul global et local. La notion d'extremum englobe celle de maximum et de minimum.
• Si f admet en $A \in \mathcal{A}$ un extremum global, alors elle admet en ce point un extremum local.
• La fonction f présente en A un extremum si, et seulement si, l'expression $f(X) - f(A)$ garde un signe constant, éventuellement au voisinage de A. C'est souvent ainsi qu'on établira l'existence d'un extremum, en effectuant le changement de variable $X = A + H$ lorsqu'on travaille au voisinage de A.
• Lorsque les inégalités de la définition sont strictes pour $X \neq A$, on parle d'*extremum strict* en A.

1. Recherche d'extremums locaux sur un ouvert

1.1 Condition nécessaire du premier ordre

Théorème 1.1 (Condition nécessaire du premier ordre de présence d'un extremum) *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$ sur un ouvert $\mathcal{U}$.*
Si f admet un extremum local en un point $A \in \mathcal{U}$, alors $\nabla f(A) = 0$. La réciproque est fausse.

Définition 1.2 *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$ sur un ouvert $\mathcal{U}$.*
*On appelle **point critique** de f tout point $A \in \mathcal{U}$ tel que $\nabla f(A) = 0$.*

Remarques 1.3 • Un point $A \in \mathcal{U}$ est critique pour f si, et seulement si, toutes les dérivées directionnelles de f s'annulent en A i.e. pour tout $V \in \mathbb{R}^n$, $\partial_V f(A) = 0$.
• Le caractère *ouvert* de $\mathcal{U}$ dans le théorème 1.1 est essentiel. En effet, la fonction $f : X \longmapsto \|X\|^2$, de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$, admet sur la boule fermée $\mathcal{B}'(0, r)$ un maximum qu'elle atteint en tout point A de la sphère de centre 0 et de rayon r, alors que son gradient $\nabla f(A) = 2A$ n'y est pas nul !
• Le théorème 1.1 peut donc être énoncé ainsi : pour une fonction de classe $\mathscr{C}^1$ sur un ouvert, les points critiques sont les seuls extremum locaux *éventuels*.
• Les points critiques d'une fonction f sur un ouvert $\mathcal{U}$ sont les points en lesquels l'hyperplan tangent au graphe de f est horizontal. En un tel point, la fonction f admet un extremum (local) si, et seulement si, le graphe de f reste (localement) du même côté de cet hyperplan tangent.

Exemple 1.4 Points critiques et extremums (locaux et globaux) de $f : (x, y) \in \mathbb{R}^2 \longmapsto 3xy - x^3 - y^3$.

1.2 Conditions du second ordre

Lemme 1.5 *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{U}$.*
Si A est point critique de f, alors

$$f(A + H) - f(A) = \frac{1}{2} q_A(H) + o(\|H\|^2), \quad H \to 0$$

où q_A désigne la forme quadratique associée à la matrice hessienne $\nabla^2 f(A)$.

En s'appuyant sur ce résultat, on va faire le lien entre la présence d'un extremum local pour f au point A et le signe de $q_{\mathrm{A}}(\mathrm{H})$ lorsque $\mathrm{H} \to 0$. Attention cependant, ce lien n'est pas aussi immédiat que pour les fonctions d'une variable : $f(\mathrm{A}+\mathrm{H}) - f(\mathrm{A})$ n'est pas toujours du signe de $q_{\mathrm{A}}(\mathrm{H})$ lorsque $\mathrm{H} \to 0$.

Remarque 1.6 Les développements limités entretiennent avec les équivalents (hors-programme) des liens relâchés pour les fonctions de $n \geqslant 2$ variables.
Par exemple, la fonction $f : (x, y) \longmapsto x^2 - y^2 + y^3$ présente le développement limité

$$f(x, y) = x^2 - y^2 + o\big(\|(x, y)\|^2\big), \quad (x, y) \to (0, 0)$$

mais on ne peut pas en déduire que $f(x, y) \sim x^2 - y^2$ lorsque $(x, y) \to (0, 0)$. En effet, on observe par exemple que

$$f(t + t^2, t) = (t + t^2)^2 - t^2 + t^3 = 3t^3 + t^4 \sim 3t^3, \quad t \to 0$$

n'est pas équivalent à $(t + t^2)^2 - t^2 \sim 2t^3$ lorsque $t \to 0$.

Lemme 1.7 *Soient* $\mathrm{S} \in \mathbf{M}_n(\mathbb{R})$ *une matrice symétrique de plus petite valeur propre* λ_1 *(resp. de plus grande valeur propre* λ_n*) et* q *la forme quadratique sur* $\mathbb{R}^n$ *canoniquement associée à* S.
On a :

$$\forall \mathrm{X} \in \mathbb{R}^n, \quad \lambda_1 \|\mathrm{X}\|^2 \leqslant q(\mathrm{X}) \leqslant \lambda_n \|\mathrm{X}\|^2$$

avec égalité dans l'inégalité de gauche (resp. de droite) si, et seulement si, X *est nul ou vecteur propre de* S *pour la valeur propre* λ_1 *(resp.* λ_n*).*

Théorème 1.8 (Condition suffisante du second ordre de présence d'un extremum) *On suppose que* f *est de classe* $\mathscr{C}^2$ *sur un ouvert* $\mathcal{U}$. *Soient* $\mathrm{A} \in \mathcal{U}$ *un point critique de* f *et* q_{A} *la forme quadratique associée à la hessienne* $\nabla^2 f(\mathrm{A})$.
Sous les conditions équivalentes suivantes :
(i) pour tout $\mathrm{H} \neq 0$, $q_{\mathrm{A}}(\mathrm{H}) < 0$ *(resp.* $q_{\mathrm{A}}(\mathrm{H}) > 0$*) ;*
(ii) les valeurs propres de la matrice hessienne $\nabla^2 f(\mathrm{A})$ *sont toutes strictement négatives (resp. strictement positives),*
la fonction f *présente en* A *un maximum local strict (resp. un minimum local strict).*
La réciproque est fausse.

Exemple 1.9 Extremums locaux de la fonction $f : (x, y) \in \mathbb{R}^2 \longmapsto x^4 + y^2$.

Théorème 1.10 (Condition nécessaire du second ordre de présence d'un extremum) *On suppose que* f *est de classe* $\mathscr{C}^2$ *sur un ouvert* $\mathcal{U}$. *Soient* $\mathrm{A} \in \mathcal{U}$ *et* q_{A} *la forme quadratique associée à la hessienne* $\nabla^2 f(\mathrm{A})$.
Si f *présente en* A *un maximum local (resp. un minimum local), alors les conditions équivalentes suivantes sont satisfaites :*
(i) pour tout $\mathrm{H} \neq 0$, $q_{\mathrm{A}}(\mathrm{H}) \leqslant 0$ *(resp.* $q_{\mathrm{A}}(\mathrm{H}) \geqslant 0$*) ;*
(ii) les valeurs propres de la matrice hessienne $\nabla^2 f(\mathrm{A})$ *sont toutes négatives ou nulles (resp. positives ou nulles).*
La réciproque est fausse.

Exemple 1.11 Extremums locaux de la fonction $f : (x, y) \in \mathbb{R}^2 \longmapsto x^3 + y^2$.

Corollaire 1.12 (Condition suffisante du second ordre d'absence d'extremum) *On suppose que* f *est de classe* $\mathscr{C}^2$ *sur un ouvert* $\mathcal{U}$. *Soient* $\mathrm{A} \in \mathcal{U}$ *et* q_{A} *la forme quadratique associée à la hessienne* $\nabla^2 f(\mathrm{A})$.
Sous les conditions équivalentes suivantes :
(i) il existe $\mathrm{H}_1, \mathrm{H}_2 \in \mathbb{R}^n$ *tels que* $q_{\mathrm{A}}(\mathrm{H}_1) > 0$ *et* $q_{\mathrm{A}}(\mathrm{H}_2) < 0$ *;*
(ii) la matrice hessienne $\nabla^2 f(\mathrm{A})$ *admet une valeur propre strictement positive et une autre strictement négative,*
la fonction f *ne présente pas en* A *d'extremum local.*

Exemple 1.13 Le point critique $(0,0,0)$ est-il un extremum local pour la fonction

$$f : (x,y,z) \in \mathbb{R}^3 \longmapsto xy + yz + zx - xyz\,?$$

1.3 Cas $n = 2$

Dans le cas d'une fonction de $n = 2$ variables, on reformule les énoncés précédents en utilisant les notations de Monge. Ces notations et le résultat ci-dessous sont hors-programme.

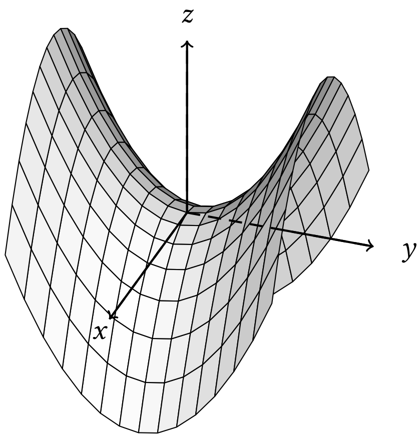
Théorème 1.14 *On suppose que $f : \mathcal{U} \longrightarrow \mathbb{R}$ est de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{U}$ de $\mathbb{R}^2$. Soient $\mathrm{A} \in \mathcal{U}$ un point critique de f et*

$$\nabla^2 f(\mathrm{A}) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

la matrice hessienne de f en A.
(i) Si $rt - s^2 > 0$, alors f admet en A un extremum local : un maximum si $r < 0$ et un minimum si $r > 0$.
(ii) Si $rt - s^2 < 0$, alors la fonction f n'admet pas d'extremum local en A.
(iii) Le cas $rt - s^2 = 0$ est indéterminé.

Remarque 1.15 Dans le cas où les deux valeurs propres de la hessienne sont non nulles de signes contraires, on dit qu'on a un point selle ou que la fonction présente un col.

Exemples 1.16 (i) La fonction $f : (x,y) \longmapsto y^2 - x^2$, représentée ci-dessous, présente un col en $(0,0)$.



(ii) Étudier les extremums locaux de la fonction $f : (x,y) \in \mathbb{R}^2 \longmapsto 3xy - x^3 - y^3$.
(iii) Les exemples 1.9 et 1.11 et montrent que le cas $rt - s^2 = 0$ est indéterminé.

2. Recherche d'extremums globaux et applications

2.1 Sur un ouvert

La méthode consiste à rechercher d'abord les extremums locaux en utilisant les techniques développées dans le paragraphe précédent, puis à vérifier s'il s'agit d'extremums globaux. Il est des situations où des majorations/minorations simples peuvent suffire. On peut également utiliser la formule de Taylor avec reste intégral déclinée ci-dessous pour les fonctions à plusieurs variables (qu'il faut savoir redémontrer).

Proposition 2.1 *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{U}$.*
Pour $\mathrm{A} \in \mathcal{U}$ et $\mathrm{H} \in \mathbb{R}^n$ tels que $[\mathrm{A}, \mathrm{A}+\mathrm{H}] \subset \mathcal{U}$, on a :

$$f(\mathrm{A}+\mathrm{H}) = f(\mathrm{A}) + \langle \nabla f(\mathrm{A}), \mathrm{H} \rangle + \int_0^1 (1-t) q_{\mathrm{A}+t\mathrm{H}}(\mathrm{H})\,\mathrm{d}t$$

où q_{B} désigne la forme quadratique associée à la hessienne $\nabla^2 f(\mathrm{B})$ en tout point $\mathrm{B} \in \mathcal{U}$.

On dispose par exemple du résultat ci-dessous (en pratique, le raisonnement doit être rappelé brièvement) :

Corollaire 2.2 *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^2$ sur un ouvert* convexe *$\mathcal{U}$.*
Si la forme quadratique q_A associée à la hessienne $\nabla^2 f(A)$ est négative (resp. positive) en tout point $A \in \mathcal{U}$, alors tout point critique de f en est un maximum global (resp. minimum global).

Exemple 2.3 Déterminer les extremums globaux de la fonction

$$f : (x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \longmapsto x^2 + y^2 + \frac{1}{x + y}.$$

Remarque 2.4 Le raisonnement présenté ci-dessus pourra parfois être employé pour montrer qu'un point critique A est un extremum local pour la fonction f : si la hessienne $\nabla^2 f$ est négative (resp. positive) en tout point d'une boule centrée en A incluse dans $\mathcal{U}$ (mais pas définie-positive au point A sans quoi le résultat est déjà connu), alors f présente un maximum local (resp. minimum local) au point A.

2.2 Sur un compact

On a admis dans un précédent chapitre le résultat théorique suivant, qui donne l'existence d'extremums globaux sous certaines hypothèses sur le domaine de définition.

Théorème 2.5 *Toute fonction $f : \mathcal{K} \longrightarrow \mathbb{R}$ continue sur une partie $\mathcal{K}$ fermée, bornée et non vide de $\mathbb{R}^n$ est bornée et atteint ses bornes (donc admet un minimum et un maximum).*

Une fois acquise l'existence des extremums globaux d'une fonction $f : \mathcal{K} \longrightarrow \mathbb{R}$ d'après le théorème précédent, il s'agit de déterminer les points en lesquels ces extremums globaux sont atteints. On introduit pour cela l'*intérieur* $\mathcal{U}$ de $\mathcal{K}$, qui est le plus grand ouvert (au sens de l'inclusion) inclus dans $\mathcal{K}$ (on peut le définir proprement comme la réunion des ouverts inclus dans $\mathcal{A}$).
Les extremums globaux de f sont alors atteints sur l'intérieur de $\mathcal{K}$ et/ou sur son complémentaire dans $\mathcal{K}$, appelé le *bord* ou la *frontière* de $\mathcal{K}$.

- ➢ Si un extremum de f est atteint sur l'ouvert $\mathcal{U}$, c'est nécessairement en un point critique (en supposant f de classe $\mathscr{C}^1$ sur $\mathcal{U}$). On recherche donc ces points critiques et on calcule la valeur de f en ces points (il n'est pas nécessaire de vérifier s'il s'agit d'extremums locaux !).
- ➢ Le paragraphe suivant fournira des moyens de limiter la recherche des extremums sur le bord de $\mathcal{K}$. En attendant, on observera en pratique sur les cas simples que le bord de $\mathcal{K}$ peut en général être paramétré avec une variable de moins que $\mathcal{K}$. On se trouve donc face au même problème (recherche d'extremum globaux sur un compact) que l'on réduit de la même façon jusqu'à arriver à un bord paramétré par une seule variable ; il est alors possible d'étudier les variations de f sur ce bord et donc d'en déterminer les extremums.

On obtient ainsi les seuls extremums globaux éventuels de f sur $\mathcal{K}$. L'existence de ces extremums globaux étant acquise, il ne reste alors qu'à comparer la valeur de f aux points précédemment obtenus.

Exemple 2.6 Minimum et maximum globaux de la fonction $f : (x, y) \in \mathbb{R}^2 \longmapsto x^2 + y^2 - xy + x + y$ sur l'ensemble $\mathcal{A} = \{(x, y) \in \mathbb{R}^2 : x + y \geqslant -3, x \leqslant 0, y \leqslant 0\}$.

2.3 Position du graphe par rapport à l'hyperplan tangent

La problématique étudiée dans cette section est classique mais ne figure pas explicitement au programme.
On suppose la fonction $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^2$ sur un ouvert $\mathcal{U}$. Dans $\mathcal{U} \times \mathbb{R} \subset \mathbb{R}^{n+1}$, le graphe de f a pour équation $y = f(x_1, \ldots, x_n)$ et son hyperplan tangent au point A :

$$y = f(A) + \sum_{j=1}^{n} \partial_j f(A)(x_j - a_j).$$

C'est donc le signe de l'expression

$$f(x_1, \dots, x_n) - f(\mathrm{A}) - \sum_{j=1}^{n} \partial_j f(\mathrm{A})(x_j - a_j)$$

qui donne la position du graphe par rapport à l'hyperplan tangent. L'expression précédente est égale à $g(\mathrm{X}) - g(\mathrm{A})$ si l'on définit la fonction

$$g : \mathrm{X} \in \mathcal{U} \longmapsto f(\mathrm{X}) - \sum_{j=1}^{n} \partial_j f(\mathrm{A}) x_j = f(\mathrm{X}) - \langle \nabla f(\mathrm{A}), \mathrm{X} \rangle .$$

Il s'agit donc de savoir si la fonction g présente un extremum au point A. Après avoir observé que g présente en A un point critique et que $\nabla^2 g = \nabla^2 f$, on pourra donc s'appuyer sur les résultats précédents, tant locaux que globaux. Par exemple,

Proposition 2.7 *Si l'une des conditions équivalentes suivantes est satisfaite :*
(i) pour tout $\mathrm{H} \neq 0$, $q_{\mathrm{A}}(\mathrm{H}) > 0$ *(resp.* $q_{\mathrm{A}}(\mathrm{H}) < 0$*);*
(ii) les valeurs propres de la hessienne $\nabla^2 f(\mathrm{A})$ *sont toutes strictement positives (resp. strictement négatives),*
alors le graphe de f *est localement au-dessus (resp. en-dessous) de son hyperplan tangent en* A.

Exemple 2.8 Déterminer une équation de l'hyperplan tangent au graphe de la fonction

$$f : (x, y) \in \mathbb{R}^2 \longmapsto x^3 - y^3 + 3xy + x + y$$

au point $\mathrm{A} = (-1, 1)$ puis étudier les positions relatives du graphe et du plan tangent au voisinage de A.

Remarque 2.9 On notera en particulier que si la matrice symétrique $\nabla^2 f(\mathrm{A})$ est positive en tout point d'un ouvert convexe, alors le graphe de f est au-dessus de chacun de ses hyperplans tangents.

3. Recherche d'extremums sous contrainte

3.1 Notion d'extremum sous contrainte

Définition 3.1 *Soient* $\mathcal{C}$ *une partie de* $\mathbb{R}^n$ *telle que* $\mathcal{A} \cap \mathcal{C} \neq \emptyset$ *et* $\mathrm{A} \in \mathcal{A} \cap \mathcal{C}$.
On dit que f *présente un* ***extremum sous la contrainte*** $\boldsymbol{\mathcal{C}}$ *au point* A *si sa restriction* $f|_{\mathcal{A}\cap\mathcal{C}}$ *admet un extremum au point* A.

Remarques 3.2 • On parle parfois d'*optimisation liée* plutôt que d'optimisation sous contrainte, car la contrainte impose des conditions entre les variables.

- Lorsque la condition d'appartenance à $\mathcal{C}$ revient à exprimer certaines variables en fonction des autres, on dit que la contrainte est *explicite*. Dans un tel cas, l'optimisation de f sous la contrainte $\mathcal{C}$ se ramène à un problème d'optimisation libre (par opposition aux problèmes d'optimisation liée) avec moins de variables. Lorsque cette technique peut être mise en oeuvre, c'est souvent la plus efficace.

Exemple 3.3 Extremums de $f : (x, y) \in \mathbb{R}^2 \longmapsto x\mathbf{e}^{x-y}$ sous la contrainte $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : y - x^2 = 0\}$.

Exemples 3.4 La microéconomie fourmille de problèmes d'optimisation sous contrainte. Voici les deux plus classiques, présentés dans un cas simple.

- Étant donnés deux biens Q_1 et Q_2, une fonction d'utilité $\mathrm{U} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ associe à tout panier (q_1, q_2) un indicateur de satisfaction du consommateur. Celui-ci cherche à maximiser $\mathrm{U}(q_1, q_2)$ sous la contrainte $p_1 q_1 + p_2 q_2 = \mathrm{R}$, où p_1, p_2 sont les prix des deux biens et R son revenu.
- Étant donnés deux inputs (travail, matières premières, ...) Q_1 et Q_2 utiles à la production d'un bien, la fonction de production $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ associe à tout panier (q_1, q_2) la quantité de bien produite. Le producteur cherche à minimiser son coût de production $p_1 q_1 + p_2 q_2$ sous la contrainte $f(q_1, q_2) = q$, où p_1, p_2 sont les prix des deux inputs et q la quantité de bien à produire.

3.2 *Cas d'une contrainte linéaire*

DÉFINITION 3.5 *Une contrainte $\mathcal{C} \subset \mathbb{R}^n$ est dite linéaire si elle s'écrit comme l'ensemble des solutions d'un système linéaire donné*

$$\mathcal{C} : \quad \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ \quad \vdots \\ a_{p,1}x_1 + \cdots + a_{p,n}x_n = b_p \end{cases} \tag{3.1}$$

pour des réels $a_{i,j}$, $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant n$, et b_i, $1 \leqslant i \leqslant p$.

Remarques 3.6
- En introduisant, pour tout $i \in [\![1, p]\!]$, la fonction
$$g_i : (x_1, \ldots, x_n) \in \mathbb{R}^n \longmapsto a_{i,1}x_1 + \cdots + a_{i,n}x_n,$$
la contrainte $\mathcal{C}$ est l'intersection des hyperplans affines d'équations $g_i(\mathrm{X}) = b_i$, $1 \leqslant i \leqslant p$.
- Si $\mathcal{H} = \bigcap_i \operatorname{Ker} g_i$ désigne l'ensemble des solutions du système homogène associé et si X_0 est une solution particulière du système complet, i.e. un élément particulier de $\mathcal{C}$, alors $\mathcal{C}$ est l'ensemble des éléments de la forme $\mathrm{X} = \mathrm{X}_0 + \mathrm{H}$ avec $\mathrm{H} \in \mathcal{H}$.
- Réciproquement, on peut démontrer que tout sous-espace affine de $\mathbb{R}^n$ (i.e. tout translaté d'un sous-espace vectoriel) est une contrainte linéaire (admissible pour peu qu'il rencontre $\mathcal{A}$).

Dans la suite de la section, on considère une contrainte linéaire ; on conserve les notations introduites plus haut.

PROPOSITION-DÉFINITION 3.7 *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$ sur un ouvert $\mathcal{U}$. Soit $\mathrm{A} \in \mathcal{U} \cap \mathcal{C}$.*
Si f admet en A un extremum local sous la contrainte $\mathcal{C}$, alors

$$\forall \mathrm{V} \in \mathcal{H}, \quad \partial_{\mathrm{V}} f(\mathrm{A}) = 0,$$

ce qui signifie que $\nabla f(\mathrm{A}) \in \mathcal{H}^{\perp}$. On dit dans ces conditions que A est un point critique de f sous la contrainte $\mathcal{C}$.
La réciproque est fausse.

Remarque 3.8 La condition énoncée étant seulement nécessaire, une étude supplémentaire au voisinage de chaque point critique sous contrainte doit être effectuée pour savoir s'il s'agit d'un extremum sous contrainte ; on pourra pour cela adapter les résultats des deux premiers paragraphes (aucun résultat ne figure au programme). Ainsi, pour une fonction f de classe $\mathscr{C}^2$, en notant q_{A} la forme quadratique associée à la hessienne $\nabla^2 f(\mathrm{A})$, on peut établir les résultats suivants.
- Si A est un point critique sous la contrainte $\mathcal{C}$ et si $q_{\mathrm{A}}(\mathrm{H}) < 0$ (resp. $q_{\mathrm{A}}(\mathrm{H}) > 0$) pour tout $\mathrm{H} \in \mathcal{H} \setminus \{0\}$, alors f présente en A un maximum local (resp. un minimum local) sous la contrainte $\mathcal{C}$ d'après la formule de Taylor-Young.
- Si $\mathcal{U}$ est convexe et $q_{\mathrm{A}}(\mathrm{H}) \leqslant 0$ (resp. $q_{\mathrm{A}}(\mathrm{H}) \geqslant 0$) pour tout $\mathrm{A} \in \mathcal{U} \cap \mathcal{C}$ et tout $\mathrm{H} \in \mathcal{H}$, la formule de Taylor avec reste intégral permet de justifier que f présente un maximum global (resp. un minimum global) sous la contrainte $\mathcal{C}$ en tout point critique sous la contrainte $\mathcal{C}$.

Attention cependant, comme on le verra en TD, le signe de la forme hessienne q_{A} sur $\mathcal{H}$ ne s'exprime pas en général sur les valeurs propres de la matrice $\nabla^2 f(\mathrm{A})$.

Exemple 3.9 Extremums de la fonction

$$f : (x, y, z) \in \mathbb{R}^3 \longmapsto xy + yz + zx$$

sous la contrainte linéaire $x + y + z = 1$.

Remarque 3.10 Chaque équation

$$g_i(\mathrm{X}) = 0 \quad \iff \quad a_{i,1}x_1 + \cdots + a_{i,n}x_n = 0, \qquad 1 \leqslant i \leqslant p,$$

définit un hyperplan normal au vecteur $(a_{i,1}, \ldots, a_{i,n}) = \nabla g_i$ où ∇g_i désigne le gradient de g_i, constant puisque g_i est affine.

Le sous-espace $\mathcal{H}$ étant défini comme intersection de ces hyperplans, son orthogonal est donc engendré par les vecteurs $\nabla g_1, \dots, \nabla g_p$: $\mathcal{H}^\perp = \mathrm{Vect}(\nabla g_1, \dots, \nabla g_p)$.

Corollaire 3.11 *On suppose $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$ sur un ouvert $\mathcal{U}$. Soit $\mathrm{A} \in \mathcal{U} \cap \mathcal{C}$.*
Si f admet en A un extremum sous la contrainte $\mathcal{C}$, alors il existe des réels $\lambda_1, \dots, \lambda_p$ tels que

$$\nabla f(\mathrm{A}) = \sum_{i=1}^{p} \lambda_i \nabla g_i.$$

La réciproque est fausse.

Remarque 3.12 Lorsque la famille $(g_1, \dots, g_p)$ est libre, les réels $\lambda_1, \dots, \lambda_p$ sont uniques ; on les appelle alors *multiplicateurs de Lagrange.*

Exemple 3.13 Déterminer les extremums de la fonction $f : (x, y, z, t) \in \mathbb{R}^4 \longmapsto x^2 + y^2 + z^2 + t^2$ sous la contrainte

$$\begin{cases} x + y + z - t = 3 \\ 2x - y + z + t = -6 \end{cases}.$$

3.3 Cas d'une contrainte définie comme ligne de niveau

On considère dans ce paragraphe une contrainte de la forme $\mathcal{C} = \{\mathrm{X} \in \mathcal{U} : \varphi(\mathrm{X}) = c\}$ où $\varphi : \mathcal{U} \longrightarrow \mathbb{R}$ est une fonction de classe $\mathscr{C}^1$ sur l'ouvert $\mathcal{U}$ et c est un réel donnés. On supposera la contrainte non critique au sens où $\nabla\varphi(\mathrm{X}) \neq 0$ pour tout $\mathrm{X} \in \mathcal{C}$.

Proposition-Définition 3.14 *Soient $f : \mathcal{U} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$ sur l'ouvert $\mathcal{U}$ et $\mathrm{A} \in \mathcal{C}$.*
On suppose que pour tout $\mathrm{X} \in \mathcal{C}$, $\nabla\varphi(\mathrm{X}) \neq 0$.
Si f admet un extremum au point A sous la contrainte $\mathcal{C}$, alors il existe $\lambda \in \mathbb{R}$ tel que $\nabla f(\mathrm{A}) = \lambda\nabla\varphi(\mathrm{A})$.
On dit dans ces conditions que A est un point critique de f sous la contrainte $\mathcal{C}$.
La réciproque est fausse.

Remarque 3.15 En pratique, pour déterminer les extremums de f sous la contrainte $\mathcal{C} : \varphi(\mathrm{X}) = c$,

- on commence par résoudre le système
$$\begin{cases} \varphi(\mathrm{X}) = c \\ \nabla f(\mathrm{X}) = \lambda\nabla\varphi(\mathrm{X}) \end{cases} \tag{3.2}$$
d'inconnue $(\mathrm{X}, \lambda) \in \mathcal{U} \times \mathbb{R}$ pour déterminer les points critiques de f sous la contrainte $\mathcal{C}$;
- puis réciproquement, pour déterminer si un point critique A de f sous la contrainte $\mathcal{C}$ est un extremum pour f, on étudie le signe de l'expression $f(\mathrm{A}+\mathrm{H})-f(\mathrm{A})$ où H est soumis à la contrainte $\varphi(\mathrm{A}+\mathrm{H}) = c$, en utilisant au besoin un développement limité.

Remarque 3.16 Pour tout $\lambda \in \mathbb{R}$, on note $\mathcal{L}_\lambda = \{\mathrm{X} \in \mathcal{U} : f(\mathrm{X}) = \lambda\}$ la ligne de niveau λ de f.

- ➢ Les valeurs prises par la fonction f sur la contrainte $\mathcal{C}$ sont les réels λ pour lesquels la ligne de niveau $\mathcal{L}_\lambda$ de f rencontre la contrainte $\mathcal{C}$, i.e. pour lesquels $\mathcal{L}_\lambda \cap \mathcal{C} \neq \emptyset$.
- ➢ La condition pour que le point $\mathrm{A} \in \mathcal{C}$ soit critique sous la contrainte $\mathcal{C}$ signifie, s'il n'est pas déjà critique au sens usuel ($\nabla f(\mathrm{A}) = 0$), que la ligne de niveau $\mathcal{L}_{f(\mathrm{A})}$ de f passant par le point A est tangente à la contrainte $\mathcal{C}$ au point A.

Souvent, sur les cas simples, la superposition des lignes de niveau de f avec la contrainte $\mathcal{C}$ permet de se forger une intuition de la situation : elle permet de détecter les points critiques sous la contrainte $\mathcal{C}$ et de savoir s'il s'agit d'extremums.

Remarque 3.17 Ce point n'est pas explicitement au programme. On s'intéresse à une contrainte non critique $\mathcal{C}$ de la forme $\varphi(\mathrm{X}) = 0$ (si $c \neq 0$, on s'y ramène facilement en remplaçant φ par $\psi = \varphi - c$).

- La résolution du système (3.2) revient à déterminer les points critiques de la fonction

$$\mathrm{L} : (x_1, \ldots, x_n, \lambda) \in \mathcal{U} \times \mathbb{R} \longmapsto f(x_1, \ldots, x_n) - \lambda\varphi(x_1, \ldots, x_n)$$

appelée *fonction lagrangienne* ou *lagrangien*. Étant donné un tel point critique (A, λ), si l'on arrive à établir que la fonction $\mathrm{X} \longmapsto \mathrm{L}(\mathrm{X}, \lambda)$ admet un extremum en A, alors f présente en A un extremum sous la contrainte $\mathcal{C}$, mais la réciproque est fausse.

- Dans le cas de fonctions de $n = 2$ variables, si f et φ sont de classe $\mathscr{C}^2$ sur $\mathcal{U}$, le lagrangien s'écrit

$$\mathrm{L} : (x, y, \lambda) \in \mathcal{U} \times \mathbb{R} \longmapsto f(x, y) - \lambda\varphi(x, y).$$

Étant donné un point critique (a, b, λ) de L, en notant

$$\begin{pmatrix} r & s \\ s & t \end{pmatrix} = \begin{pmatrix} \partial_1^2 \mathrm{L}(a, b, \lambda) & \partial_{1,2}^2 \mathrm{L}(a, b, \lambda) \\ \partial_{1,2}^2 \mathrm{L}(a, b, \lambda) & \partial_2^2 \mathrm{L}(a, b, \lambda) \end{pmatrix}$$

la matrice hessienne en $\mathrm{A} = (a, b)$ de la fonction $(x, y) \longmapsto \mathrm{L}(x, y, \lambda)$, on a l'alternative suivante :

  ➢ si $rt - s^2 > 0$, alors f présente un extremum en A sous la contrainte $\mathcal{C}$, plus précisément un maximum si $r < 0$ et un minimum si $r > 0$;

  ➢ le cas $rt - s^2 \leqslant 0$ est indéterminé.

Exemple 3.18 Déterminer les extremums de la fonction $f : (x, y) \in \mathbb{R}^2 \longmapsto x + y$ sous la contrainte $x^2 + y^2 = 1$.

3.4 Retour sur l'optimisation sur un compact

Soient $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ sur $\mathbb{R}^n$ et $c \in \mathbb{R}$ tels que :

➢ la partie $\mathcal{K} = \{\mathrm{X} \in \mathbb{R}^n : \varphi(\mathrm{X}) \leqslant c\}$, automatiquement fermée, soit bornée et non vide ;

➢ la contrainte $\mathcal{C} = \{\mathrm{X} \in \mathbb{R}^n : \varphi(\mathrm{X}) = c\}$ soit non critique.

Étant donnée une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$, on se trouve alors en situation d'appliquer la méthode présentée plus tôt. La fonction f admet sur $\mathcal{K}$ des extremums globaux. S'ils sont atteints sur l'ouvert $\mathcal{U} = \{\mathrm{X} \in \mathbb{R}^n : \varphi(\mathrm{X}) < c\}$, c'est nécessairement en un point critique, s'ils sont atteints sur $\mathcal{C}$, c'est nécessairement en un point critique sous la contrainte $\mathcal{C}$.

Exemple 3.19 Déterminer les extremums de la fonction

$$f : (x, y) \in \mathbb{R}^2 \longmapsto x + y + \sqrt{1 - x^2 - y^2}$$

sous la contrainte $x^2 + y^2 \leqslant 1$.