### IV.3 Optimisation sous contrainte

Soit $f : \Omega \to \mathbb{R}$, et $X \subset D$. Optimiser $f$ sous la contrainte $X$ consiste à étudier les extrema de $f$ en se restreignant à $X$ (la contrainte). Il s'agit donc d'étudier les extrema d'une restriction. Nous commençons par une CN similaire à celle qu'on a trouvé dans le cas global (qui d'ailleurs peut être vu comme conséquence de celle-ci).

**Proposition IV.10** (CN du premier ordre pour une restriction)**.** *Soit $\Omega$ un ouvert de $E$, $X$ une partie de $\Omega$, et $f : \Omega \to \mathbb{R}$. Si $f|_X$ admet un extremum local en $a \in X$, alors*

$$\forall h \in T_a X, \quad df(a) \cdot h = 0.$$

*En d'autres termes, $df(a)$ s'annule en tout vecteur tangent à $X$, c'est-à-dire $T_a X \subset \operatorname{Ker} df(a)$.*

Éléments de preuve. Dériver $f(\alpha(t))$, où $\alpha : [-\varepsilon, \varepsilon] \to X$ vérifie $\alpha(0) = a$. □

**Remarque IV.11.** *En prenant $X = \mathring{D}$, puisque $X$ est un voisinage de $a$, $T_a X = E$, et on retrouve donc le théorème 16.4.5.*

2. *Si $E$ est euclidien, la proposition 16.4.15 affirme que si $f|_X$ admet un extremum local en $a$, $\nabla f(a) \perp T_a X$.*

Si l'ensemble $X$ est défini par une équation $g(x) = 0$, on peut combiner le résultat précédent et le théorème 16.3.21.

**Théorème IV.12** (Optimisation sous contrainte)**.** *Soit $f, g : \Omega \to \mathbb{R}$ de classe $C^1$, et $X = \{x \in \Omega, g(x) = 0\}$. Si $f|_X$ admet un extremum local en $a \in X$ tel que $dg(a) \neq 0$, alors $df(a)$ et $dg(a)$ sont colinéaires.*

## Exercices d'optimisation sous contrainte

**Exemple IV.13.** *Déterminer les extrema de $f(x, y) = x^2 + xy - 2y^2$ sur le cercle unité $g(x, y) = x^2 + y^2 - 1 = 0$.*

**Solution**
**Existence.** Le cercle unité est compact et $f$ est continue, donc $f$ admet un minimum et un maximum sur ce cercle.
**Multiplicateurs de Lagrange.** On résout :

$$\begin{cases} \nabla f = \lambda \nabla g \\ g(x, y) = 0 \end{cases}$$

Les gradients :

$$\nabla f = (2x + y,\ x - 4y), \quad \nabla g = (2x,\ 2y)$$

D'où le système linéaire en $x, y$ :

$$\begin{cases} 2x + y = 2\lambda x \\ x - 4y = 2\lambda y \end{cases}$$

Soit :

$$\begin{cases} (2 - 2\lambda)x + y = 0 \\ x + (-4 - 2\lambda)y = 0 \end{cases}$$

**Recherche de $\lambda$.** Pour que ce système homogène admette une solution non nulle $(x, y) \neq (0, 0)$, il faut que son déterminant soit nul :

$$\det \begin{pmatrix} 2 - 2\lambda & 1 \\ 1 & -4 - 2\lambda \end{pmatrix} = 0$$

Calculons :

$$(2-2\lambda)(-4-2\lambda)-1=0$$
$$-8-4\lambda+8\lambda+4\lambda^2-1=0$$
$$4\lambda^2+4\lambda-9=0$$
$$\lambda^2+\lambda-\frac{9}{4}=0$$

Le discriminant est $\Delta=1+9=10$, donc :

$$\lambda=\frac{-1\pm\sqrt{10}}{2}$$

**Recherche de** $(x,y)$**.** Pour chaque $\lambda$, le système devient une seule équation (les deux sont proportionnelles).
Pour $\lambda=\frac{-1+\sqrt{10}}{2}$ : La première équation : $(2-2\lambda)x+y=0$. Calculons $2-2\lambda=2-(-1+\sqrt{10})=3-\sqrt{10}$. Donc $(3-\sqrt{10})x+y=0$, soit $y=(\sqrt{10}-3)x$.
Pour $\lambda=\frac{-1-\sqrt{10}}{2}$ : $2-2\lambda=2-(-1-\sqrt{10})=3+\sqrt{10}$. Donc $(3+\sqrt{10})x+y=0$, soit $y=-(3+\sqrt{10})x$.
**Contrainte** $g(x,y)=0$ **i.e.** $x^2+y^2=1$**.** Pour le premier cas : $y=(\sqrt{10}-3)x$. Alors :

$$x^2+(\sqrt{10}-3)^2x^2=x^2[1+(\sqrt{10}-3)^2]=1$$

Or $(\sqrt{10}-3)^2=10-6\sqrt{10}+9=19-6\sqrt{10}$. Donc :

$$1+19-6\sqrt{10}=20-6\sqrt{10}$$

Ainsi :

$$x^2=\frac{1}{20-6\sqrt{10}},\quad x=\pm\frac{1}{\sqrt{20-6\sqrt{10}}}$$

et $y=(\sqrt{10}-3)x=\pm\frac{\sqrt{10}-3}{\sqrt{20-6\sqrt{10}}}$.
Pour le second cas : $y=-(3+\sqrt{10})x$. Alors :

$$x^2+(3+\sqrt{10})^2x^2=x^2[1+(3+\sqrt{10})^2]=1$$

Or $(3+\sqrt{10})^2=19+6\sqrt{10}$. Donc :

$$1+19+6\sqrt{10}=20+6\sqrt{10}$$

Ainsi :

$$x^2=\frac{1}{20+6\sqrt{10}},\quad x=\pm\frac{1}{\sqrt{20+6\sqrt{10}}}$$

et $y=-(3+\sqrt{10})x=\mp\frac{3+\sqrt{10}}{\sqrt{20+6\sqrt{10}}}$.
**Valeurs de** $f$**.** Calculons $f=x^2+xy-2y^2$ pour chaque cas.
Pour le premier cas ($\lambda=\frac{-1+\sqrt{10}}{2}$), on trouve $f=\frac{-1+\sqrt{10}}{2}$.
Pour le second cas ($\lambda=\frac{-1-\sqrt{10}}{2}$), on trouve $f=\frac{-1-\sqrt{10}}{2}$.
**Conclusion.** Les extremums sont :

$$\boxed{\max f=\frac{-1+\sqrt{10}}{2},\quad \min f=\frac{-1-\sqrt{10}}{2}}$$

Les points où ils sont atteints sont respectivement $(x,y)=\left(\pm\frac{1}{\sqrt{20-6\sqrt{10}}},\ \pm\frac{\sqrt{10}-3}{\sqrt{20-6\sqrt{10}}}\right)$ pour le maximum et $(x,y)=\left(\pm\frac{1}{\sqrt{20+6\sqrt{10}}},\ \mp\frac{3+\sqrt{10}}{\sqrt{20+6\sqrt{10}}}\right)$ pour le minimum.

**Exemple IV.14.** *Montrer l'inégalité arithmético-géométrique :*

$$\frac{x_1+\cdots+x_n}{n} \geq \sqrt[n]{x_1\cdots x_n} \quad (x_i>0)$$

**Solution**
On maximise $f(x_1,\ldots,x_n)=\sqrt[n]{x_1\cdots x_n}$ sous la contrainte $g(x_1,\ldots,x_n)=x_1+\cdots+x_n-n=0$, avec $x_i>0$.
**Existence.** Soit $K=\{x\in\mathbb{R}^n \mid x_i\geq 0 \text{ et } g(x)=0\}$. L'ensemble $K$ est fermé (image réciproque du fermé $\{0\}$ par l'application continue $g$) et borné (car $x_i\geq 0$ et $\sum x_i=n$ implique $0\leq x_i\leq n$). $K$ est compact car c'est un fermé bonrné en dimension finie. La fonction $f$ est continue sur $K$ (prolongée par continuité en posant $f=0$ dès qu'un $x_i=0$). D'après le théorème des extremums, $f$ atteint son maximum sur $K$. Puisque $f(x)>0$ lorsque tous les $x_i>0$ et que $f(x)=0$ lorsqu'un $x_i=0$, le maximum est strictement positif et n'est donc pas atteint sur la partie où un $x_i$ s'annule. Par conséquent, le maximum est atteint en un point où tous les $x_i>0$.
**Recherche du maximum par la méthode des extrema liés.** Au point où le maximum est atteint, le gradient de $f$ est colinéaire au gradient de $g$. Il existe donc un réel $\lambda$ tel que :

$$\nabla f=\lambda\nabla g$$

Calculons les dérivées partielles. En écrivant $f=(x_1\cdots x_n)^{1/n}$, on a :

$$\frac{\partial f}{\partial x_i}=\frac{1}{n}(x_1\cdots x_n)^{1/n}\cdot\frac{1}{x_i}=\frac{f}{nx_i}$$

Comme $\frac{\partial g}{\partial x_i}=1$, la condition $\nabla f=\lambda\nabla g$ donne pour tout $i$ :

$$\frac{f}{nx_i}=\lambda\cdot 1 \quad\Rightarrow\quad \frac{f}{nx_i}=\lambda$$

Cette égalité étant vraie pour tout $i$, on en déduit :

$$\frac{f}{nx_1}=\frac{f}{nx_2}=\cdots=\frac{f}{nx_n}$$

Puisque $f>0$ au point considéré, on simplifie par $\frac{f}{n}$ et on obtient :

$$\frac{1}{x_1}=\frac{1}{x_2}=\cdots=\frac{1}{x_n}$$

Ainsi, tous les $x_i$ sont égaux : $x_1=x_2=\cdots=x_n$. Notons $a$ cette valeur commune. La contrainte $g(x)=0$ s'écrit :

$$a+a+\cdots+a=n \quad\Rightarrow\quad na=n \quad\Rightarrow\quad a=1$$

Le seul point vérifiant la condition est donc $(1,1,\ldots,1)$. En ce point, on a :

$$f(1,\ldots,1)=\sqrt[n]{1\cdots 1}=1$$

C'est donc la valeur maximale de $f$ sous la contrainte.
On a ainsi montré que pour tout $x_1,\ldots,x_n>0$ vérifiant $x_1+\cdots+x_n=n$ :

$$\sqrt[n]{x_1\cdots x_n}\leq 1$$

**Passage au cas général.** Soient maintenant $y_1,\ldots,y_n>0$ quelconques. Posons :

$$x_i=\frac{ny_i}{\sum_{j=1}^n y_j}$$

Alors on vérifie que $x_1 + \cdots + x_n = n$ et $x_i > 0$. En appliquant l'inégalité obtenue à ces $x_i$ :

$$\sqrt[n]{x_1 \cdots x_n} \leq 1$$

En remplaçant $x_i$ par leur expression :

$$\sqrt[n]{\frac{ny_1}{\sum y_j} \cdot \frac{ny_2}{\sum y_j} \cdots \frac{ny_n}{\sum y_j}} \leq 1$$

$$\sqrt[n]{\frac{n^n y_1 y_2 \cdots y_n}{(\sum y_j)^n}} \leq 1$$

$$\frac{n \sqrt[n]{y_1 y_2 \cdots y_n}}{\sum_{j=1}^n y_j} \leq 1$$

Ce qui équivaut à :

$$\sqrt[n]{y_1 y_2 \cdots y_n} \leq \frac{\sum_{j=1}^n y_j}{n}$$

**Étude du cas d'égalité.** Analysons maintenant pour quels $y_i > 0$ on obtient l'égalité :

$$\sqrt[n]{y_1 y_2 \cdots y_n} = \frac{y_1 + \cdots + y_n}{n}$$

L'égalité dans l'inégalité finale provient de l'égalité dans l'inégalité intermédiaire $\sqrt[n]{x_1 \cdots x_n} \leq 1$ appliquée aux $x_i = \frac{ny_i}{\sum y_j}$. En effet, on a :

$$\frac{n \sqrt[n]{y_1 \cdots y_n}}{\sum y_j} = 1 \iff \sqrt[n]{x_1 \cdots x_n} = 1$$

Or, d'après l'étude du problème de maximisation, la valeur maximale $1$ n'est atteinte que lorsque tous les $x_i$ sont égaux entre eux, c'est-à-dire lorsque $x_1 = x_2 = \cdots = x_n$. En effet, la condition $\nabla f = \lambda \nabla g$ nous a conduit à l'égalité des $x_i$, et le point $(1, \ldots, 1)$ est l'unique point vérifiant à la fois cette condition et la contrainte.
Ainsi, l'égalité $\sqrt[n]{x_1 \cdots x_n} = 1$ équivaut à $x_1 = x_2 = \cdots = x_n$. En revenant à la définition des $x_i$ en fonction des $y_i$ :

$$\frac{ny_1}{\sum_{j=1}^n y_j} = \frac{ny_2}{\sum_{j=1}^n y_j} = \cdots = \frac{ny_n}{\sum_{j=1}^n y_j}$$

En multipliant chaque égalité par $\frac{\sum y_j}{n} > 0$, on obtient :

$$y_1 = y_2 = \cdots = y_n$$

Réciproquement, si tous les $y_i$ sont égaux à une même valeur $y > 0$, alors :

$$\frac{y_1 + \cdots + y_n}{n} = \frac{ny}{n} = y \quad \text{et} \quad \sqrt[n]{y_1 \cdots y_n} = \sqrt[n]{y^n} = y$$

L'égalité est bien vérifiée.
On peut donc conclure : l'égalité dans l'inégalité arithmético-géométrique a lieu si et seulement si tous les nombres $y_1, \ldots, y_n$ sont égaux.

## IV.4 Étude à l'ordre 2

Si $f : I \to \mathbb{R}$ est une fonction définie sur un intervalle ouvert de $\mathbb{R}$, et si $a \in I$ est un point critique de $f$, on peut déterminer s'il s'agit effectivement d'un extremum par un argument de convexité locale, ce qui, en cas de régularité suffisante, se ramène à l'étude du signe de $f''(a)$. Nous recherchons une condition similaire lorsque $f$ est une fonction définie sur un ouvert de $\mathbb{R}^n$, à l'aide de la matrice hessienne $H_f(a)$.

**Proposition IV.15** (CN du second ordre)**.** *Soit $f : \Omega \to \mathbb{R}$ (où $\Omega$ est un ouvert de $\mathbb{R}^n$) de classe $C^2$. Si $f$ admet un minimum local (resp. maximum local) en $a \in \Omega$, alors $a$ est un point critique de $f$ et $H_f(a) \in S_+(\mathbb{R})$ (resp. $-H_f(a) \in S_+(\mathbb{R})$).*

Éléments de preuve. D'après la formule de Taylor-Young, le signe de $f(a+h) - f(a)$ est celui de $h^T H_f(a) h$. □

**Théorème IV.16** (CS du second ordre)**.** *Soit $f : \Omega \to \mathbb{R}$ (où $\Omega$ est un ouvert de $\mathbb{R}^n$) de classe $C^2$, et soit $a \in \Omega$. Si $a$ est un point critique de $f$ et si $H_f(a) \in S_{++}(\mathbb{R})$ (resp. $-H_f(a) \in S_{++}(\mathbb{R})$), alors $f$ admet un minimum local strict (resp. maximum local strict) en $a$.*

Éléments de preuve. Toujours le contrôle du signe par Taylor-Young, en travaillant avec une norme adaptée à $H_f(a)$. □

**Remarque IV.17** (Avertissement)**.** *La CN du second ordre n'est pas une CS.*

- *La CS du second ordre n'est pas une CN.*

**Exemple IV.18.** *Étudier les CN et CS du second ordre pour les fonctions suivantes, en $0$.*

1. $f : (x, y) \mapsto (x^2 + y^2)^2$.
2. $g : (x, y) \mapsto x^3 y^3$.

**Méthode IV.19** (Étude d'un point critique)**.** *Pour savoir si un point critique correspond à un extremum local, on étudiera le signe de la hessienne,*

- *soit directement par mises sous forme canonique successives de la forme quadratique $h \mapsto h^T H_f(a) h$,*
- *soit par étude des valeurs propres de $H_f(a)$.*

*On peut alors conclure :*

- *Si $H_f(a)$ n'est pas positive (i.e. s'il existe une valeur propre strictement négative, ou s'il existe $h$ tel que $h^T H_f(a) h < 0$), alors $f$ n'admet pas d'extremum local en $a$.*
- *Si $H_f(a)$ est strictement positive (si $\mathrm{Spec}(H_f(a)) \subset \mathbb{R}_+^*$), alors $f$ admet un minimum local.*

*On ne peut pas conclure de cette façon si ces conditions ou les conditions symétriques (pour un maximum) ne sont pas satisfaites, à savoir dans le cas où $0 \in \mathrm{Spec}(H_f(a))$.*

**Corollaire IV.20** (Explicitation pour $n = 2$)**.** *Soit $\Omega$ un ouvert de $\mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ de classe $C^2$, et $a$ un point critique de $f$. Alors :*

- *Si $\det(H_f(a)) > 0$ et $\mathrm{tr}(H_f(a)) > 0$, alors $f$ admet un minimum local strict en $a$.*
- *Si $\det(H_f(a)) > 0$ et $\mathrm{tr}(H_f(a)) < 0$, alors $f$ admet un maximum local strict en $a$.*
- *si $\det(H_f(a)) < 0$, alors $f$ n'admet pas d'extremum local en $a$.*