Sur les variantes du théorème de Kantorovich pour la méthode de Newton

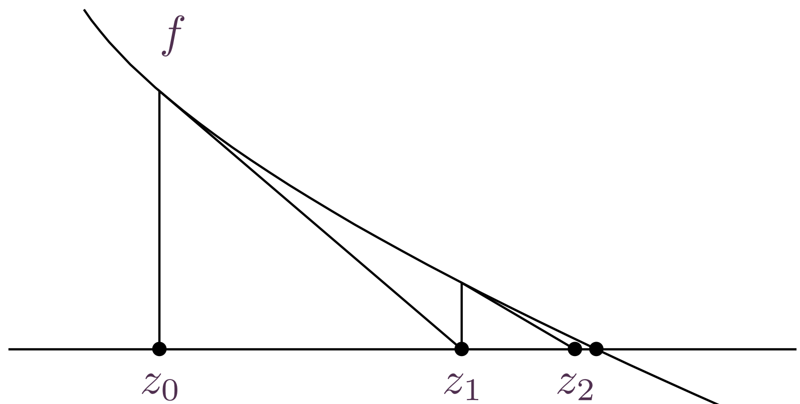
par Joelle Saadé

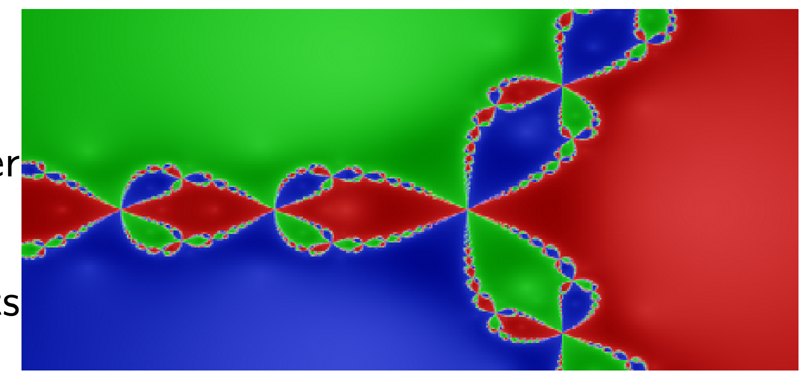
Laboratoire XLIM, UMR 7252 CNRS
Université de Limoges
123, avenue Albert Thomas
87060 Limoges Cedex, France

Travail en collaboration avec Grégoire Lecerf

4 novembre 2015

Suite de Newton pour une fonction f et une valeur initiale z_0 : $z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}$.





Il est difficile de supposer qu'un point va converger vers un zéro ou non.

Dans la pratique on a besoin de critères suffisants rapides.

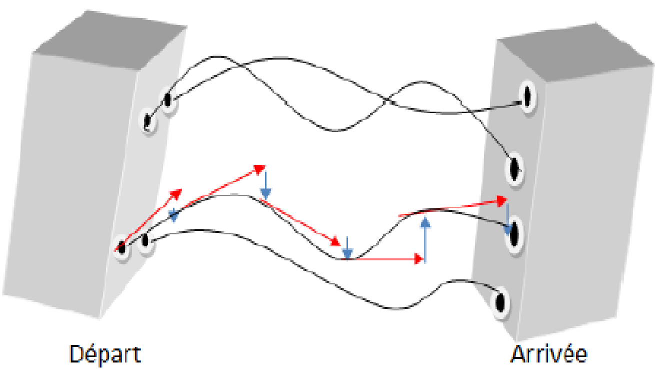
$$\left\{\begin{array}{lcl} P_1(x,y) & = & x^2-y^2+x+3=0 \\ P_2(x,y) & = & x^2+2\,x\,y+7\,y^2-8\,y+2=0 \end{array}\right. \quad \text{et} \quad \left\{\begin{array}{lcl} Q_1(x,y) & = & x^2-1=0 \\ Q_2(x,y) & = & y^2-1=0 \end{array}\right.$$

Homotopie linéaire, par exemple :

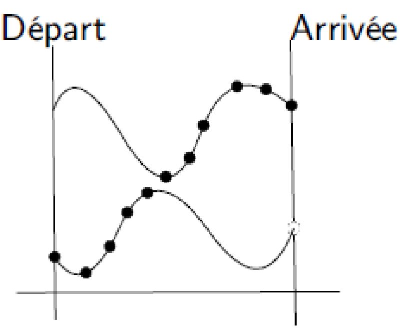
$$H(x,t)=(1-t)\,P(x)+t\,Q(x), t\in[0,1].$$

Solutions de Q $\longrightarrow$ Solutions de P de $t=1$ jusqu'à $t=0$.

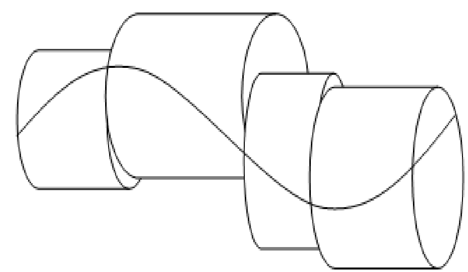
Le cas générique



Risque d'erreur



Certification



(Lecerf, Saadé, 2015, Conséquence du théorème de Wang 1999) :

→ **(Fenyo, Ostrowski, Ortega, Gragg et Tapia, 1954-1974, Théorème de Kantorovich, Version : condition Lipschitzienne classe $\mathcal{C}^1$)**

→ **(Ezquerro, González et Hernández, 2013, Théorème de Kantorovich, Version : condition Lipschitzienne classe $\mathcal{C}^l$)**

→ **(Giusti, Lecerf, Salvy, Yakoubsohn, 2007, α-théorie de Smale 1986)**

Notation. $\mathbb{X}, \mathbb{Y}$ *des espaces de Banach sur* $\mathbb{R}$ *ou* $\mathbb{C}$.

$\Omega \subseteq \mathbb{X}$ *un ouvert,* $f \in \mathcal{C}^1(\Omega, \mathbb{Y})$, $x_0 \in \Omega$ *et* $\|f\| := \sup_{x \in \mathbb{X}, \|x\|=1} \|f(x)\|$.

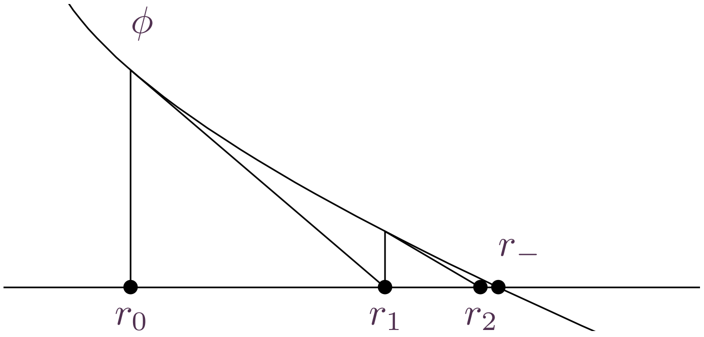
Théorème. (Lecerf, Saadé, 2015, D'après Kantorovich) *Supposons* $Df(x_0)$ *inversible. Soit* $\beta \geqslant \|Df(x_0)^{-1} f(x_0)\|$, *et* $L\colon [0, R] \to \mathbb{R}_{\geqslant 0}$ *continue croissante et positive, qui satisfait* $B(x_0, R) \subseteq \Omega$ *et*

$$\|Df(x_0)^{-1}(Df(b) - Df(a))\| \leqslant L(r)\,\|b-a\|, \forall r \in [0, R] \text{ et } \forall a, b \in \bar{B}(x_0, r).$$

Soit $\phi(r) = \beta - r + \int_0^r L(s)\,(r-s)\,\mathrm{d}s$.

On suppose que ϕ *admet un unique zéro*

r_- *dans* $[0, R)$, *et que* $\phi(R) \leqslant 0$.



Alors la suite de Newton $r_0 = 0$, $r_{k+1} = r_k - \frac{\phi(r_k)}{\phi'(r_k)}$ *est définie dans* $[0, r_-]$, *et converge vers* r_-.

La suite $x_{k+1} = x_k - Df(x_k)^{-1} f(x_k)$ *est définie dans* $\bar{B}(x_0, r_-)$, *et converge vers l'unique zéro* ζ *de f dans* $B(x_0, R)$.

De plus, on a $\|\zeta - x_k\| \leqslant r_- - r_k$ *et* $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$.

Lemme classique. *Soit* $A\colon \mathbb{X}\to\mathbb{X}$ *un opérateur linéaire tel que* $\|A\|<1$*. Alors* $\mathrm{Id}-A$ *est inversible, d'inverse* $(\mathrm{Id}-A)^{-1}=\sum_{k\geqslant 0}A^k$*, et on a* $\|(\mathrm{Id}-A)^{-1}\|\leqslant(1-\|A\|)^{-1}$*.*

Lemme. (Lecerf, Saadé, 2015) *La condition Lipschitzienne classique du théorème*

$$\|Df(x_0)^{-1}(Df(b)-Df(a))\|\leqslant L(r)\,\|b-a\|, \forall r\in[0,R]\,\mathrm{et}\,\forall a,b\in\bar{B}(x_0,r)$$

est equivalente à la condition de Wang (1999): pour tout $[a,b]\subset B(x_0,R)$ *tel que* $\|a-x_0\|+\|b-a\|\leqslant R$*,*

$$\|Df(x_0)^{-1}(Df(b)-Df(a))\|\;\leqslant\;\int_{\|x_0-a\|}^{\|x_0-a\|+\|b-a\|}L(s)\,\mathrm{d}s.$$

Notation. *Pour tout* $f\in\mathcal{C}^\ell(\Omega,\mathbb{Y})$*,* $a,b\in\Omega$*, et* $l\in\{0,\ldots,\ell\}$*, on note*

$$R_l(f;a,b)=f(b)-\sum_{k=0}^{l}D^kf(a)\,\frac{(b-a)^k}{k!},$$

le reste de Taylor de f *en* b *à l'ordre* l*, au voisinage de* a*.*

Lemme de majoration. *Pour tout* $[a,b]\subset B(x_0,R)$ *tel que* $\|a-x_0\|+\|b-a\|\leqslant R$*, on a :*

$$\|R_1(Df(x_0)^{-1}f;a,b)\|\;\leqslant\;R_1(\phi;\|a-x_0\|,\|a-x_0\|+\|b-a\|).$$

- $(r_k)_{k\geqslant 0}$ converge vers r_-: $\phi''(r)=L(r)$ dans $[0,R]$, $\phi'(0)=-1$, ϕ convexe, décroissante.

On veut prouver par récurrence : $\|x_{k+1}-x_k\|\leqslant r_{k+1}-r_k$ pour tout $k\geqslant 0$.

- $k=0$: $\|x_1-x_0\|=\|Df(x_0)^{-1}f(x_0)\|\leqslant \beta=r_1-r_0$.
- Supposons le résultat vrai pour $k\geqslant 0$, on a

$$\|x_k-x_0\| = \sum_{i=0}^{k-1}\|x_{i+1}-x_i\| \leqslant \sum_{i=0}^{k-1}(r_{i+1}-r_i) = r_k-r_0 = r_k \leqslant r_-.$$

- On décompose :

$$\|x_{k+1}-x_k\|=\|Df(x_k)^{-1}f(x_k)\|=\|Df(x_k)^{-1}Df(x_0)\|\,\|Df(x_0)^{-1}f(x_k)\|.$$

Majoration de $\|Df(x_k)^{-1}Df(x_0)\|$:

- On utilise la condition de Wang et la majoration du reste de Taylor :

$$\|Df(x_0)^{-1}(Df(x_k)-Df(x_0))\|\leqslant 1+\phi'(r_k)<1,$$

- On déduit que $Df(x_k)$ est inversible et que

$$\|Df(x_k)^{-1}Df(x_0)\| \leqslant \frac{1}{1-\|Df(x_0)^{-1}(Df(x_k)-Df(x_0))\|}\leqslant \frac{-1}{\phi'(r_k)}.$$

Majoration de $\|Df(x_0)^{-1}\, f(x_k)\|$:

- Par définition de x_k :

$$f(x_k) \;=\; f(x_{k-1}) + Df(x_{k-1})\,(x_k - x_{k-1}) + R_1(f; x_{k-1}, x_k) \;=\; R_1(f; x_{k-1}, x_k).$$

- On combine l'hypothèse de récurrence $\|x_k - x_{k-1}\| \leqslant r_k - r_{k-1}$, la définition de r_k, et on applique le lemme de majoration du reste de Taylor :

$$\begin{aligned} \|R_1(Df(x_0)^{-1}\, f; x_{k-1}, x_k)\| &\leqslant R_1(\phi; r_{k-1}, r_k) \\ &= \phi(r_k) - \phi(r_{k-1}) - \phi'(r_{k-1})\,(r_k - r_{k-1}) \\ &= \phi(r_k). \end{aligned}$$

- On déduit $\|Df(x_0)^{-1}\, f(x_k)\| \leqslant \phi(r_k)$.

Ceci démontre l'hypothèse de récurrence pour $k+1$.

On déduit facilement :

- La suite $(x_k)_{k \geqslant 0}$ converge vers un zéro ζ de f.

La preuve de l'unicité suit des arguments dans la même veine.

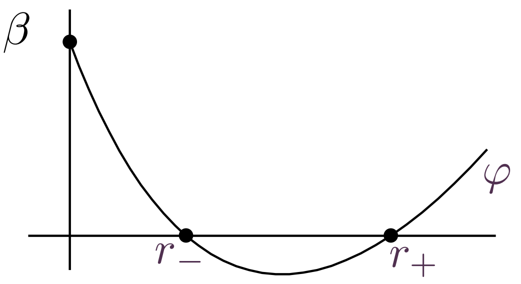
Corollaire. (Kantorovich, 1948) *Soit $f \in \mathcal{C}^1(\Omega, \mathbb{Y})$, $x_0 \in \Omega$ tel que $Df(x_0)$ inversible. Soient β, λ tels que*

$$\beta \geqslant \|Df(x_0)^{-1} f(x_0)\|,\ 0 < \beta\lambda < 1/2,\ B(x_0, r_+) \subset \Omega, \text{ et}$$

$$\|Df(x_0)^{-1}(Df(b) - Df(a))\| \leqslant \lambda \|b - a\|,$$

$\forall a, b \in B(x_0, r_+)$

où $r_- \leqslant r_+$ sont les deux zéros de $\varphi(r) = \lambda r^2/2 - r + \beta$.



Alors, la suite de Newton $r_0 = 0$, $r_{k+1} = r_k - \frac{\varphi(r_k)}{\varphi'(r_k)}$ est bien définie dans $[0, r_-]$, et converge vers r_-. La suite $(x_k)_{k \geqslant 0}$ définie par $x_{k+1} = x_k - Df(x_k)^{-1} f(x_k)$ est bien définie dans $\bar{B}(x_0, r_-)$ et converge vers l'unique zéro ζ de f dans $B(x_0, r_+)$.

De plus, on a $\|\zeta - x_k\| \leqslant r_- - r_k$ et $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$.

Démonstration. On applique le théorème précedent avec $L(r) = \lambda$, $R = r_+$, et $\phi = \varphi$. □

Avertissement. $f(x) = x^3 + 2x^2 + 4x + 1$ et $x_0 = 0$ dans $\mathbb{X} = \mathbb{Y} = [-1, 1]$.

Avec $\beta = \frac{1}{4}$, $\lambda = \frac{10}{4}$ et $\beta\lambda = \frac{10}{16} > 1/2$, le critère échoue pour des polynômes de degré trois.

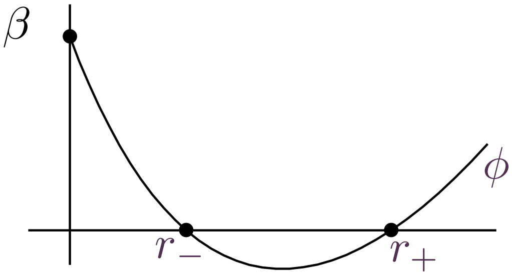
Corollaire. (Ezquerro, González et Hernández, 2013) *$f \in \mathcal{C}^\ell$, $\beta \geqslant \|Df(x_0)^{-1}\ f(x_0)\|$, $\gamma_i \geqslant \|Df(x_0)^{-1}\ D^i f(x_0)\|$, pour $i \in \{2, \ldots, \ell\}$ et L_ℓ: $[0, R] \to \mathbb{R}_{\geqslant 0}$ croissante :*

$$\|Df(x_0)^{-1}(D^\ell f(b) - D^\ell f(a))\| \leqslant L_\ell(r)\,\|b-a\|, \ \forall\, r \in [0, R] \text{ et } a, b \in \bar{B}(x_0, r).$$

On suppose que

$$\phi_\ell(r) = \beta - r + \gamma_2 \frac{r^2}{2!} + \cdots + \gamma_\ell \frac{r^\ell}{\ell!} + \int_0^r L_\ell(s) \frac{(r-s)^\ell}{\ell!}\,\mathrm{d}s$$

admet $0 < r_- < r_+$ comme premiers zéros positifs.



Alors, la suite de Newton $r_0 = 0$, $r_{k+1} = r_k - \frac{\phi_\ell(r_k)}{\phi'_\ell(r_k)}$ est bien définie dans $[0, r_-]$, et converge vers r_-. La suite $(x_k)_{k\geqslant 0}$ définie par $x_{k+1} = x_k - Df(x_k)^{-1}\ f(x_k)$ est bien définie dans $\bar{B}(x_0, r_-)$et converge vers l'unique zéro ζ de f dans $B(x_0, r_+)$.

$\ell = 1$: $\|Df(x_0)^{-1}(Df(b) - Df(a))\| \leqslant K_1\|b-a\|$ ... ϕ de degré 2.

$\ell = 2$: $\|Df(x_0)^{-1}(D^2 f(b) - D^2 f(a))\| \leqslant K_2\|b-a\|$ ... ϕ de degré 3.

Exemple. $f(x) = x^3 + 2x^2 + 4x + 1$ et $x_0 = 0$ dans $\mathbb{X} = \mathbb{Y} = [-1, 1]$, avec $\ell = 3$, $\beta = \frac{1}{4}$, $\gamma_2 = \frac{3}{2}$, $L_3(r) = \frac{3}{2}$ et $\phi(r) = \frac{1}{4} - r + \frac{5}{4}r^2 + \frac{1}{4}r^3$ de discriminant positif, le critère s'applique.

Corollaire. (Giusti, Lecerf, Salvy, Yakoubsohn, 2007)

$f \in \mathcal{C}^\infty$, $\beta \geqslant \|Df(x_0)^{-1} f(x_0)\|$, $\gamma_i \geqslant \|Df(x_0)^{-1} D^i f(x_0)\|$, pour $i \geqslant 2$.

Soit $\phi_\infty \colon [0, R] \to \mathbb{R}$, $\phi_\infty(r) = \beta - r + \sum_{l \geqslant 2} \gamma_l \frac{r^l}{l!}$.

Supposons que ϕ_∞ admet un unique zéro $r_- \geqslant 0 \in [0, R)$, tel que $B(x_0, R) \subseteq \Omega$ et $\phi_\infty(R) \leqslant 0$.

Alors, les trois lignes de convergence...

Démonstration. $L(r) = \phi''_\infty(r) = \sum_{l \geqslant 2} \gamma_l \frac{r^{l-2}}{(l-2)!}$ □

Cas particulier de l'α-théorème de Smale (1986) : Xinghua Wang (1990)

- On considère $\gamma \geqslant \left\| Df(x_0)^{-1} \frac{D^l f(x_0)}{l!} \right\|^{\frac{1}{l-1}}$ pour $l \geqslant 2$, et on prend $\gamma_l = l!\, \gamma^{l-1}$,
- $R < 1/\gamma$,
- La condition devient $\alpha = \beta \gamma < 3 - 2\sqrt{2}$.

Alors $\phi_\infty(r) = \beta - r + \frac{\gamma r^2}{1 - \gamma r} = \frac{\beta - (\alpha + 1) r + 2 \gamma r^2}{1 - \gamma r}$, et les conditions du corollaire sont vérifiées.

Par ailleurs, on dispose de formules explicites simples pour borner la convergence.

→ **1948 :** première version du critère de convergence, dû à KANTOROVICH.

Conditions : $f \in \mathcal{C}^2$, majoration sur $\|f(x_0)\|$, $\|f(x_0)^{-1}\|$ et $\|D^2 f\|$.

→ **1954–1974 :** FENYO, OSTROWSKI, ORTEGA, GRAGG et TAPIA.

Conditions : $f \in \mathcal{C}^1$, majoration sur $\|f(x_0)\|$, $\|f(x_0)^{-1}\|$ et $\|Df(x) - D\,f(y)\| \leqslant K\,\|x - y\|$

→ **1986 :** Apparition de l'α-théorie de SMALE pour des fonctions analytiques.

→ **1990 :** XINGHUA WANG

Amélioration des conditions de α-théorie de SMALE, en introduisant une fonction majorante.

→ **1993 :** ZHENG DA HUANG

Conditions : $f \in \mathcal{C}^2$, majoration sur $\|f(x_0)\|$, $\|f(x_0)^{-1}\|$, $\|f''(x_0)\|$ et $\|D^2 f(x) - D^2 f(y)\| \leqslant K\,\|x - y\|$, extension en degré 3.

→ **1999 :** XINGHUA WANG

Conditions : $f \in \mathcal{C}^1$, une forme générale pour la condition Lipschitzienne, introduit la fonction générale L continue croissante.

Donne une relation explicite entre l'α-théorie et Kantorovich.

→ **2007 :** Giusti, Lecerf, Salvy, Yakoubsohn.

Revisitent l'α-théorie en terme de séries majorantes.

→ **2013 :** Ezquerro, González et Hernández

Conditions : $f \in \mathcal{C}^l; \ell \geqslant 2$ majoration sur $\|Df(x_0)^{-1}\|, \|f(x_0)\|, \|D^i f(x_0)\|$ et $\|(D^\ell f(b) - D^\ell f(a))\|$

→ **2015 :** Lecerf, Saadé, https://hal.archives-ouvertes.fr/hal-01196890

Nos contributions sont :

Conséquence du théorème de Wang 1999 :

→ **(Fenyo, Ostrowski, Ortega, Gragg et Tapia, 1954-1974, Théorème de Kantorovich, Version : condition Lipschitzienne classe $\mathcal{C}^1$)**

→ **(Ezquerro, González et Hernández, 2013, Théorème de Kantorovich, Version : condition Lipschitzienne classe $\mathcal{C}^l$)**

→ **(Giusti, Lecerf, Salvy, Yakoubsohn, 2007, α-théorie de Smale 1986)**

Merci de votre attention