Ch 2. Extrema -

§1 - Introduction. Notations

Soit $f : A \subset \mathbb{R}^n \to \mathbb{R}$, $n \geq 1$ A non nécessair^t ouvert.

. Déf. 1. soit $x_0 \in A$. On dit que x_0 est un point de minimum global de f sur A si $f(x_0) = \min_{x \in A} f(x)$, i.e. si $\forall x \in A$, $f(x) \geq f(x_0)$, et $x_0 \in A$

. soit $x_0 \in A$. On dit que x_0 est un point de minimum local de f sur A si $\exists r > 0$ tel que $f(x_0) = \min_{\substack{x \in A \\ \text{et } \|x - x_0\| \leq r}} f(x)$

. on définit de même un point de maximum (global, local)

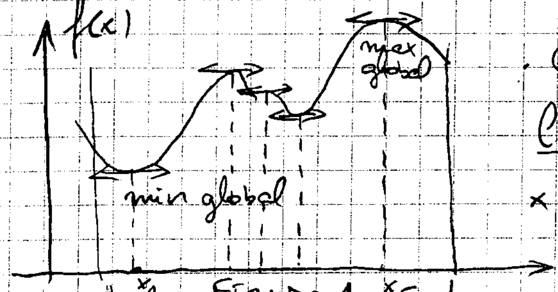


Figure 1

Ex. $A = [a, b[$ ...

. On rappelle, cf Ch 1, p 2, que l'intérieur de A est l'ensemble $\mathring{A}$ des $x \in A$ tq $\exists r > 0$ tq $B(x, r) \subset A$. Il se peut que $\mathring{A}$ soit vide.

. On dit que $A \subset \mathbb{R}^n$ est compact si et seulement si : (i) A est fermé : si une suite $(x_k)_{k \in \mathbb{N}}$ d'éléments de A converge dans $\mathbb{R}^n$ vers une limite x quand $k \to +\infty$, alors $x \in A$: l'adhérence $\overline{A}$ de A est égale à A

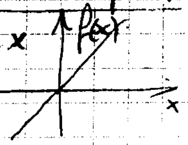
(ii) et si A est borné : $\exists R > 0$ tq $A \subset B(0, R)$.

. Thm 1 ****

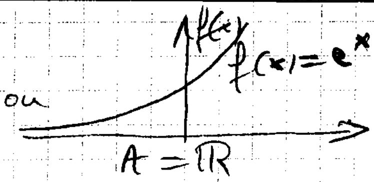
Soit $f : A \subset \mathbb{R}^n \to \mathbb{R}$. On suppose (H) f continue et A compact. Alors f atteint sa borne $\left\{\begin{matrix}\text{inférieure } m \\ \text{supérieure } M\end{matrix}\right\}$ sur A en (au moins) un point $\left\{\begin{matrix} x_m \\ x_M \end{matrix}\right\}$. On a donc : $f(x_m) = \min_{x \in A} f(x) = m$; $f(x_M) = \max_{x \in A} f(x) = M$.

. Rem. FAUX si A non compact

Ex. $A = \mathbb{R}$, $f(x) = x$

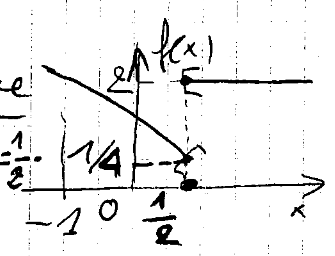


ou



- - - -

ou si f discontinue



ex. $f(\frac{1}{2}) = 2 > \min f(x) = \frac{1}{2}$.

§2_ Conditions d'ordre 1 :

Thm 2: Soit $f : A \to \mathbb{R}$. On suppose que $\mathring{A} = \mathrm{Int}(A) \neq \emptyset$, que f admet un minimum local en $a \in \mathring{A}$, et que f est différentiable en a. Alors $\nabla f(a) = 0$, i.e $\frac{\partial f}{\partial x_i}(a) = 0$, $\forall i = 1, \dots, n$.

Dém. Soit $x = a + t\vec{e}_i$, $\vec{e}_i = (0, \dots, 0, \underset{i^{\text{ème}}}{1}, 0, \dots 0)$

Pour $|t| < r$, $x \in A$, car $a \in \mathring{A}$, et $f_a : \{t \mapsto f(a + t\vec{e}_i)\}$ admet un point de minimum local pour $t = 0$. Donc $(f_a)'(t)\big|_{t=0} = 0$, i.e. $Df_a(\vec{e}_i) = \frac{\partial f}{\partial x_i}(a) \cdot 1 = 0$.

⚠ Réciproque fausse : cf Figure 1, cf $\{x \mapsto x^3, A = \mathbb{R}\}$...

Déf 2 On dit que a est un point critique pour $f \in C^1$ si $Df_a = \nabla f(a) = 0$.

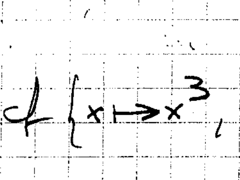
Etude pratique pour la recherche des extrema de f, $f \in C^1$.

(i) * montrer l'$\exists$ d'extrema, si c'est possible : sous l'hypothèse (H) du Thm 1

(ii) * chercher les points critiques de f sur $\mathring{A}$ [même sans l'hyp. (H)]

(iii) * étudier f sur la frontière $\mathrm{Fr}(A) = A \setminus \mathring{A}$, e.g. en paramétrant les morceaux de la frontière de A, cf TD

Ex $A = \{x \geq 0, y \geq 0, x + y \leq 1\} \subset \mathbb{R}^2$



(iv) * si e.g. $f|_{F_1}$ présente un maximum local sur F_1 pour $x = \frac{1}{2}$, en ce point $\{x \mapsto f(x, y=0)\}$ est maximale et $\frac{1}{2} \in \mathring{\overbrace{[0,1]}}$, donc $\frac{\partial f}{\partial x}(\frac{1}{2}, 0) = 0$ ... sinon étudier →

f en $(0,0)$ et en $(1,0)$ ...

* finalement, comparer les valeurs de f en ces différents candidats, cf TD.

§3 - Conditions d'ordre 2 :

Soit $f : \Omega \to \mathbb{R}$, avec Ω ouvert $\subset \mathbb{R}^n$. Si f admet un (point de) minimum local en $a \in \mathring{\Omega} = \Omega$, alors a est un point critique pour f. Réciproque fausse.

Thm 3 : Soit $f : \Omega$ ouvert $\to \mathbb{R}$, $f \in C^2(\Omega ; \mathbb{R})$ et a un point critique de f : $\nabla f(a) = 0$. Alors

(i) Si a est un point de minimum local de f, alors la matrice Hessienne $H(f)(a) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a), & \cdots, & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a), & \cdots, & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$

est (symétrique) positive :

$\forall x \in \mathbb{R}^n$, ${}^t x\, H(f)(a) x := \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, x_i x_j \geq 0$

(ii) si a est un point de maximum local de f, alors $H(f)(a)$ est négative :

$\forall x \in \mathbb{R}^n$, ${}^t x\, H(f)(a) x \leq 0$.

Rem : penser à $\{x \mapsto +x^2\}$ ou $\{x \mapsto -x^2\}$ si $n = 1$

⚠ Réciproque fausse : ex $f(x,y) = \frac{x^2}{2} - \frac{y^4}{4}$

Dém (idée) : si f admet un point de minimum local en a, alors $\nabla f_a = \nabla f(a) = 0$, et comme $f \in C^2$, elle admet en a le DL_2 suivant : $f(x) = f(a) + \langle 0, x-a \rangle + \frac{1}{2!} \boxed{{}^t(x-a) H(f)(a)(x-a)} + \|x-a\|^2 \theta(x)$, où $\theta(x) \underset{(x \to a)}{\to} 0$.

Donc $f(x) - f(a) \geq 0$. Donc le terme quadratique reste ≥ 0 ...

Thm 4 : Réciproque fausse si $H(f)(a)$ positive, mais VRAIE si $H(f)(a)$ définie-positive : $\forall x \neq 0$, ${}^t x\, H(f)(a) x > 0$

§4. Rappels d'analyse matricielle :

. Changements de base : Soit $\vec{x} = \sum_{i=1}^{n} x_i \vec{e_i} \in \mathbb{R}^n$, où $\vec{e_i} = (0, \dots, 0, \underset{i^{\text{ème}}}{1}, 0, \dots, 0)$ et soit $(\vec{e}'_i)_{1 \le i \le n}$ une autre base de $\mathbb{R}^n$. Soit P la matrice de changement de base :

$$P = \begin{array}{c} \begin{array}{ccc} \vec{e}'_1 & \dots & \vec{e}'_n \end{array} \\ \begin{pmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{pmatrix} \end{array} \begin{array}{c} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{array}$$

. Alors $\underset{\uparrow}{\vec{e}'_j} = \sum_{i=1}^{n} p_{i,j} \underset{\uparrow}{\vec{e}_i}$.

nouveaux $\leftarrow$ anciens vecteurs de base

. Soit $\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i = \sum_{j=1}^{n} x'_j \vec{e}'_j$. On a alors $x = Px'$, où

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{pmatrix} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

anciennes composantes $\leftarrow$ nouvelles composantes

. Si $\mathcal{A}$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ (un endomorphisme de $\mathbb{R}^n$), alors dans la base $(\vec{e_i})_{1 \le i \le n}$ $\vec{y} = \mathcal{A}(\vec{x})$ est repéré par $y = Ax$, où :

$$\mathcal{A}(\vec{e}_j) = \sum_{i=1}^{n} a_{i,j} \vec{e}_i \qquad A = \begin{array}{c} \begin{array}{ccc} \mathcal{A}(\vec{e}_1) & \dots & \mathcal{A}(\vec{e}_n) \end{array} \\ \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} \end{array} \begin{array}{c} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{array}$$

. Dans la "nouvelle" base $(\vec{e}'_1, \dots, \vec{e}'_n)$, la <u>même</u> application linéaire $\mathcal{A}$ est repérée par $\underline{A' = P^{-1} A P}$. En effet, $\underset{\underset{P y'}{\parallel}}{y} = Ax = APx'$, donc $\boxed{y' = P^{-1} A P x'}$

. En particulier, <u>si la nouvelle base</u> $(\vec{e}'_i)_{1 \le i \le n}$ <u>est orthonormale</u>, i.e. si $P^{-1} = {}^tP$ (la matrice <u>P est orthogonale</u>), alors on a <u>aussi</u>

$$\underline{A' = P^{-1} A P = {}^tP A P}$$

Diagonalisation d'une matrice A :

Définition 3 : on dit qu'une matrice carrée $n\times n$: A est diagonalisable s'il existe une base (de $\mathbb{R}^n$) formée de vecteurs propres de A, i.e de vecteurs propres de l'endomorphisme associé $\mathcal{A}$: $\mathcal{A}\vec{w_i} = \lambda_i \vec{w_i}$, $\lambda_i \in \mathbb{R}$, $\vec{w_i} \in \mathbb{R}^n$

Thm 5 : ****

(i) En particulier, si A est symétrique réelle, alors elle est diagonalisable : il existe une base ortho-normale de $\mathbb{R}^n$ formée de vecteurs propres de A, et toutes les valeurs propres de A sont réelles.

(ii) Dans une telle base de vecteurs propres $\{\vec{w_1}, \dots, \vec{w_n}\}$, l'endomorphisme $\mathcal{A}$ associé à A dans la base $\{\vec{e_1}, \dots, \vec{e_n}\}$ est donc repéré par

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \text{ et } P^{-1} = {}^tP.$$

(iii) On associe ainsi à une matrice A (symétrique) réelle la forme bi-linéaire $q(\vec{x},\vec{y}) = {}^tx\,A\,y = \sum_{i,j=1}^{n} a_{i,j}x_i y_j$, et la forme quadratique associée : $Q(\vec{x}) := q(\vec{x},\vec{x}) = {}^tx\,Ax$.

Dans la base $\perp$normale de vecteurs propres $\{\vec{w_1}, \dots, \vec{w_n}\}$, on a donc $\vec{x} = \sum_{i=1}^{n} x_i \vec{e_i} = \sum_{j=1}^{n} x'_j \vec{w_j}$, avec $x = Px' \Leftrightarrow x' = P^{-1}x = {}^tPx$. Donc la forme quadratique Q est repérée par

$$Q(\vec{x}) = q(\vec{x},\vec{x}) = {}^tx\,Ax = {}^tx'\,\overbrace{{}^tPAP}^{=D}\,x' = {}^tx' D x' = \sum_{i=1}^{n} \lambda_i (x'_i)^2.$$

(iv) On dit que la forme quadratique Q associée à A est définie-positive si $\forall \vec{x} \in \mathbb{R}^n$, $Q(\vec{x}) = {}^tx\,Ax \geq 0$, et $Q(\vec{x}) = 0$ ssi $\vec{x} = \vec{0}$.

Cette condition équivaut à : $\forall i = 1, \dots, n$, $\lambda_i > 0$.

Application pratique : La matrice Hessienne d'une application $f : \mathbb{R}^n \to \mathbb{R}$ est ($\frac{1}{2}$ définie-) positive si toutes ses valeurs propres — qui sont réelles (matrice symétrique) sont ≥ 0, et définie-positive si ses valeurs propres sont toutes > 0). Ceci permet de caractériser les points critiques de f, i.e. les points $a \in \mathbb{R}^n$ où $\nabla f(a) = 0$.

Rem. (i) Pour une matrice 2×2, les valeurs propres de A sont racines de $\det(A - \lambda I) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A = 0$, où $\operatorname{tr} A = a_{1,1} + a_{2,2} = \lambda_1 + \lambda_2$ et $\det A = \lambda_1 \lambda_2$ ...

(ii) Dans le cas où A n'est pas symétrique, elle peut ne pas être diagonalisable, e.g. si elle a un bloc de Jordan ex. $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Elle est par contre diagonalisable si toutes ses valeurs propres sont distinctes, mais elle peut alors avoir des valeurs propres ou des vecteurs propres complexes...

§5. Optimisation avec contraintes : indications, extrema liés

Exemple : $(\mathcal{P})$ $\left\{ \begin{array}{l} \text{Trouver } (x_0, y_0) \in (C) := \{(x,y) \in \mathbb{R}^2 ; \underbrace{x^2 + y^2 - R^2}_{:= g(x,y)} = 0\} \\ \text{tq } f(x_0, y_0) = \inf\{f(x,y) ; g(x,y) = 0\} \end{array} \right.$

. Rem. Ici, la courbe (C) est fermée (image réciproque de $\{0\}$, qui est fermé dans $\mathbb{R}$, par l'application continue g de $\mathbb{R}^2$ dans $\mathbb{R}$), et bornée, donc compacte. Donc, en supposant f continue, le pb $(\mathcal{P})$ admet au moins une solution (x_0, y_0). $\square$

. même si on n'est pas sûr de l'existence de (x_0, y_0), on peut chercher des conditions nécessaires du 1er ordre pour qu'un point (x_0, y_0) soit solution de $(\mathcal{P})$.

A) Contraintes du type égalité ; extrema liés : exemple.

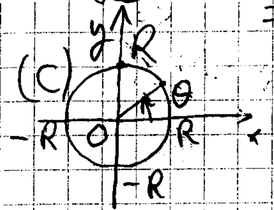
. On suppose donc $f, g \in C^1(\mathbb{R}^2;\mathbb{R})$ et
$\forall M=(x,y) \in (C)$, $\nabla g(M) \neq 0$. Alors

Thm 6 : Pour que $M_0=(x_0,y_0)$ soit solution de $(\mathcal{P})$, il est nécessaire (mais pas suffisant !) que $\nabla f(M_0)$ soit normal ($:\perp$) à la courbe (C) en M_0

. Idée de démonstration. Il y a 2 cas :

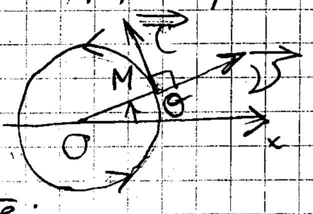
1er cas : on connaît une représentation paramétrique de la courbe (C) : ex.

$$\begin{cases} g(x,y) = x^2+y^2 = R^2 & : \text{équation} \\ \begin{cases} x(\theta) := R\cos\theta \\ y(\theta) := R\sin\theta \end{cases} & : \text{(représentation paramétrique)} \end{cases}$$



Alors la courbe (C) est l'image de $I=[0,2\pi]$ (ou de $\mathbb{R}$) par l'application $\theta \mapsto (R\cos\theta, R\sin\theta) =: M(\theta)$.

En tout point $M \in (C)$, le vecteur
$\frac{d\overrightarrow{M}}{d\theta} := \dot{M}(\theta) = (-R\sin\theta, R\cos\theta)$
est tangent à la courbe (C). On note :



$\vec{\tau} = \frac{\pm \dot{M}(\theta)}{\|\dot{M}(\theta)\|}$ le vecteur unitaire de la tangente à la courbe (C) au point M. Le signe $+$ ou $-$ définit l'orientation de la courbe (C).

Rem. De plus, on peut changer de paramètre : e.g. θ peut être fonction du temps t : $\theta = \Theta(t)$.
Alors $\frac{d\overrightarrow{M}}{dt}(t) = \frac{d\overrightarrow{M}}{d\theta}(\Theta(t)) \cdot \underbrace{\frac{d\theta}{dt}}_{\in \mathbb{R}}$ est le vecteur vitesse,

et $\frac{d^2\overrightarrow{M}}{dt^2} = \frac{d}{dt}\left(\frac{d\overrightarrow{M}}{dt}\right)$ est le vecteur accélération. □

. Dans tous les cas, pour une courbe (C) dans $\mathbb{R}^2$, le vecteur $\vec{\nu} = \begin{pmatrix}\nu_1 \\ \nu_2\end{pmatrix} := \begin{pmatrix}\tau_2 \\ -\tau_1\end{pmatrix}$ est un vecteur unitaire de la normale à (C) au point M.

. 1er cas (suite) : Si on connait une représentation paramétrique de la courbe (C), alors le pb :

(P) Trouver $(x_0, y_0) \in C$ tq $f(x_0, y_0) = \inf\{f(x,y) ; g(x,y) = 0\}$,

équivaut à : trouver $\theta_0 \in \mathbb{R}$ tq $F(\theta_0) = \inf_{\theta \in \mathbb{R}} F(\theta)$

où $F(\theta) := f(M(\theta)) = f(R\cos\theta, R\sin\theta)$.

. On cherche donc un minimum local de $\{\theta \mapsto F(\theta)\}$.

Comme F est C^1 (composée de fonctions C^1), la condition nécessaire du 1er ordre est :

$$0 = \frac{dF}{d\theta}(\theta_0) = \langle \nabla f(x(\theta_0), y(\theta_0)), (\dot{x}(\theta_0), \dot{y}(\theta_0)) \rangle$$

$$= \frac{\partial f}{\partial x}(M(\theta_0))\dot{x}(\theta_0) + \frac{\partial f}{\partial y}(M(\theta_0))\dot{y}(\theta_0). \text{ Donc}$$

$$\boxed{\frac{dF}{d\theta}(\theta_0) \underset{\text{ici}}{=} \frac{\partial f}{\partial x}(M_0)\cdot(-R\sin\theta_0) + \frac{\partial f}{\partial y}(R\cos\theta_0) = 0 \quad \text{(cond néc. d'ordre 1 en } M_0) \quad (CN)}$$

$$\iff \overrightarrow{\nabla f}(M_0) \perp \overrightarrow{\frac{dM}{d\theta}(\theta_0)}$$

$$\iff \nabla f(M_0) \text{ normal à la courbe } (C) \text{ en } M_0$$

On va voir que ceci équivaut à $\boxed{\overrightarrow{\nabla f}(M_0) \,//\, \overrightarrow{\nabla g}(M_0)}$

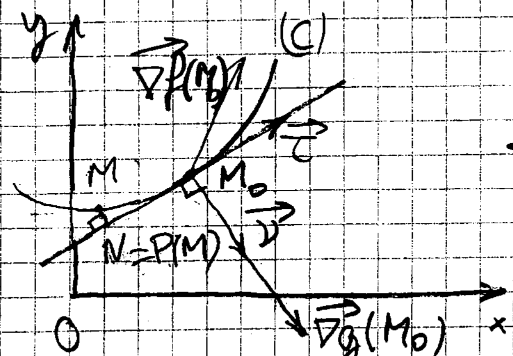
→ 2ème cas : d'abord l'équation de la courbe (C) est : $g(x,y) = 0$. Donc $\forall\theta$, $g(x(\theta), y(\theta)) = 0$. Dérivons ceci / à θ :

$$\frac{dG}{d\theta}(\theta) := \frac{d}{d\theta} g(x(\theta), y(\theta)) = \langle \overrightarrow{\nabla g}(M(\theta)), \underbrace{\overrightarrow{\frac{dM}{d\theta}}(\theta)}_{\text{vecteur tangent}} \rangle = 0, \ \forall\theta,$$

donc $\overrightarrow{\nabla g}(M(\theta)) \perp \overrightarrow{\frac{dM}{d\theta}}(\theta)$.

Donc la cond néc (CN) se réécrit : (CN') $\boxed{\overrightarrow{\nabla f}(M_0) \,//\, \overrightarrow{\nabla g}(M_0)}$

. Retrouvons cette condition sans connaître de représentation paramétrique de la courbe (C). Supposons que M_0 ne vérifie pas la condition (CN'), et montrons qu'alors M_0 n'est pas solution de (P)



On suppose ici que $\vec{\nabla} g(M_0) \neq \vec{0}$, que $g \in C^1(\mathbb{R}^2, \mathbb{R})$ et que $< \vec{\nabla} f(M_0), \vec{\tau} > \neq 0$. Alors, en remplaçant chaque point M voisin de M_0 sur la courbe (C) par sa projection N sur la tangente à la courbe (C) en M_0 et en faisant un DL_1 de f, on a :

$$f(M) - f(M_0) = < \vec{\nabla} f(M_0), \overrightarrow{M_0 M} > + \|\overrightarrow{M_0 M}\| \, \theta(M) \quad \text{où } \theta(M) \to 0 \ (M \to M_0)$$

$$= < \vec{\nabla} f(M_0), \overrightarrow{M_0 N} > + \underbrace{< \vec{\nabla} f(M_0), \overrightarrow{NM} > + \|\overrightarrow{M_0 M}\| \, \theta(M)}$$

$$= < \vec{\nabla} f(M_0), h(M) \vec{\tau} > + \underbrace{\|\overrightarrow{M_0 M}\| \left(\theta_1(M) + \theta(M) \right)}_{\text{idem}} .$$

Donc

$$f(M) - f(M_0) = h(M) \underbrace{< \vec{\nabla} f(M_0), \vec{\tau} >}_{\neq 0 \ (\text{ex. } >0 \text{ sur la figure})} + \ldots \quad \text{où } \theta_1(M) \text{ et } \theta(M) \to 0 \ (M \to M_0),$$

avec $h(M) \sim \pm \|\overrightarrow{M_0 M}\| \sim \pm \|\overrightarrow{M_0 N}\|$ (qd $M \to M_0$).

Donc ceci s'annule en changeant de signe quand $h(M)$ s'annule : si (CN) n'est pas vérifiée, M_0 n'est pas solution de (P).

. Rem. (CN) est nécessaire, mais pas suffisante. De plus, il peut y avoir plusieurs solutions de (P) :
ex. $\inf \{ y^2 ; \ x^2 + y^2 - 1 = 0 \}$. $\square$

Résumé : Minimisation avec contraintes du type égalité

Thm 7 (des extrema liés) : Soient $f, g_1, \ldots, g_p \in C^1(\mathbb{R}^n; \mathbb{R})$, avec $p < n$. On suppose que $\nabla g_1, \ldots, \nabla g_p$ sont linéairement indépendants au point $M_0 = a$, où a est est solution de

$$(\mathcal{P}) \quad \operatorname{Inf}\{f(x);\ g_1(x) = \cdots = g_p(x) = 0\} := \operatorname{Inf}\{f(x);\ x \in C\}$$

Alors la condition nécessaire :

$$(CN) \begin{cases} \exists \lambda_1, \ldots, \lambda_p \in \mathbb{R} \text{ tq } \nabla f(a) - \sum_{i=1}^{p} \lambda_i \nabla g_i(a) = 0, \\ \text{et } g_1(a) = \cdots = g_p(a) = 0 \end{cases}$$

est vérifiée. On trouve $a = (a_1, \ldots, a_n)$ et les λ_i (les multiplicateurs de Lagrange) en cherchant les points critiques de $\phi : (x, \lambda) \mapsto f(x) - \sum_{i=1}^{p} \lambda_i g_i(x)$, i.e. en résolvant

$$\begin{cases} \forall j = 1, \ldots, n, \ \dfrac{\partial f}{\partial x_j}(a) - \sum_{i=1}^{p} \lambda_i \dfrac{\partial g_i}{\partial x_j}(a) = 0, \text{ et} \\ \forall i = 1, \ldots, p, \ \dfrac{\partial \phi}{\partial \lambda_i}(a) = -g_i(a) = 0. \end{cases}$$

La condition n'est pas suffisante

B) Minimisation avec contraintes du type inégalité :

Thm 8 : Extrema liés, contraintes du type inégalité

Soient $f, g_1, \dots, g_p \in C^1(\mathbb{R}^n; \mathbb{R})$. On considère le pb :

$(\mathcal{P}_C)$: $\operatorname{Inf}\{f(x) ; x \in C\}$, où

$C = \bigcap_{1 \le i \le p} C_i$, et $C_i := \{x \in \mathbb{R}^n ; g_i(x) \le 0\}$. On suppose que

$\exists x \in \mathbb{R}^n$ tq $\forall i, g_i(x) < 0$. (H_1).

Comme g_i est continue $\forall i$, $\overset{\circ}{C_i} = \operatorname{Int}(C_i) \supset \{x \in \mathbb{R}^n ; g_i(x) < 0\}$, et $\overset{\circ}{C} = \bigcap_{i=1} \overset{\circ}{C_i} \neq \emptyset$.

Alors si $a \in C$ est solution de $(\mathcal{P}_C)$, nécessairement,

$$\boxed{\begin{array}{l} \exists \lambda = (\lambda_i) \in \mathbb{R}^p \text{ tq :} \\ \nabla f(a) - \sum_{1 \le i \le p} \lambda_i \nabla g_i(a) = 0, \text{ et tq} \qquad (CN) \\ \forall i = 1, \dots, p, \quad \lambda_i \le 0, \ g_i(a) \le 0, \text{ et } \lambda_i g_i(a) = 0. \end{array}}$$

La condition (CN) est nécessaire, mais pas suffisante !

Idée de dém.

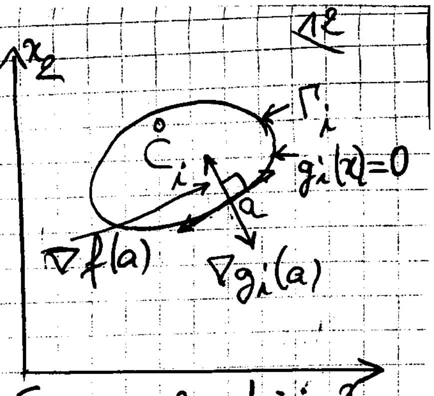
. On cherche d'abord les points critiques de f dans $\overset{\circ}{C}$. Ils vérifient : $\nabla f(a) = 0$, donc ils vérifient (CN) avec $\lambda_1 = \dots = \lambda_p = 0$.

. On cherche ensuite une solution a de $(\mathcal{P}_C)$ tq $a \in \operatorname{Fr}(C_i)$, donc $g_i(a) = 0$. Supposons d'abord que $\forall j \neq i$, $g_j(a) < 0$. Dans ce cas, $f(a) = \operatorname{Inf}\{f(x) ; g_i(x) \le 0\} \underset{\text{aussi}}{=} \operatorname{Inf}\{f(x) ; g_i(x) = 0\}$

[Exercice : justifier ces deux affirmations]

. Mais ici, on peut se déplacer non seulement sur la courbe $\Gamma_i = \{x \in \mathbb{R}^n ; g_i(x) = 0\}$, mais aussi vers l'intérieur de C_i : le Thm 7 (ici, c'est comme si on avait 1 seule contrainte) devient : $\exists \lambda_i \le 0$ tel que $\nabla f(a) - \lambda_i \nabla g_i(a) = 0$.

<u>Rappel</u>: $\forall h \in \mathbb{R}^n$, $\nabla f(a) \cdot h \geq 0 \Leftrightarrow \overrightarrow{\nabla f(a)}$ et $\vec{h}$ font un angle aigu. Donc ici, pour diminuer f, <u>on n'a pas intérêt à se déplacer sur la courbe Γ_i</u>, car $\nabla f(a)$ est $\perp$ à la tangente à Γ_i au point a, et <u>on n'a pas intérêt à se déplacer vers $\mathring{C}_i$</u>, car $\nabla f(a) \cdot \nabla g_i(a) < 0$.



Ex. $n=2$, et ici $\forall j \neq i$, $g_j(a) < 0 = g_i(a)$.

. <u>Rem.</u> $\forall i$, $\lambda_i g_i(a) = 0$: seules les <u>contraintes "actives"</u>, <u>ici $g_i(a)=0$</u>, <u>interviennent</u> : si $\forall i = 1, \dots, p$, $g_i(a) < 0$, alors la cond. nécessaire (CN) est : $\nabla f(a) = 0$, <u>comme dans le cas sans contrainte</u>. $\square$

. <u>Rem. importante</u> : Dans le cas où f (<u>la fonction-coût</u>) et les contraintes $x \in C_i = \{x ; g_i(x) \leq 0\}$ sont <u>convexes</u>, les conditions ci-dessus sont aussi <u>suffisantes</u>, cf cours d'optimisation. $\square$

§ 6. <u>Thm d'inversion locale. Thm des fonctions implicites</u> :

Ce thm a une version multi-dimensionnelle. <u>On se limite au cas de deux variables</u> : Soit $f : \underset{\text{ouvert}}{\Omega} \subset \mathbb{R}^2 \to \mathbb{R}$. Soit $(x_0, y_0) \in \Omega$. Peut-on (<u>localement</u>) représenter l'ensemble $(C) := \{(x,y) \in \underline{\omega} \subset \Omega ; f(x,y) = 0\}$ sous la forme $(C) \underset{\text{aussi}}{=} \{(x, \underset{\substack{\| \\ y}}{\varphi(x)}), x \in \underline{I} = \text{intervalle} \ni x_0, \varphi \in C^1(\mathbb{R})\}$?

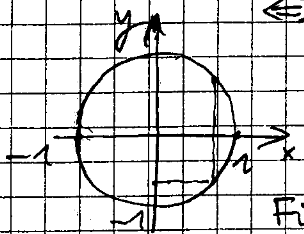
. Autrement dit, <u>peut-on exprimer $\genfrac{}{}{0pt}{}{y}{\text{ou } x}$ en fonction de $\genfrac{}{}{0pt}{}{x}{\text{ou } y}$</u>, au moins dans un voisinage de (x_0, y_0) ?

. <u>Rem.</u> Si oui, (C) est une <u>courbe</u>, d'éq. $\{f(x,y)=0\}$, et (localement) d'équation $\left\{ \begin{array}{l} y = \varphi(x) \\ \text{ou } x = \psi(y) \end{array} \right.$. <u>Ce n'est pas toujours</u>

<u>le cas</u> :

- <u>Exemple 1</u> : $f(x,y)\equiv 0$, $\Omega =]0,1[\times]0,1[$
- <u>Exemple 2</u> : $f(x,y) = x^2+y^2-1 = 0$

$\iff y = +\sqrt{1-x^2} = \varphi_1(x)$ ou $y = -\sqrt{1-x^2} = \varphi_2(x)$, $|x|\le 1$.



<u>Figure 1</u>

- <u>Thm 9</u> <u>(fonctions implicites)</u> ****

(i) Soit $f\in C^1(\Omega;\mathbb{R})$, Ω ouvert $\subset\mathbb{R}^2$. Soit $(x_0,y_0)\in\Omega$ tq $f(x_0,y_0)=0$. On suppose que $\boxed{\frac{\partial f}{\partial y}(x_0,y_0)\neq 0.}$

Alors $\exists$ deux (petits) intervalles ouverts I, J tq $x_0\in I$, $y_0\in J$, et $\forall (x,y)\in I\times J\subset\Omega$, $f(x,y)=0 \iff y=\varphi(x)$

où φ est une fonction C^1, déterminée de manière unique, tq $y_0=\varphi(x_0)$, et $\varphi'(x_0) = \dfrac{\frac{\partial f}{\partial x}(x_0,y_0)}{\frac{\partial f}{\partial y}(x_0,y_0)}$

(ii) Ce thm a une version multi-dimensionnelle pour $f:\mathbb{R}^n\to\mathbb{R}^p$, $n>p$, sous la condition qu'une <u>matrice jacobienne</u> convenable soit <u>inversible</u>...

⚠ <u>Interprétation géométrique</u> : éviter le cas $(x_0,y_0)=(\pm 1,0)$ dans la Figure 1. <u>Notez qu'en ces points</u>, $\frac{\partial f}{\partial y}(x_0,y_0)=0$.

- <u>Idée de démonstration</u> :

(i) on se ramène au cas où $\begin{cases}(x_0,y_0)=(0,0) \text{ (par translation) et}\\ \partial_x f(x_0,y_0)=0 \text{ et } \partial_y f(x_0,y_0)=1,\end{cases}$

en remplaçant f par $\tilde f(x,y) := f\left(x, -\dfrac{\partial_x f(0,0)}{\partial_y f(0,0)}x + \dfrac{y}{\partial_y f(0,0)}\right)$

(vérifier que $\partial_x\tilde f(0,0)=0$, et $\partial_y\tilde f(0,0)=1$).

. On pose ensuite, après avoir remplacé f par $\tilde{f}$, $F(x,y) := y - f(x,y)$, et pour chaque x fixé $\in]x_0-\delta, x_0+\delta[$, on montre que l'application $\{y \mapsto F(x,y)\}$ est contractante de rapport $k = \frac{1}{2}$ (e.g.) : $|F(x,y_1) - F(x,y_2)| \leq \frac{1}{2}|y_1 - y_2|$.

Donc l'application $\{y \mapsto F(x,y)\}$ a un unique point fixe : $\forall x \in I, \exists !\, y := \varphi(x)$, tq $y = F(x, \underbrace{\varphi(x)}_{y}) \Leftrightarrow y - f(x, \underbrace{\varphi(x)}_{y}) = y$

$$\Updownarrow$$

$$f(x, \underbrace{\varphi(x)}_{y}) = 0.$$

. Donc $\forall x \in I, y \in J$, on a $f(x, \varphi(x)) = 0$.

. En calculant la dérivée de la fonction composée $\{x \mapsto f(x, \varphi(x))\}$, on obtient

$$\frac{d}{dx} f(x,\varphi(x)) = \partial_x f(x,\varphi(x)) \cdot 1 + \partial_y f(x,\varphi(x)) \cdot \varphi'(x) = \frac{d}{dx}(0) = 0 \; ;$$

donc $\varphi'(x) = -\dfrac{\partial_x f(x,\varphi(x))}{\partial_y f(x,\varphi(x))}$, $\forall x \in I$. □

. **thm 10 : (****)** : Soit $f : \Omega_{\text{ouvert}} \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction C^1, soit $x_0 \in \Omega$ un point où la matrice jacobienne $Df(x_0)$ est inversible : $\det(Df(x_0)) \neq 0$. Alors $\exists U$ (resp. V) voisinage ouvert de x_0 (resp. de y_0) tq f soit une bijection de U sur V et tq f^{-1} soit C^1 de V sur U. On dit alors que f est un C^1-difféomorphisme de U sur V.

Dém. admise. Noter que si $f(x) = Ax$, alors $\forall x_0 \in \Omega = \mathbb{R}^n$, $Df(x_0) = A$. Dans ce cas, si $\det(A) \neq 0$, $f^{-1} : y \mapsto A^{-1}y$ (resp. $f : x \mapsto Ax$) est une bijection de $V = \mathbb{R}^n$ sur $U = \mathbb{R}^n$ (resp. de $U = \mathbb{R}^n$ sur $V = \mathbb{R}^n$).

□