APPROXIMATION NUMERIQUE ET OPTIMISATION

G. ALLAIRE

6 novembre 2018

CHAPITRE IV (suite)

☞ Contraintes d'inégalité.

☞ Point selle, théorème de Kuhn et Tucker.

☞ Algorithmes d'optimisation sans contraintes.

Minimisation avec contraintes d'inégalité

Soit $F(v) = (F_1(v), ..., F_M(v))$ une application de V dans $\mathbb{R}^M$. On considère

$$\inf_{v \in K} J(v) \quad \text{avec} \quad K = \{v \in V,\ F(v) \leq 0\}$$

où $F(v) \leq 0$ signifie que $F_i(v) \leq 0$ pour $1 \leq i \leq M$.

Définitions. Soit u tel que $F(u) \leq 0$.

1. On appelle ensemble des contraintes actives en u
$I(u) = \{i \in \{1, \ldots, M\}, F_i(u) = 0\}$.

2. On dit que les contraintes d'inégalité sont qualifiées en $u \in K$ si la famille $(F_i'(u))_{i \in I(u)}$ est libre.

Contraintes et variations

- Les contraintes inactives $F_i(u) < 0$ ne jouent aucun rôle.
- Pour les contraintes actives $F_i(u) = 0$ on peut faire des variations dans certaines directions (mais pas dans d'autres).
- Il existe d'autres conditions de qualification.

Théorème. Soit $u \in V$ tel que $F(u) \leq 0$. On suppose que J et les contraintes $(F_i)_{1 \leq i \leq M}$ sont dérivables en u. On suppose de plus que les contraintes sont qualifiées en u. Alors, si u est un minimum local de J sur K, il existe $\lambda_1, \ldots, \lambda_M \in \mathbb{R}^+$, appelés **multiplicateurs de Lagrange**, tels que

$$J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0, \quad \lambda_i \geq 0, \quad F_i(u) \leq 0,$$

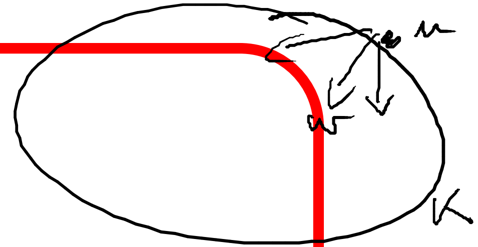
$$\lambda_i = 0 \text{ si } F_i(u) < 0.$$

Remarque. Les conditions sur les λ_i peuvent se réécrire

$$\lambda \cdot F(u) = 0, \quad \lambda \geq 0 \text{ et } F(u) \leq 0.$$

et s'appellent conditions de complémentarité.

Remarque. Pour calculer les multiplicateurs de Lagrange λ_i des contraintes actives on peut utiliser les égalités $F_i(u) = 0$ si $i \in I(u)$.



Preuve

Soit $K(u)$ le cône des directions admissibles au point u

$$K(u) = \left\{ \begin{array}{c} w \in V \text{ tel que } \exists\,(v^n)_{n\geq 0} \in K\,,\ \exists\,(\varepsilon^n)_{n\geq 0} \in \mathbb{R}_+^* \text{ vérifiant} \\ \lim\limits_{n\to+\infty} v^n = u\,,\ \lim\limits_{n\to+\infty} \varepsilon^n = 0\,,\ \lim\limits_{n\to+\infty} \dfrac{v^n - u}{\varepsilon^n} = w \end{array} \right\}$$

Rappel (Proposition 4.2.12): si u est un minimum local, alors

$$\langle J'(u), w\rangle \geq 0 \quad \forall\, w \in K(u) = K^e(u)$$

Il faut donc caractériser $K(u)$!

Si $w \in K(u)$ $\exists\, v_n \in K$ $v_n \to u$ $v_n = u + \varepsilon_n w + o(\varepsilon_n)$

$\Rightarrow F_i(v_n) \leq 0$ Taylor $\Rightarrow$ $\underbrace{F_i(u)}_{=0} + \cancel{\varepsilon_n} \langle F_i'(u), w\rangle + o(\cancel{\varepsilon_n}) \leq 0$

$o(1) \to 0$ $\Rightarrow$ $\boxed{\langle F_i'(u), w\rangle \leq 0}$

$i \in I(u)$ contrainte active

Pour les contraintes inactives $F_i(u) < 0 \Rightarrow F_i(v_n) \leq 0$ si ε_n suffisamment petit

$$\Rightarrow K(u) \subset K^e(u) = \{ w \in V \text{ tq } \langle F_i'(u), w\rangle \leq 0 \ \forall\, i \in I(u) \}$$

$K^e(u) = \{ w \in V \text{ tq } \langle F_i'(u), w \rangle \leq 0 \quad \forall i \in I(u) \}$ fermé

$K^i(u) = \{ \text{———} < 0 \text{———} \}$ ouvert

$\forall w \in \overset{i}{K}(u) \quad F_i(v_n) = \underset{\underset{0}{\parallel}}{F_i(u)} + \varepsilon_n \underbrace{\langle F_i'(u), w \rangle}_{< 0} + o(\varepsilon_n) \qquad i \in I(u)$

≤ 0 si ε_n suffisamment petit

si $F_i(u) < 0 \Rightarrow F_i(v_n) \leq 0$ par continuité si ε_n petit

$\Rightarrow \quad v_n \in K \quad \Rightarrow \quad w \in K(u) \quad \Rightarrow \quad K^i(u) \subset K(u)$

$$\boxed{K^i(u) \subset K(u) \subset K^e(u)} \qquad \overline{K^i(u)} = K^e(u)$$

Exercice vérifier $K(u)$ est fermé

$$\Rightarrow \quad \underline{\underline{K(u) = \overline{K(u)} = K^e(u)}}$$

$F'_i(u)$

Lemme de Farkas. Soient $a_1, \ldots, a_M$ une famille libre de V. On considère les ensembles

$$\mathcal{K} = \Big\{ w \in V \,, \quad \langle a_i, w \rangle \leq 0 \ \text{ pour } \ 1 \leq i \leq M \Big\} \,,$$

et

$$\hat{\mathcal{K}} = \Big\{ q \in V \,, \quad \exists\, \lambda_1, \ldots, \lambda_M \geq 0 \,,\ q = -\sum_{i=1}^{M} \lambda_i a_i \Big\} \,.$$

Alors pour tout $p \in V$, on a l'implication

$$\langle p, w \rangle \geq 0 \ \forall\, w \in \mathcal{K} \implies p \in \hat{\mathcal{K}} \,.$$

$J'(u)$

Remarque. La réciproque étant évidente, il s'agit en fait d'une équivalence.

Si Farkas est vrai, alors $J'(u) = -\sum_{i=1}^{n} \lambda_i F'_i(u)$ avec $\lambda_i = 0$ si $i \notin I(u)$

Preuve du lemme de Farkas

Supposons que la famille (a_i) est orthogonale $\quad a_i \perp a_j \quad i \neq j$

$$w \in K = \{ \langle w, a_i \rangle \leq 0 \quad \forall i \}$$

$$\langle p, w \rangle \geq 0 \quad \forall w \in K$$

$$\Longrightarrow \langle p, a_i \rangle \leq 0 \quad \forall i$$

$$\Longrightarrow p = - \sum_{i=1}^{n} \lambda_i a_i$$

$$\lambda_i \geq 0$$

Lagrangien

Définition. On appelle Lagrangien du problème la fonction

$$\mathcal{L}(v,\mu) = J(v) + \sum_{i=1}^{M} \mu_i F_i(v) = J(v) + \mu \cdot F(v) \qquad \forall (v,\mu) \in V \times (\mathbb{R}^+)^M.$$

La nouvelle variable **positive** $\mu \in (\mathbb{R}^+)^M$ est appelée multiplicateur de Lagrange pour la contrainte $F(v) \leq 0$.

Lemme. Le problème de minimisation est équivalent à

$$\inf_{v \in V,\ F(v) \leq 0} J(v) = \inf_{v \in V} \sup_{\mu \in (\mathbb{R}^+)^M} \mathcal{L}(v,\mu).$$

Preuve

$$\sup_{\mu \geq 0} \mathcal{L}(v,\mu) = J(v) + \mu \cdot F(v) = \begin{cases} +\infty & \text{si } \exists i \quad F_i(v) > 0 \\ J(v) & \text{si } F(v) \leq 0 \end{cases}$$

Stationnarité du Lagrangien

Définition du Lagrangien: $\mathcal{L}(v, \mu) = J(v) + \mu \cdot F(v) \qquad \forall (v, \mu) \in V \times (\mathbb{R}^+)^M$.

On peut réécrire le précédent théorème sous la forme suivante.

Théorème. On suppose que les contraintes sont qualifiées en u tel que $F(u) \leq 0$. Alors, si u est un minimum local, il existe $\lambda_1, \ldots, \lambda_M \geq 0$, appelés multiplicateurs de Lagrange, tels que

$$\frac{\partial \mathcal{L}}{\partial v}(u, \lambda) = 0 \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \mu}(u, \lambda) \cdot (\mu - \lambda) \leq 0 \quad \forall \mu \in (\mathbb{R}^+)^M$$

Remarque. La deuxième condition est l'inéquation d'Euler pour la **maximisation** du Lagrangien par rapport à μ dans le convexe fermé $(\mathbb{R}^+)^M$.

Preuve

Les conditions d'optimalité sont

$$J'(u) + \sum_{i=1}^{M} \lambda_i F_i'(u) = 0 \, , \quad \lambda_i \geq 0 \, , \quad F_i(u) \leq 0 \, , \quad \lambda_i = 0 \text{ si } F_i(u) < 0 \, .$$

Cette condition est bien la stationnarité du Lagrangien puisque

$$\frac{\partial \mathcal{L}}{\partial v}(u, \lambda) = J'(u) + \lambda \cdot F'(u) = 0,$$

et que l'inéquation d'Euler

$$\frac{\partial \mathcal{L}}{\partial \mu}(u, \lambda) \cdot (\mu - \lambda) = F(u) \cdot (\mu - \lambda) \leq 0 \quad \forall \mu \in (\mathbb{R}^+)^M$$

implique que $\lambda_i = 0$ si $F_i(u) < 0$.

Exemple: régularisation d'un signal

Soit A une matrice carrée d'ordre n, symétrique définie positive.

Soit $b \in \mathbb{R}^n$ et $\epsilon > 0$. On considère

$$\min_{x \in \mathbb{R}^n \text{ tel que } \|x-b\|^2 \leq \epsilon^2} Ax \cdot x.$$

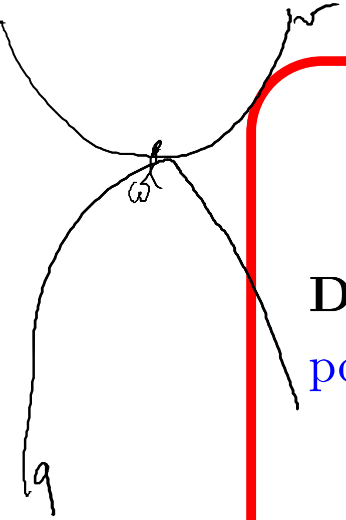
Il existe une unique solution $x^* \in \mathbb{R}^n$ qui vérifie

$$\text{soit } x^* = 0 \quad \text{ et } \|b\| < \epsilon,$$

$$\text{soit } x^* = (\mathrm{Id} + \lambda^{-1} A)^{-1} b \text{ avec } \lambda > 0 \text{ tel que } \|x^* - b\| = \epsilon.$$

Condition d'optimalité

$$2Ax^* + \underbrace{\lambda}_{>0} 2(x^* - b) = 0$$



Point selle

Définition. Soit un Lagrangien $\mathcal{L}(v,q)$. On dit que $(u,p) \in V \times P$ est un point-selle (ou col, ou min-max) de $\mathcal{L}$ sur $V \times P$ si

$$\forall\, q \in P \qquad \mathcal{L}(u,q) \leq \mathcal{L}(u,p) \leq \mathcal{L}(v,p) \quad \forall\, v \in V\ .$$

Proposition. Soit des fonctions $J, F_1, \ldots, F_M$ dérivables sur V. Soit $K = \{F(v) = 0\}$ avec $P = \mathbb{R}^M$, ou bien $K = \{F(v) \leq 0\}$ avec $P = (\mathbb{R}_+)^M$. Pour $(v,q) \in V \times P$, on considère

$$\mathcal{L}(v,q) = J(v) + q \cdot F(v).$$

Si (u,p) est un point-selle de $\mathcal{L}$ sur $V \times P$, alors $u \in K$ et u est un minimum global de J sur K. De plus, on a

$$J'(u) + \sum_{i=1}^{M} p_i F_i'(u) = 0\ .$$

(u,p) point selle

Preuve

$$J(u) + q.F(u) \leq J(u) + p.F(u) \leq J(v) + p.F(v) \quad \forall v \in V$$

$\forall q \in K = \mathbb{R}^n$ ou $\mathbb{R}^n_+$

$$q.F(u) \leq p.F(u)$$

si $K = \mathbb{R}^n \Rightarrow F(u) = 0$ si $K = \mathbb{R}^n_+$, $F(u) \leq 0$

Si $F(v) = 0$ $(v \in K)$, alors $J(u) + pF(u) \leq J(v)$

si $K = \mathbb{R}^n$ $F(u) = 0 \Rightarrow J(u) \leq J(v) \; \forall v \in K$

si $K = \mathbb{R}^n_+$ $p.F(u) = 0 \Rightarrow J(u) \leq J(v) \; \forall v \in K$

$$\Rightarrow \quad u = \arg\min_{v \in K} J(v)$$

De plus $v \to J(v) + p.F(v)$ est min en u $\Rightarrow$ $J'(u) + p.F'(u) = 0$

Théorème de Kuhn et Tucker. Soit des fonctions $J, F_1, \ldots, F_M$ convexes et dérivables sur V. Soit $K = \{F(v) \leq 0\}$, $P = (\mathbb{R}_+)^M$ et le Lagrangien

$$\mathcal{L}(v, q) = J(v) + q \cdot F(v) \quad \forall\, (v, q) \in V \times P\ .$$

Soit $u \in K$ où les contraintes sont qualifiées. Alors les 3 affirmations suivantes sont équivalentes:

1. u est un minimum global de J sur K,
2. $\exists\, p \in (\mathbb{R}_+)^M$ tel que (u, p) est un point-selle du Lagrangien $\mathcal{L}$ sur $V \times P$,
3. $\exists\, p \in (\mathbb{R}_+)^M$ tel que

$$F(u) \leq 0\,, \quad p \geq 0\,, \quad p \cdot F(u) = 0\,, \quad J'(u) + p \cdot F'(u) = 0\,.$$

Remarques.

Les conditions d'optimalités sont donc nécessaires et suffisantes dans ce cas.

Le théorème de Kuhn et Tucker ne s'applique que pour des contraintes d'inégalité.

On sait déjà que point selle $\Longrightarrow$ min global de J sur K

Preuve

Min global de J sur K $\Longrightarrow$ conditions d'optimalité $\exists\, p \geq 0$ $\quad F(u) \leq 0$

$$J'(u) + p \cdot F'(u) = 0 \quad \text{et} \quad p \cdot F(u) = 0$$

Montrons que ces conditions d'optimalité $\Longrightarrow$ (u,p) point selle

$v \longrightarrow J(v) + p \cdot F(v)$ fct convexe dont la dérivée s'annule en u

$\Longrightarrow$ u est un min global de cette fonction

$$\Longrightarrow \quad J(u) + p \cdot F(u) \leq J(v) + p \cdot F(v) \quad \forall v \in V$$

On veut $\quad \cancel{J(u)} + q \cdot F(u) \leq \cancel{J(u)} + p \cdot F(u)$

on veut $\quad \underbrace{q \cdot F(u)}_{\leq 0} \leq p \cdot F(u) = 0 \qquad \forall q \in \mathbb{R}_+^m$

$\Longrightarrow$ (u,p) point selle

Exemple essentiel

Soit A une matrice carrée d'ordre n, symétrique définie positive.

Soit B une matrice rectangulaire de taille $m \times n$ avec $m \leq n$ et $rg(B) = m$.

Soit $b \in \mathbb{R}^n$ et $c \in \mathbb{R}^m$. On considère

$$\min_{x \in \mathbb{R}^n \text{ tel que } Bx \leq c} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\}.$$

transposé de B

Il existe une unique solution $x^* \in \mathbb{R}^n$ qui vérifie

$$Ax^* - b = B^* p \quad \text{avec } p \in (\mathbb{R}^+)^m \qquad \text{et} \qquad Bx^* \leq c.$$

Comment calculer p ? Problème combinatoire !

On prédit l'ensemble I des contraintes actives avec $\{1, ..., m\} = I \cup I^c$.

On résout

$$\min_{x \in \mathbb{R}^n \text{ tel que, pour tout } i \in I, \ (Bx)_i = c_i} J(x)$$

qui admet une unique solution x_I et un unique multiplicateur de Lagrange p_I.

Exemple essentiel (fin)

On teste: si $p_I \geq 0$ et si $(Bx)_i < c_i$ pour $i \in I^c$, on a trouvé la bonne solution grâce à Kuhn-Tucker.

Sinon, on recommence avec un nouveau choix de I.

Il y a un nombre fini (mais exponentiel) de choix possibles de I !

Dualité

Définition. Soit un Lagrangien $\mathcal{L}(v,q)$ défini sur $V \times P$.

$$\text{Pour } v \in V \text{ posons } \mathcal{J}(v) = \sup_{q \in P} \mathcal{L}(v,q).$$

$$\text{Pour } q \in P \text{ posons } \mathcal{G}(q) = \inf_{v \in V} \mathcal{L}(v,q).$$

On appelle problème primal

$$\inf_{v \in V} \mathcal{J}(v)\ ,$$

et problème dual

$$\sup_{q \in P} \mathcal{G}(q)\ .$$

Exemple. Lagrangien $\mathcal{L}(v,q) = J(v) + q \cdot F(v)$ avec $P = \mathbb{R}^M$ (ou $\mathbb{R}^M_+$). Dans ce cas $\mathcal{J}(v) = J(v)$ si $F(v) = 0$ (ou $F(v) \leq 0$) et $\mathcal{J}(v) = +\infty$ sinon, tandis qu'il n'y a pas de contraintes pour le problème dual (autres que $q \in P$).

Lemme (dualité faible). On a toujours

$$\inf_{v\in V} \mathcal{J}(v) \geq \sup_{q\in P} \mathcal{G}(q).$$

Preuve: $\inf \sup \mathcal{L}(v,q) \geq \sup \inf \mathcal{L}(v,q)$.

$\mathcal{L} \geq \inf_{v} \mathcal{L}$

$\Rightarrow \sup_{q} \mathcal{L} \geq \sup_{q} \inf_{v} \mathcal{L}$

Théorème (dualité forte). Le couple (u,p) est un point-selle de $\mathcal{L}$ sur $V \times P$ si et seulement si

$$\mathcal{J}(u) = \min_{v\in V} \mathcal{J}(v) = \max_{q\in P} \mathcal{G}(q) = \mathcal{G}(p) \; .$$

Remarque. Le problème dual est souvent plus simple que le primal (pas de contraintes). Si on peut résoudre le dual, on obtient la solution du primal grâce à une minimisation sans contrainte.

Preuve

$$\inf_v \mathcal{L}(v,p) \le \mathcal{L}(u,q) \le \mathcal{L}(u,p) \le \mathcal{L}(v,p) \le \max_q \mathcal{L}(v,q)$$

$$\inf_v \mathcal{L}(v,p) = \mathcal{G}(p) \qquad \max_q \mathcal{L}(v,q) = \mathcal{J}(v)$$

$$\underline{\mathcal{G}(p) \le \mathcal{L}(u,p) \le \mathcal{J}(v)}$$

min en v

$$\mathcal{L}(u,p) \le \mathcal{G}(p) \quad (=)$$

$$\max_q \mathcal{L}(u,q) = \mathcal{L}(u,p) \le \mathcal{J}(u)$$

$$\max_q \mathcal{L}(u,q) = \mathcal{J}(u)$$

$$\Longrightarrow \quad \mathcal{L}(u,p) = \mathcal{J}(u) = \mathcal{G}(p)$$

Exemple

A $n \times n$, sym. déf. ≥ 0, B $m \times n$ avec $rg(B) = m \leq n$, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}^m$.

$$\min_{x \in \mathbb{R}^n \text{ tel que } Bx=c} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\}.$$

Fonction duale, pour $p \in \mathbb{R}^m$,

$$\mathcal{G}(p) = \min_{x \in \mathbb{R}^n} \left\{ \mathcal{L}(x,p) = \frac{1}{2} Ax \cdot x - b \cdot x + p \cdot (Bx - c) \right\}.$$

On calcule $\quad \mathcal{G}(p) = -\dfrac{1}{2} A^{-1}(b - B^* p) \cdot (b - B^* p) - p \cdot c.$

Problème dual (sans contraintes !)

$$\max_{p \in \mathbb{R}^m} \mathcal{G}(p)$$

Une fois calculé le point de maximum p^*, il reste une minimisation (sans contraintes !) pour trouver le point de minimum x^*

$$\min_{x \in \mathbb{R}^n} \left\{ \mathcal{L}(x,p^*) = \frac{1}{2} Ax \cdot x - b \cdot x + p^* \cdot (Bx - c) \right\}.$$

Algorithmes d'optimisation sans contraintes.

- ☞ Gradient à pas optimal (ordre 1).
- ☞ Gradient à pas fixe (ordre 1).
- ☞ Algorithme de Newton (ordre 2).

Remarque. Tous les algorithmes d'optimisation sont itératifs. On se limite aux algorithmes déterministes utilisant la connaissance des dérivées.

Dans tout ce qui suit, V est un espace vectoriel de dimension finie, muni d'un produit scalaire $\langle u, v \rangle$ et d'une norme associée.

Principe des algorithmes de gradient

On veut résoudre

$$\inf_{v \in V} J(v) \ .$$

Initialisation: choisir u^0 dans V.

Itérations: pour $n \geq 0$

$$u^{n+1} = u^n - \mu w^n \qquad \text{avec } \mu > 0$$

avec $w^n =$ direction de descente et $\mu =$ pas de descente. Si μ est petit, on a

$$J(u^{n+1}) = J(u^n) - \mu \langle J'(u^n), w^n \rangle + \mathcal{O}(\mu^2),$$

donc pour descendre il vaut mieux prendre w^n colinéaire à $J'(u^n)$.

Conclusion: algorithme de gradient

$$u^{n+1} = u^n - \mu^n \, J'(u^n) \qquad \text{avec } \mu^n > 0.$$

Gradient à pas optimal

On veut résoudre

$$\inf_{v \in V} J(v) \ .$$

Initialisation: choisir u^0 dans V.

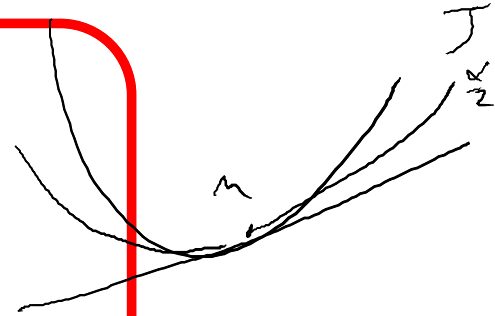
Itérations: pour $n \geq 0$

$$u^{n+1} = u^n - \mu^n \, J'(u^n) \ ,$$

où $\mu^n \in \mathbb{R}$ est choisi à chaque étape tel que

$$J(u^{n+1}) = \inf_{\mu \in \mathbb{R}^+} J\big(u^n - \mu \, J'(u^n)\big) \ .$$

μ^n est appelé pas de descente optimal.



Théorème Soit J une fonction différentiable. On suppose que J est α-convexe, pour $\alpha > 0$,

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

et que J' est Lipschitzien sur tout borné de V,

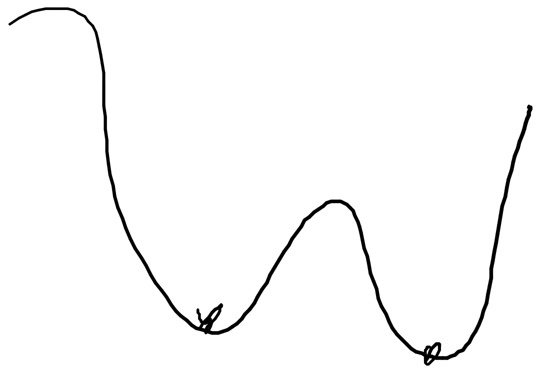
$$\forall\, M > 0\,, \quad \exists\, C_M > 0\,, \quad \|v\| + \|w\| \leq M \;\Rightarrow\; \|J'(v) - J'(w)\| \leq C_M \|v - w\|\,.$$

Alors l'algorithme de gradient à pas optimal converge: $\forall\, u^0$, la suite (u^n) converge vers la solution u.

Remarque. Si J n'est pas α-convexe:

- ☞ l'algorithme peut ne pas converger (il oscille entre plusieurs solutions),
- ☞ l'algorithme peut converger vers un minimum local,
- ☞ le résultat de l'algorithme peut varier selon l'initialisation.

J pas convexe



Preuve

$f(\mu) = J(u^n - \mu J'(u^n))$ fct convexe strict.

μ_n vérifie $f'(\mu_n) = 0$

$$\parallel$$

$$< J'(u_n), J'(u^n - \mu_n J'(u_n)) > = 0$$

$$\Leftrightarrow \quad < J'(u_n), J'(u^{n+1}) > = 0$$

α - convexité

$$J(u_n) \geq J(u^{n+1}) + \underbrace{< J'(u^{n+1}), \underset{-\mu_n J'(u_n)}{\underset{\parallel}{u^{n+1} - u_n}} >}_{\substack{\parallel \\ 0}} + \frac{\alpha}{2} \|u^{n+1} - u_n\|^2$$

$$J(u_n) - J(u_{n+1}) \geq \frac{\alpha}{2} \|u^{n+1} - u_n\|^2$$

la suite $J(u_n)$ décroît, donc CV $\Rightarrow \|u^{n+1} - u_n\| \longrightarrow 0$

$$\|J'(u_n)\|^2 + \|J'(u_{n+1})\|^2 + 2\underset{\substack{\parallel \\ 0}}{< J'(u_n), J'(u_{n+1}) >} = \| J'(u^{n+1}) - J'(u_n) \|^2 \leq C \|u^{n+1} - u^n\|^2$$

$$\Rightarrow \quad \|J'(u_n)\|^2 \leq C \|u^{n+1} - u_n\|^2 \longrightarrow 0$$

$$\Rightarrow \quad J'(u_n) \longrightarrow 0$$

α convexité $\Rightarrow$ $\langle J'(u) - J'(v), u - v \rangle \geq \alpha \|u - v\|^2$

$v = u^n$ et $u = \lim_{n \to +\infty} u^n$ pour une sous-suite

$J'(u) = \lim J'(u_n) = 0$

$\|J'(u_n)\| \cdot \|u - u_n\| \geq \langle \underbrace{J'(u)}_{= 0} - J'(u_n), u - u_n \rangle \geq \alpha \|\cancel{u} - u_n\|^2$

$\Rightarrow$ $\alpha \|u - u_n\| \leq \|J'(u_n)\| \longrightarrow 0$

$u_n \to u$ toute la suite

$J'(u) = 0 \Rightarrow u$ pt de min global

Variante: gradient à pas fixe

On veut résoudre

$$\inf_{v \in V} J(v) \ .$$

Initialisation: choisir u^0 dans V. Itérations: pour $n \geq 0$

$$u^{n+1} = u^n - \mu \, J'(u^n) \ ,$$

Théorème. On suppose que J est α-convexe, différentiable et que J' est Lipschitzien sur V. Alors, si $\mu > 0$ est suffisamment petit, l'algorithme de gradient à pas fixe converge : $\forall \, u^0$, la suite (u^n) converge vers la solution u.

Preuve.

Preuve (suite)

Algorithmes de Newton (ordre 2)

Idée principale: si J'' est une matrice définie positive

$$J(w) \approx J(v) + J'(v) \cdot (w - v) + \frac{1}{2} J''(v)(w - v) \cdot (w - v),$$

dont le minimum est $w = v - \left(J''(v)\right)^{-1} J'(v)$.

Algorithme: $u^{n+1} = u^n - \left(J''(u^n)\right)^{-1} J'(u^n)$.

☞ Converge plus vite si u^0 est proche du minimum u

$$\|u^{n+1} - u\| \leq C \|u^n - u\|^2 \, .$$

☞ Nécessite d'inverser un système linéaire de matrice $J''(u^n)$.

☞ En fait c'est un algorithme de recherche de zéro de J'...