

Méthode du gradient à pas optimal:

Référence: [Cialet P] [Oraus X-ENS Analyse 4]

Énoncé: Soit $J: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonctionnelle elliptique (i.e. J est $\mathcal{C}^1$ sur $\mathbb{R}^n$ et $\exists \alpha > 0$ tq $\forall (x, y) \in (\mathbb{R}^n)^2$

$$\langle \nabla J(x) - \nabla J(y), x - y \rangle \geq \alpha \|x - y\|^2. \quad (1)$$

Alors J a un unique point de minimum et la méthode du gradient à pas optimal converge.

Étapes: (E1) Existence et unicité du minimum

(E1.1) Caractérisation $J(x) - J(y) \geq \langle \nabla J(y), x - y \rangle + \frac{\alpha}{2} \|x - y\|^2$

(E1.2) Coercivité + continuité $\Rightarrow$ existence

(E1.2) α -convexité $\Rightarrow$ unicité

(E2) Construction de la suite minimisante

$x_0 \in \mathbb{R}^n$, $\forall k \in \mathbb{N}$ on cherche p_k tq $J(x_k - p_k \nabla J(x_k))$ est minimale. Étude pour $x \in \mathbb{R}^n$ de l'app.

$$\varphi_x: p \in \mathbb{R} \mapsto J(x - p \nabla J(x)).$$

(E3) Convergence vers x^*

(E3.1) Mq si $\nabla J(x_k) \xrightarrow[k \rightarrow +\infty]{} 0$ alors $x_k \xrightarrow[k \rightarrow +\infty]{} x^*$

(E3.2) Mq si $\|\nabla J(x_k) - \nabla J(x_{k+1})\| \xrightarrow[k \rightarrow +\infty]{} 0$ alors $\nabla J(x_k) \xrightarrow[k \rightarrow +\infty]{} 0$

Pour cela mq $(x_k) \in \overline{B(0, A)}$ compact $\rightarrow \nabla J$ unif. cont.

$$\text{et } \|x_k - x_{k+1}\| \xrightarrow[k \rightarrow +\infty]{} 0$$

Étape 1: Existence et unicité du minimum

Comme J est $\mathcal{C}^1$ sur $\mathbb{R}^n$, on peut écrire la formule de Taylor avec reste intégral à l'ordre 1.

$$\text{Soit } (x, y) \in (\mathbb{R}^n)^2 \quad J(y) = J(x) + \int_0^1 \langle \nabla J(x + t(y-x)), y-x \rangle dt$$

$$\text{Alors } J(y) - J(x) - \langle \nabla J(x), y-x \rangle = \int_0^1 \langle \nabla J(x + t(y-x)) - \nabla J(x), y-x \rangle dt$$

Et comme J est elliptique, $\exists \alpha > 0$ tel que

$$\forall t \in [0, 1], \langle \nabla J(x + t(y-x)) - \nabla J(x), t(y-x) \rangle \geq \alpha \|t(y-x)\|^2$$

$$\text{donc } \forall t \in [0, 1], \langle \nabla J(x + t(y-x)) - \nabla J(x), y-x \rangle \geq \alpha t \|y-x\|^2$$

$$\text{donc } J(y) - J(x) - \langle \nabla J(x), y-x \rangle \geq \alpha \|y-x\|^2 \int_0^1 t dt$$

$$\text{donc } J(y) - J(x) - \langle \nabla J(x), y-x \rangle \geq \frac{\alpha}{2} \|y-x\|^2 \quad (2)$$

Ainsi, cette nouvelle caractérisation nous permet de montrer que J est coercive.

$$\forall x \in \mathbb{R}^n, J(x) \geq J(0) + \langle \nabla J(0), x \rangle + \frac{\alpha}{2} \|x\|^2$$

$$\forall x \in \mathbb{R}^n, J(x) \geq J(0) - \underbrace{\|\nabla J(0)\| \cdot \|x\|}_{\xrightarrow{\|x\| \rightarrow +\infty} +\infty} + \frac{\alpha}{2} \|x\|^2$$

$$\text{Donc } J(x) \xrightarrow{\|x\| \rightarrow +\infty} +\infty$$

$$\text{Ainsi } \exists A \in \mathbb{R} \text{ tq } \forall x \in \mathbb{R}^n \setminus B(0; A) \quad J(x) \geq J(0)$$

Donc on peut restreindre la recherche de minimum au domaine $\overline{B(0; A)}$

Comme $\mathbb{R}^n$ est de dimension finie et $\overline{B(0;A)}$ est un fermé-borné, $\overline{B(0;A)}$ est compact et comme J est continue, elle y est bornée et atteint ses bornes.

$$\text{Donc } \exists x^* \in \overline{B(0;A)} \quad \inf_{v \in \mathbb{R}^n} (J(v)) = \inf_{v \in \overline{B(0;A)}} (J(v)) = J(x^*) = J_{\min}$$

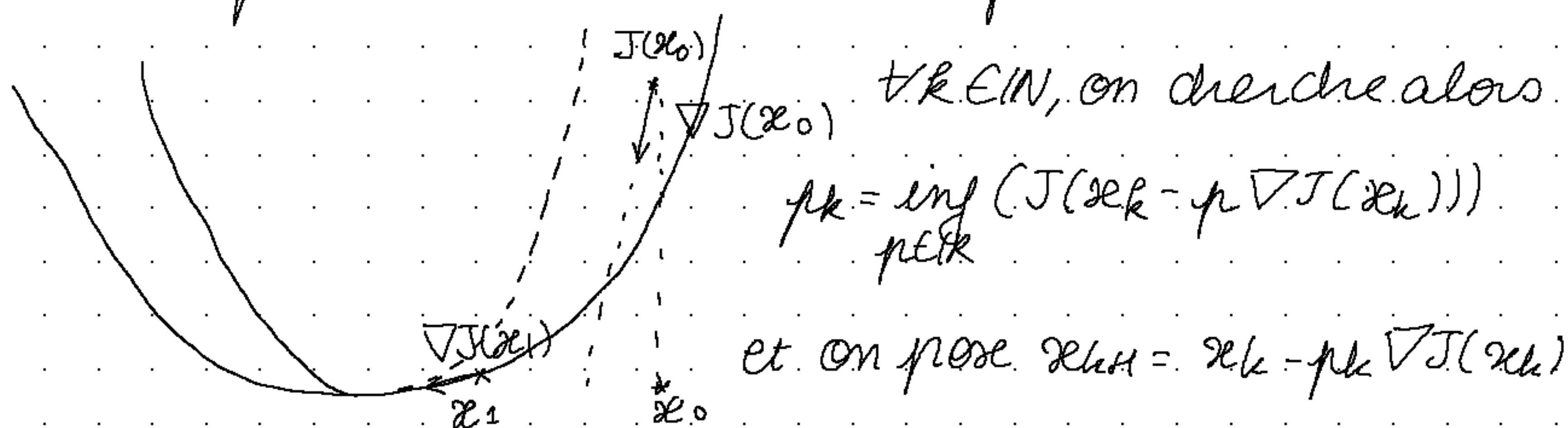
Supposons $\exists \tilde{x} \in \mathbb{R}^n$ tq $J(\tilde{x}) = J(x^*) = \inf_{x \in \mathbb{R}^n} (J(x))$.

$$\text{On a alors } \underbrace{J(\tilde{x}) - J(x^*)}_{=0} = \underbrace{\langle \nabla J(x^*), \tilde{x} - x^* \rangle}_{=0 \text{ car } x^* \text{ est un minimum}} \geq \frac{\alpha}{2} \|\tilde{x} - x^*\|^2$$

donc $0 \geq \frac{\alpha}{2} \|\tilde{x} - x^*\|^2 \geq 0$ donc $\tilde{x} = x^*$.

Étape 2 : Construire la suite minimisante.

Principe : le gradient indique la direction de la plus forte pente. On peut alors prendre $x_0 \in \mathbb{R}$ et construire x_1 en cherchant le pas p_0 le plus optimal pour minimiser $J(x_0 - p_0 \nabla J(x_0))$.



Pour pouvoir à chaque fois trouver p_k , cela demande d'étudier, pour $x \in \mathbb{R}^n$, la fonction $\phi_x : \mathbb{R} \rightarrow \mathbb{R}$

$$t \mapsto J(x - t \nabla J(x))$$

φ_λ est $\mathcal{C}^1$ comme composée de fonctions $\mathcal{C}^1$ et on a

$$\|x - t \nabla J(x)\| \xrightarrow[t \rightarrow \pm\infty]{} +\infty \text{ donc par composition des limites}$$

$$\varphi_\lambda(t) \xrightarrow[t \rightarrow \pm\infty]{} +\infty \text{ donc } \varphi_\lambda \text{ est continue et coercive,}$$

elle admet donc un minimum et on notera $\tilde{z}$ un point de minimum. Supposons que $\tilde{z} \in \mathbb{R}$ soit un autre point de minimum.

Comme φ_λ est $\mathcal{C}^1$, elle est différentiable et en $\tilde{z}$ et $\tilde{z}$,

$$\varphi'_\lambda(\tilde{z}) = \varphi'_\lambda(\tilde{z}) = 0$$

$$\text{or } \varphi'_\lambda(t) = \langle \nabla J(x - t \nabla J(x)) ; -\nabla J(x) \rangle \text{ pour tout } t \in \mathbb{R}$$

$$\text{Donc on en conclut } \langle \nabla J(x - \tilde{z} \nabla J(x)) ; \nabla J(x) \rangle = 0$$

$$\text{et } \langle \nabla J(x - \tilde{z} \nabla J(x)) ; \nabla J(x) \rangle = 0 \quad (*)$$

Et comme J est elliptique, la relation (1) donne:

$$\underbrace{\langle \nabla J(x - \tilde{z} \nabla J(x)) - \nabla J(x - \tilde{z} \nabla J(x)) ; (\tilde{z} - \tilde{z}) \nabla J(x) \rangle}_{=0} \geq \frac{\alpha}{\|\nabla J(x)\|^2} (\tilde{z} - \tilde{z})^2$$

$$\text{Donc } \|\nabla J(x)\| = 0 \text{ ou } (\tilde{z} - \tilde{z}) = 0$$

Si $\nabla J(x) \neq 0$ alors on a un unique point $\tilde{z}$ de minimum

Si $\nabla J(x) = 0$ alors on a en fait trouvé le minimum de J .

$$\text{En effet, } \forall y \in \mathbb{R}^n \quad J(y) - J(x) \geq \underbrace{\langle \nabla J(x) ; y - x \rangle}_{=0} + \frac{\alpha}{2} \|y - x\|^2$$

(relation (2))

$$\text{Donc } \forall y \in \mathbb{R}^n \quad J(y) \geq J(x) + \frac{\alpha}{2} \|y - x\|^2$$

En particulier donc pour $y = x^*$, $J(x^*) \geq J(x) + \frac{\alpha}{2} \|x - x^*\|^2$

or $J(x) \geq J(x^*)$ car x^* minimise J donc $x = x^*$.

Donc si $\nabla J(x) = 0$ on a alors $x = x^*$.

Ainsi on peut construire la suite $(x_k)_{k \in \mathbb{N}}$ avec $x_0 \in \mathbb{R}^n$,

$\forall k \in \mathbb{N}$ si $\nabla J(x_k) = 0$, $x_{k+1} = x_k$

si $\nabla J(x_k) \neq 0$, $\mu_k = \min_{\mu \in \mathbb{R}} (J(x_k - \mu \nabla J(x_k)))$

$$x_{k+1} = x_k - \mu_k \nabla J(x_k)$$

Étape 3: Mq $(x_k)_{k \in \mathbb{N}}$ converge vers x^* .

Par ellipticité de J , pour $k \in \mathbb{N}$ fixé, la relation (1) donne

$$\langle \nabla J(x^*) - \nabla J(x_k), x^* - x_k \rangle \geq \alpha \|x^* - x_k\|^2$$

or $\nabla J(x^*) = 0$ donc

$$\langle \nabla J(x_k), x_k - x^* \rangle \geq \alpha \|x^* - x_k\|^2$$

et par Cauchy-Schwarz,

$$\|\nabla J(x_k)\| \|x_k - x^*\| \geq \alpha \|x^* - x_k\|^2$$

▷ si $x_k = x^*$, alors la suite est stationnaire et $\forall k' \geq k$

$x_{k'} = x^*$ donc en particulier on a la convergence.

▷ si $x_k \neq x^*$, alors $\|\nabla J(x_k)\| \geq \alpha \|x^* - x_k\|$.

Reste alors à mq $\nabla J(x_k) \xrightarrow[k \rightarrow +\infty]{} 0$.

$$\|\nabla J(x_k)\|^2 = \langle \nabla J(x_k), \nabla J(x_k) \rangle$$

on a vu en (*) que pour p minimisant la fonction

$$\varphi_x: p \mapsto J(x - p \nabla J(x)) \quad (x \in \mathbb{R}^n),$$

$$\langle \nabla J(x - p \nabla J(x)); \nabla J(x) \rangle = 0.$$

Autrement dit, deux directions de descente consécutives sont orthogonales.

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

$$= \langle \nabla J(x_k) - \nabla J(x_{k+1}); \nabla J(x_k) \rangle$$

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

Donc comme on a supposé $x_k \neq x^*$, $\nabla J(x_k) \neq 0$ donc

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

On a alors à mg $\|\nabla J(x_k) - \nabla J(x_{k+1})\| \xrightarrow[k \rightarrow +\infty]{} 0$

Mg $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall k \in \mathbb{N}, (k \geq N \Rightarrow \|\nabla J(x_k) - \nabla J(x_{k+1})\| \leq \varepsilon)$

Comme $\nabla J: x \mapsto \nabla J(x)$ est continue car J est $\mathcal{C}^1$,

de plus comme J est coercive, $\exists A \in \mathbb{R}^+$ tel que

$$\forall x \in \mathbb{R}^n \setminus \overline{B(0; A)}, J(x) \geq J(x_0).$$

donc comme $(J(x_k))_{k \in \mathbb{N}}$ est décroissante $(x_k)_k \in \overline{B(0; A)}^{\text{in}}$

donc $(x_k)_k$ est dans un fermé-borné en dim. finie

donc dans un compact.

On a alors ∇J est uniformément continue sur $\overline{B(0; A)}$

Donc $\forall \varepsilon > 0, \exists \eta > 0, \forall (x, y) \in \overline{B(0; A)}, \|x - y\| \leq \eta \Rightarrow \|\nabla J(x) - \nabla J(y)\| \leq \varepsilon$

Reste à mg $\forall \eta > 0, \exists N \in \mathbb{N}, \forall k \geq N, \|x_k - x_{k+1}\| \leq \eta$.

Par ellipticité de J , $\forall k \in \mathbb{N}$

$$\begin{aligned} J(x_k) - J(x_{k+1}) &\geq \underbrace{\langle \nabla J(x_{k+1}), x_k - x_{k+1} \rangle + \frac{\alpha}{2} \|x_k - x_{k+1}\|^2}_{= \mu_k \nabla J(x_k)} \\ &= 0 \quad \text{car } \langle \nabla J(x_{k+1}), \nabla J(x_k) \rangle = 0 \end{aligned}$$

$$\text{Donc } J(x_k) - J(x_{k+1}) \geq \frac{\alpha}{2} \|x_k - x_{k+1}\|^2 \geq 0.$$

et par construction, $(J(x_k))_k$ est décroissante et

minorée donc convergente. Donc $J(x_k) - J(x_{k+1}) \xrightarrow[k \rightarrow +\infty]{} 0$

$$\text{donc } \|x_k - x_{k+1}\| \xrightarrow[k \rightarrow +\infty]{} 0$$

$$\text{Donc } \forall \eta > 0, \exists N \in \mathbb{N}, \forall k \in \mathbb{N}, (k \geq N \Rightarrow \|x_k - x_{k+1}\| \leq \eta)$$

Ainsi, par continuité uniforme de ∇J ,

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall k \in \mathbb{N}, (k \geq N \Rightarrow \|\nabla J(x_k) - \nabla J(x_{k+1})\| \leq \varepsilon)$$

$$\text{Donc } \|\nabla J(x_k) - \nabla J(x_{k+1})\| \xrightarrow[k \rightarrow +\infty]{} 0$$

$$\text{Donc } \|\nabla J(x_k)\| \xrightarrow[k \rightarrow +\infty]{} 0 \quad \text{donc } \|x_k - x^*\| \xrightarrow[k \rightarrow +\infty]{} 0$$

$$\text{Donc } x_k \xrightarrow[k \rightarrow +\infty]{} x^*$$

□

Convergence de la méthode du gradient à pas optimal.

Généralité sur les pb d'optimisation

La fonction que l'on cherche à minimiser s'appelle fonctionnelle en optimisation.

$f: V \rightarrow \mathbb{R}$ on veut la minimiser sur U

si $U=V \rightarrow$ sans contrainte

sinon U sera souvent de la forme

$$U = \{v \in V \mid \varphi_i(v) \leq 0 \quad \varphi_j(v) = 0\}$$



Contrainte
inégalité



Contrainte
égalité

• si $f(v) = \frac{1}{2} \langle Av; v \rangle - \langle b; v \rangle \rightarrow$ fonctionnelle quadratique

A sup. définie positive

$\hookrightarrow f$ strictement convexe.

• si $f(v) = \sum_{i=1}^m a_i v_i$ alors f est une fonctionnelle linéaire

théorème:

$J: v \in V \rightarrow J(v) = \frac{1}{2} a(v; v) - \underbrace{f(v)}_{\text{forme linéaire continue}}$
forme bilinéaire continue sym.
on supp. $\exists \alpha > 0 \quad a(v; v) \geq \alpha \|v\|_V^2$

$\forall U$ convexe non-vide et fermée il existe un unique élément u tq $J(u) = \inf_{v \in U} (J(v))$

$a(v; v)$ forme bilinéaire, continue, symétrique

$f: V \rightarrow \mathbb{R}$ est dite elliptique si elle est $\mathcal{C}^1$ sur V et si il existe une constante $\alpha > 0$ tq

$$\langle \nabla f(v) - \nabla f(u); v - u \rangle \geq \alpha \|v - u\|^2$$

une fonction elliptique est strictement convexe et coercive. Elle vérifie l'inégalité

$$J(v) - J(u) \geq \langle \nabla J(u); v - u \rangle + \frac{\alpha}{2} \|u - v\|^2 \quad \forall (u, v) \in V^2$$

si U est une partie convexe, fermée et non-vide de $(H; \langle \cdot, \cdot \rangle)$ et J une fonctionnelle elliptique, le pb

trouver $u \in U$ tq $J(u) = \inf_{v \in U} (J(v))$

a une solution et une seule.

$$\langle \nabla J(v) - \nabla J(u); v - u \rangle \geq \alpha \|u - v\|^2$$

$$J(v) = J(u) + \langle \nabla J(u); v - u \rangle + \|v - u\| \varepsilon(\|v - u\|)$$

$$J(v) - \langle \nabla J(v); v - u \rangle = J(u) + \langle \nabla J(u) - \nabla J(v); v - u \rangle + \|v - u\| \varepsilon(\|v - u\|)$$

$$\langle \nabla J(v) - \nabla J(u), v - u \rangle \geq \alpha \|u - v\|^2$$

$$J(v) = J(u) + \langle \nabla J(u), v - u \rangle + \|v - u\| \varepsilon(\|v - u\|)$$

$$J(v) - \langle \nabla J(v), v - u \rangle = J(u) + \langle \nabla J(u) - \nabla J(v), v - u \rangle + \|v - u\| \varepsilon(\|v - u\|)$$

Donc

$$J(u) - J(v) + \langle \nabla J(v), v - u \rangle = \langle \nabla J(v) - \nabla J(u), v - u \rangle + \|v - u\| \varepsilon(\|v - u\|)$$

$$J(u) - J(v) + \langle \nabla J(v), v - u \rangle \geq \alpha \|u - v\|^2 + \|v - u\| \varepsilon(\|v - u\|)$$

donc $J(u) - J(v) + \langle \nabla J(v), v - u \rangle$

On peut écrire Taylor avec R.T

$$J(v) = J(u) + \int_0^1 \langle \nabla J(u + t(v - u)), v - u \rangle (1 - t) dt$$

$$J(v) - J(u) = \int_u^v \langle \nabla J(t), v - t \rangle dt$$

$$\int_a^b \frac{f^{(n)}(t)}{(n-1)!} (b-t)^{n-1} dt$$

$$\int_u^v f'(t)(v-t) dt \sim \int_u^v \langle \nabla f(h), v-h \rangle dh$$

$$h = u + t(v - u) \quad dh = dt(v - u)$$

$$\int_0^1 \langle \nabla f(u + t(v - u)), v - u - t(v - u) \rangle (v - u) dt$$

$$\int_a^b f'(t)(b-t) dt = \int_0^1 f'(a+s(b-a))(b-a-s(b-a))(b-a) ds$$

$$= \int_0^1 f'(a+s(b-a))(1-s)(b-a)^2 ds$$

$$\int_a^b f'(t)(b-t) dt \quad t = a + s(b-a)$$

$$\int_a^b \frac{f^{(n+1)}(t)}{n!} (b-t)^n dt$$

$$t = a + s(b-a) \quad dt = (b-a) ds$$

$$\int_0^1 \frac{f^{(n+1)}(a+s(b-a))}{n!} (b-a-s(b-a))^n (b-a) ds$$

$$= \int_0^1 (b-a)^{n+1} (1-s)^n \frac{f^{(n+1)}(a+s(b-a))}{n!} ds$$

$$\varphi: t \mapsto f(a+t(b-a)) \quad \varphi \in \mathcal{C}^1$$

$$\varphi'(t) = \langle \nabla f(a+t(b-a)); b-a \rangle$$

$$\Rightarrow \varphi(1) - \varphi(0) = \int_0^1 \langle \nabla f(a+t(b-a)); b-a \rangle dt$$

$$f(b) - f(a) = \nearrow$$

$$J(v) - J(u) = \int_0^1 \langle \nabla J(u + t(v-u)); v-u \rangle dt$$

$$J(v) - J(u) = \langle \nabla J(u); v-u \rangle$$

$$= \int_0^1 \langle \nabla J(u + t(v-u)) - \nabla J(u); v-u \rangle dt$$

$$\text{or } \langle \nabla J(u + t(v-u)) - \nabla J(u); u + t(v-u) - u \rangle \geq \alpha \|t(v-u)\|^2$$

$$\text{donc } \langle \nabla J(u + t(v-u)) - \nabla J(u); v-u \rangle \geq \alpha t \|v-u\|^2$$

$$\text{donc } J(v) - J(u) = \langle \nabla J(u); v-u \rangle \geq \frac{\alpha}{2} \|v-u\|^2$$

et par la caractérisation, on a alors J α -convexe

~~Existence et unicité de la solution~~: admis

On se place maintenant dans le cas $V = \mathbb{R}^n$.

On suppose que J est une fonctionnelle elliptique.

Elle vérifie donc $J(v) - J(u) \geq \langle \nabla J(u); v-u \rangle + \frac{\alpha}{2} \|v-u\|^2$

$$\langle \nabla J(v) - \nabla J(u); v-u \rangle \geq \alpha \|v-u\|^2$$

Méthode du gradient à pas optimal:

Soit $u_0 \in V$, $\forall k \in \mathbb{N}$.

$$\left\{ \begin{array}{l} \langle \nabla(u_k - p_k \nabla J(u_k)) \rangle = \inf_{p \in \mathbb{R}} (J(u_k - p \nabla J(u_k))) \\ u_{k+1} = u_k + p_k \nabla J(u_k) \end{array} \right.$$

[Cours X-ENS]

$$J: \mathbb{R}^n \rightarrow \mathbb{R}$$

J fonctionnelle elliptique i.e. $J \in \mathcal{C}^1$ et $\alpha > 0$.

$$\forall (x, y) \in \mathbb{R}^n, \langle \nabla J(x) - \nabla J(y), x - y \rangle \geq \|x - y\|^2 \alpha$$

Alors J vérifie

$$\forall (x, y) \in \mathbb{R}^n \quad J(x) - J(y) \geq \langle \nabla J(y), x - y \rangle + \frac{\alpha}{2} \|x - y\|^2$$

Ng J a une borne inf et l'atteint en un unique point.

J est α -convexe et J est coercive.

$$\text{En effet } \forall x \in \mathbb{R}^n, J(x) \geq J(0) + \langle \nabla J(0), x \rangle + \frac{\alpha}{2} \|x\|^2$$

$$J(x) \geq J(0) - \|\nabla J(0)\| \|x\| + \frac{\alpha}{2} \|x\|^2$$

donc J est coercive.

$$\text{Soit } x_0 \in \mathbb{R}^n, \exists A \in \mathbb{R} \text{ tq } \forall x \in \mathbb{R}^n \text{ tq } \|x\| > A$$

$$\text{alors } \frac{\alpha}{2} \|x\|^2 + J(0) - \|\nabla J(0)\| \|x\| \geq A$$

$$\text{i.e. } J(x) \geq A$$

Donc on peut restreindre la recherche du mini.

à un ensemble borné : $\overline{B(0; A)}$ ~~et~~ et fermé

dans un e.v. de dim finie donc un compact.

Comme J est continue elle atteint ses bornes.

et on a unicité car si $(x, y) \in \mathbb{R}^n$ tq

$J(x) = \min(J) = J(y)$ alors

$$0 \geq \underbrace{\langle \nabla J(x), x-y \rangle}_{=0} + \frac{\alpha}{2} \|x-y\|^2$$

car x est un point de minimum

Donc $0 \geq \|x-y\|$ donc $x=y$.

Donc unicité de minimum.