Chapitre IV Optimisation

4.1 ~~Classification des~~ Position des problèmes d'optimisation

Soit $K \subset \mathbb{R}^n$ une partie et $J : K \to \mathbb{R}$ une fonction.

$$(P) \begin{cases} \text{Trouver } u \in K \\ \text{tq } J(u) = \inf\limits_{v \in K} J(u) \end{cases}$$

J = fonction coût $\quad K$ = ensemble de contraintes.

Applications nombreuses, en physique, chimie, en économie.

4.2 Rappels théoriques sur les problèmes d'optimisation

(a) Conditions suffisantes d'existence d'une solution à (P)

Prop: Si $K \subset \mathbb{R}^n$ est une partie compacte de $\mathbb{R}^n$ et $J : K \to \mathbb{R}$ continue

$$\exists u_0 \in K \qquad J(u_0) = \inf_{u \in K} J(u).$$

~~Coercitif~~

Déf: $J : \mathbb{R}^n \to \mathbb{R}^n$ est dite <u>coercive</u> si $\lim\limits_{\|u\| \to +\infty} J(u) = +\infty$.

Prop: Si $F \subset \mathbb{R}^n$ est une partie fermée non vide et $J : \mathbb{R}^n \to \mathbb{R}^n$ continue coercive

$$\exists u_0 \in F \qquad J(u_0) = \inf_{u \in F} J(u).$$

Preuve: L'hypothèse signifie

$$\forall M > 0 \quad \exists \underset{=r(M)}{r} > 0 \qquad \|x\| > r \Rightarrow f(x) \geq M.$$

Soit $v \in F$ posons $M = J(v) + 1$ et $r = r(M)$.

Alors $K = F \cap \bar{B}(0, r)$ compact et donc $f|_K$ a un minimum en $u_0 \in K$

On a $J(u_0) \leq J(v) \leq M - 1$ et donc si $u \in F \cap \{\|x\| > r\}$ $\quad J(u) > J(u_0)$

Donc u_0 est le minimum de $J/_F$.

(b). Conditions d'existence d'un solution de (P)

Prop Si $U \subset \mathbb{R}^n$ ouvert et $J : U \to \mathbb{R}$ de classe C^2. Soit $u_0 \in U$ un minimum local de J alors

$$\frac{\partial J}{\partial x_i}(u_0) = 0$$

$$\left(\frac{\partial^2 J}{\partial x_i \partial x_j}(u_0) \right) \geq 0$$

Réciproqt si

$$\nabla J = \left(\frac{\partial J}{\partial x_i}(u_0)\right) = 0 \qquad i = 1 \dots n$$

$$H(f) = \left(\frac{\partial^2 J}{\partial x_i \partial x_j}(u_0)\right) \text{ définie positive}$$

Alors u_0 est un minimum local (qu'on dira non dégénéré).

<u>Prop</u> (mult. de Lagrange) Soit $J \in C^1(\mathbb{R}^n;\mathbb{R})$ et $f : \mathbb{R}^n \xrightarrow{C^1} \mathbb{R}$.

On pose $P = f^{-1}(0)$

Si $u_0 \in P$ ~~est~~ vérifie $\begin{cases} J(u_0) = \min\limits_{u \in P} J(u) \\ \left(\frac{\partial f}{\partial x_i}(u_0)\right) \neq 0 \end{cases}$.

Alors $\exists \lambda \in \mathbb{R}$ $\quad \frac{\partial J}{\partial x_i}(u_0) = \lambda . \frac{\partial f}{\partial x_i}(u_0)$.

Ces deux propositions sont des classiques du cours de calcul différentiel.

43 Algorithmes d'optimisation pour les problèmes sans contraintes

(a) <u>méthode de Newton</u> Elle revient à étudier le système $\frac{\partial J}{\partial x_i} = 0 \quad i = 1 \dots n$.

<u>Prop</u> Si $J \in C^2(\mathbb{R}^n;\mathbb{R})$ et u_0 vérifie $\frac{\partial J}{\partial x_i}(u_0) = 0$ et $\left(\frac{\partial^2 J}{\partial x_i \partial x_j}(u_0)\right)$ inversible

Alors u_0 est un pt fixe attractif de la méthode de Newton :

$$u_{n+1} = \Phi(u_n)$$

avec $$\Phi(x) = x - \left[\frac{\partial^2 J}{\partial x_i \partial x_j}(x)\right]^{-1} \begin{pmatrix} \frac{\partial J}{\partial x_1} \\ \vdots \\ \frac{\partial J}{\partial x_n} \end{pmatrix} = x - H(J)^{-1} \nabla J(x).$$

Variante : méthode de la corde.

(b) <u>méthode de relaxation pour les fonctions elliptiques</u>

<u>Déf</u> $J \in C^2(\mathbb{R}^n;\mathbb{R})$ $\quad J$ est [illegible] <u>elliptique</u>) ssi

$\exists \alpha > 0 \qquad H(J) \geq \alpha I_n$

(ie la matrice sym. $H(J) - \alpha I_n$ est ~~déf~~ définie semi-positive).

Chercher les minima de fonctions i... : ~~es~~ peu restrictif :

Lemme : Soit u_0 un min. local ~~non~~ dégénéré de $J \in C^2(\mathbb{R}^n;\mathbb{R})$

alors $\exists$ $\tilde{J} \in C^2(\mathbb{R}^n, \mathbb{R})$ " " tq

$$\begin{cases} J(u_0) = \min\limits_{u \in \mathbb{R}^n} J(u) \\ J = \tilde{J} \text{ dans un voisinage de } u_0 \end{cases}$$

La preuve sort du programme de ce cours.
Elle est relativement simple.

Proposition : Soit $J : \mathbb{R}^n \to \mathbb{R}$ ~~est~~ $\alpha > 0$ comme dans la définition

(1) $\forall (x,y) \in \mathbb{R}^n$ $\quad J(x) - J(y) \geq (\nabla J(y), x-y) + \frac{\alpha}{2}\|x-y\|^2$.
(2) J coercive
(3) ~~Si~~ Soit $U \subset \mathbb{R}^n$ une partie convexe fermée non vide.

Alors $\exists! \, u_0 \in U$ $\quad J(u_0) = \min\limits_{u \in U_0} J(u)$.

Preuve : • la formule de Taylor ~~appliquée~~ à $j(t) = f(y + t(x-y))$ en $t=0$

donne :
$$j(1) = j(0) + j'(0) + \frac{j''(c)}{2} \quad \text{pour un certain } c \in]0,1[.$$

$j(1) = J(x)$
$j(0) = J(y)$

$$j'(0) = (\nabla J(y); (x-y)) = d_y J_y (x-y)$$

$$j'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(y + t(x-y)) \cdot (x_i - y_i)$$

$$j''(t) = \sum_{i=1}^{n}\sum_{j=1} \frac{\partial^2 f}{\partial x_i \partial x_j}(y + t(x-y))\,(x_i - y_i)(x_j - y_j)$$

$$= \left(H(f)[y + t(x-y)] \cdot (x-y) ; (x-y) \right)$$

$$\geq \alpha \|x-y\|^2$$ par la propriété caractéristique de la définition. D'où (1).

• En particulier $J(x) \geq (\nabla J(0), x) + \frac{\alpha}{2}\|x\|^2 \geq -B\|x\| + \frac{\|x\|^2}{2}$

$B = \|\nabla J(0)\|$ et $\lim\limits_{x \to \infty} J(x) = +\infty$ D'où (2).

L'existence d'un minimum dans (3) découle. Soit $u_0 \in U$ un tel minimum et $v \in U$ alors tout $v_t = u_0 + t(v-u_0)$ $t \in [0,1]$ on a :

$$J(v_t) - J(u_0) \geq (\nabla J(u_0) ; (v_t - u_0)) + \frac{\alpha}{2} \| v_t - u_0 \|^2$$

$$\geq t(\nabla J(u_0) ; v - u_0) + \frac{\alpha}{2} \| v - u_0 \|^2 t^2 .$$

De plus $0 \leq \lim_{t \to 0^+} \frac{J(v_t) - J(u_0)}{t} = (\nabla J(u_0) ; v - u_0)$ donc

car u_0 est un minimum

Donc $(\nabla J(u_0) ; v - u_0) \geq 0$

De sorte que $J(v_t) - J(u_0) \geq \frac{\alpha}{2} \| v - u_0 \|^2 t^2 > 0$ si $v \neq u_0$.

D'où l'unicité dans (3).

<u>Méthode de relaxation</u> : Soit $J : \mathbb{R}^n \longrightarrow \mathbb{R}$ unif^t convexe

Soit $x_0 \in \mathbb{R}^n$. $(e_1 \dots e_n)$ base canonique

On pose $x_0^{(0)} = x_0$

$x_0^{(1)} = x_0^{(0)} + t_0^{(1)} e_1$ avec $J(x_0^{(0)} + t_0^{(1)} e_1) = \min_{t^{(1)} \in \mathbb{R}} J(x_0^{(0)} + t^{(1)} e_1)$

$x_0^{(k+1)} = x_0^{(k)} + t_0^{(k+1)} e_{k+1}$ avec $J(x_0 + t_0^{(k)} e_{k+1}) = \min_{t^{(k+1)} \in \mathbb{R}} J(x_0^{(k)} + t^{(k+1)} e_{k+1})$

Ce qui définit $x_0^{(0)} \dots x_0^{(n)}$.

On pose alors $x_1 = x_0^{(n)}$ puis on itère de façon à obtenir

une suite $(x_i)_{i \in \mathbb{N}}$ [et même $n+1$ suites : $(x_i^{(k)})_{i \in \mathbb{N}}$ $k = 0 \dots n$].

<u>Théorème</u> : Si J coercive. pour tout $x_0 \in \mathbb{R}^n$ la méthode de relaxation converge vers l'unique minimum de J sur $\mathbb{R}^n$.

Preuve : • Appliquons l'inégalité (1) de la proposition :

$$J(x_i^{(k)}) - J(x_i^{(k+1)}) \geq \left(\nabla J(x_i^{(k)}) \cdot (x_i^{(k)} - x_i^{(k+1)})\right)$$

$$+ \frac{\alpha}{2} \| x_i^{(k)} - x_i^{(k+1)} \|^2.$$

mais $\left(\nabla J(x_i^{(k+1)})\right)\left(\underbrace{x_i^{(k)} - x_i^{(k+1)}}_{= t_i^{(k)} e_{k+1}}\right) = 0$ car $x_i^{(k+1)}$ est le point où J a un min. sur $x_i^{(k)} + \mathbb{R} e_{k+1}$.

Donc

$$J(x_i^{(k)}) - J(x_i^{(k+1)}) \geq \frac{\alpha}{2} \| x_i^{(k)} - x_i^{(k+1)} \|^2$$

et

$$J(x_i^{(n)}) - J(x_i^{(0)}) \geq \frac{\alpha}{2} \sum_{k=0}^{n-1} \| x_i^{(k)} - x_i^{(k+1)} \|^2$$

$$= \frac{\alpha}{2} \| x_i^{(n)} - x_i^{(0)} \|^2 = \frac{\alpha}{2} \| x_{i+1} - x_i \|^2$$

car les $(x_i^{(k)} - x_i^{(k+1)})$ sont 2 à 2 orthogonaux.

Donc $(J(x_i^{(n)}))_{n \in \mathbb{N}}$ est décroissante minorée donc converge.

et

$$\lim_{n \to \infty} \| x_{i+1} - x_i \| = 0$$

(ainsi que $\lim_{n \to \infty} \| x_i^{(k)} - x_i \| = 0$)

•

Soit u_0 le point où le min de J est atteint :

$$0 \geq J(u_0) - J(x_i) \geq \left(\nabla J(x_i) ; u_0 - x_i\right) + \frac{\alpha}{2} \| u_0 - x_i \|^2.$$

$$\frac{\alpha}{2}\|u_0 - x_i\|^2 \leq -\sum_{\ell=1}^{n} \frac{\partial J}{\partial x_\ell}(x_i^0)\cdot(u_{0,\ell} - x_{i,\ell})$$

Comme $\frac{\partial J}{\partial x_\ell}(x_i^{(k)}) = 0$

$$\sum_{\ell=1}^{n} \frac{\partial J}{\partial x_\ell}(x_i)(x_{i\ell} - u_{0,\ell}) = \sum_{\ell=1}^{n} \frac{\partial J}{\partial x_\ell}(x_i) - \frac{\partial J}{\partial x_\ell}(x_i^{(\ell)}) \cdot (x_{i,\ell} - u_{0,\ell})$$

$$\leq \|u_0 - x_i^0\| \cdot \sum_{\ell=1}^{n} \left| \frac{\partial J}{\partial x_\ell}(x_i) - \frac{\partial J}{\partial x_\ell}(x_i^{(\ell)}) \right|$$

Donc

$$\frac{\alpha}{2}\|u_0 - x_i\| \leq \sum_{\ell=1}^{n} \left| \frac{\partial J}{\partial x_\ell}(x_i) - \frac{\partial J}{\partial x_\ell}(x_i^{(\ell)}) \right|$$

- Comme J coercive $(x_i)_{i\in\mathbb{N}}$ $(x_i^{(\ell)})_{i\in\mathbb{N}}$ restent dans ~~une boule fermée de ay~~ un compact. K

Or $\frac{\partial J}{\partial x_\ell}$ est continue donc unif^t continue sur K

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \begin{matrix}\|x-y\| < \delta(\varepsilon) \\ x, y \in K\end{matrix} \Rightarrow \left| \frac{\partial J}{\partial x_\ell}(x) - \frac{\partial J}{\partial x_\ell}(y) \right| < \varepsilon$$

Or $x_i - x_i^{(\ell)} \xrightarrow[i\to\infty]{} 0$ $\exists i_0$ tq pour $i \geq i_0$ $|x_i - x_i^{(\ell)}| < \delta(2\varepsilon/\alpha n)$

et on a pour $i \geq i_0$ $\|u_0 - x_i\| \leq \varepsilon$.

Donc $\lim_{i\to\infty} x_i = u_0$. $\square$.

(c) méthode du gradient à pas optimal pour les fonctions unif^t convexes.

$J \in C^2(\mathbb{R}^n, \mathbb{R})$ elliptique.

$x_0 \in \mathbb{R}^n$

~~On pose~~ (

$$\begin{cases} x_{k+1} = x_k - \rho_k \nabla J(x_k) \\ \rho_k \text{ choisi tel que } \quad J(x_{k+1}) = \min\limits_{t \in \mathbb{R}} J(x_k - t \nabla J(x_k)) \quad (*) \end{cases}$$

Noter que $x_{k+1} = x_k$ si $\nabla J(x_k) = 0$ auquel cas x_k est le min. recherché

<u>Thm</u>: $(x_k)_{k \in \mathbb{N}}$ converge vers l'unique minimum de J sur $\mathbb{R}^n$.

<u>Preuve</u>: Cette preuve est une variante de la précédente et ne sera pas donnée en amphi

- Comme précédt on a

$$J(x_k) - J(x_{k+1}) \geq \frac{\alpha}{2} \|x_k - x_{k+1}\|^2$$

Donc $J(x_k)$ converge et $\|x_k - x_{k+1}\| \xrightarrow[k \to \infty]{} 0$.

- le fait que $J(x_{k+1})$ soit le minimum comme en $(*)$

donne $\quad (\nabla J(x_{k+1}) \cdot \nabla J(x_k)) = 0$

Donc $\quad \|\nabla J(x_k)\|^2 = (\nabla J(x_k) ; \nabla J(x_k) - \nabla J(x_{k+1}))$

$$\leq \|\nabla J(x_k)\| \cdot \|\nabla J(x_k) - \nabla J(x_{k+1})\|.$$

et $\quad \|\nabla J(x_k)\| \leq \|\nabla J(x_k) - \nabla J(x_{k+1})\|$

(même si $\nabla J(x_k) = 0 \ldots$)

- Comme $(x_k)_{k \in \mathbb{N}}$ bornée (reste dans un compact) et ∇J uniformt continue sur ce compact l'argt final du théorème précédent s'applique.

(d) méthode de gradient conjugué

C'est une variante de la précédente On fixe $k \geq 1$ et on suppose défini $(x_0, x_1, \dots, x_\ell)$ $\ell \geq k$.

On définit $x_{\ell+1}$ comme le point dans l'espace affine $x_\ell + V_\ell$ $V_\ell = \mathrm{Vect}(\nabla J(x_\ell)\ \nabla J(x_{\ell-1}) \dots \nabla J(x_{\ell-k+1}))$

où $J(x_\ell + V)$ atteint son minimum.

Cette méthode est a priori meilleure que les précédentes mais nous n'effectuerons pas son analyse détaillée, faute de temps.