

Leçon 219 - Extrema : existence, caractérisation, recherche. Exemples et applications

Vincent LOUATRON, encadré par Paul BLOCHAS

Table des matières

1	Introduction	2
2	Remerciements	3
3	Régularité et conditions d'existence d'extrema	4
3.1	Extrema globaux et continuité	4
3.2	Extrema locaux et différentiabilité	5
4	Convexité	7
4.1	Généralités sur les fonctions convexes	7
4.2	Espaces de Hilbert	7
5	Optimisation sous contrainte et extrema liés	10
6	Méthodes numériques	13
7	Etude du module des fonctions holomorphes	17
A	Figures	18
A.1	Graphe de $f : x \mapsto x^3 - x$	18
A.2	Graphe de $f : x \mapsto x^2 + y^2$	19
A.3	Graphe de $f : x \mapsto x^2 - y^2$	19
A.4	Graphe de $f : x \mapsto x^3 + y^3$	20
	Références	20

1 Introduction

La recherche d'extrema de fonctions est à la base des problèmes d'optimisation en analyse. Les premières motivations pour minimiser ou maximiser une fonction se retrouvent dans la vie quotidienne : trouver le chemin le plus court, calculer un coût minimal pour faire des économies, etc... Plus concrètement, on peut par exemple penser au problème du voyageur de commerce, qui consiste à trouver le plus court chemin reliant un certain nombre de points. Malgré sa formulation relativement simple en langage courant, ce problème d'optimisation se révèle être non trivial. Ce sont tout autant de motivations provenant de la vie quotidienne qui ont poussé les mathématiciens à pousser leurs recherches dans les domaines liés à l'optimisation (mathématiques fondamentales, mais aussi appliquées : algorithmique génétique, intelligence artificielle...).

L'étude théorique des méthodes numériques sont d'une importance cruciale lorsque l'on souhaite effectuer des calculs pratiques comme le calcul de solutions approchées d'équations ou bien le calcul de valeurs propres d'une matrice donnée. Par exemple, il est intéressant de disposer de critères précis de convergence pour les méthodes de descente de gradient, qui permettent de trouver les points critiques fonctions suffisamment régulières. Le calcul de valeurs propres peut quant à lui intervenir pour étudier le conditionnement de certaines matrices, comme par exemple la matrice associée à la discrétisation en dimension 1 de l'opérateur laplacien Δ . Le conditionnement d'une matrice permet de majorer l'erreur relative commise lors de la résolution numérique d'équations linéaires, ce qui en fait un outil approprié pour étudier la convergence de schémas numériques.

Plusieurs approches sont possibles pour étudier les extrema d'une fonction donnée. Dans le cas général il est rare d'obtenir un condition nécessaire et suffisante donnant directement les solutions au problème d'optimisation. On peut cependant passer d'une part par les conditions suffisantes pour vérifier l'existence d'une solution, puis par des conditions nécessaires permettant d'exhiber des candidats pour la ou les solutions au problème. La théorie d'optimisation est naturellement reliée à celle du calcul différentiel, mais aussi à des théories apparues historiquement plus tard telles que l'étude des équations aux dérivées partielles.

La leçon 219 est ainsi reliée aux leçons de calcul différentiel que sont les leçons 214 et 215, et à la leçon

222 sur les équations aux dérivées partielles. On peut également citer le lien avec la convexité (229, 253) et plus simplement à la dérivabilité en dimension 1, qui apparaît dans la leçon 228. Les espaces de Hilbert (213) sont également en lien avec cette leçon. Les problèmes d'optimisation amenant à étudier des courbes (et inversement : l'étude de certaines courbes peut requérir une recherche d'extrema), on peut aussi citer la leçon 267.

La première partie présentée traite de l'existence et de la recherche d'extrema, en particulier de certaines conditions nécessaires, d'autres suffisantes, servant à trouver les extrema d'une fonction. On y fait naturellement le lien avec la différentiabilité. Par la suite, on développe le point de vue de la convexité, qui est un cadre agréable en optimisation dans les espaces vectoriels normés. On présente notamment un théorème d'optimisation dans le cadre hilbertien, qui fait le lien via le théorème de projection avec la théorie des espaces de Hilbert. En troisième partie, on présente le théorème des extrema liés, première approche à la théorie de l'optimisation sous contrainte. La quatrième partie comprend l'étude de méthodes numériques telles que la méthode du gradient. Enfin, on évoque l'étude (en module) d'extrema de fonctions à variables complexes avec notamment un théorème fondamental lorsqu'il s'agit d'étudier les extrema de fonctions holomorphes : le principe du maximum.

2 Remerciements

Je tiens à remercier mon encadrant Paul Blochas pour m'avoir donné de précieux conseils et pour avoir effectué plusieurs relectures du mémoire. La plupart des notes de bas de page sont le fruit de ses remarques faites lors du passage à l'oral ainsi que par échanges de mails. Merci également à Hugo Chemin avec qui j'ai préparé la leçon.

Dans toute la suite, on se place dans le cadre d'un espace vectoriel normé E . U désignera un ouvert de E et $f : U \subset E \rightarrow \mathbb{R}$ une application à valeurs réelles définie sur U .¹

3 Régularité et conditions d'existence d'extrema

Dans cette partie, l'objectif va être de présenter un panorama de la notion d'extremum, et en particulier d'exhiber les premières conditions nécessaires et conditions suffisantes d'existence d'extrema.

3.1 Extrema globaux et continuité

Définition 1. Soient $f : U \rightarrow \mathbb{R}$ et $a \in U$. On dit que $f(a)$ est un *minimum* (resp. *maximum*) *global* pour f si pour tout $x \in U$, $f(x) \geq f(a)$ (resp $f(x) \leq f(a)$).

Exemple 2. $f : x \mapsto x^2$ admet 0 pour *minimum global*.

L'étude des points critiques développée dans la suite s'appuie largement sur les développements limités et le calcul différentiel. Le principe même d'un développement limité, c'est à dire une simplification locale d'une fonction suffisamment régulière en un polynôme, montre bien que ces derniers sont un cas particulier important pour l'étude d'extrema de fonctions. Pour donner un exemple simple, si on peut montrer qu'une fonction f est équivalente en un point c à $\alpha(x - c)^{2n}$ avec $n \in \mathbb{N}$ et $\alpha \in \mathbb{R}^*$ alors on peut directement conclure que f a un extremum en c .

Avant d'entamer des considérations liées aux points critiques, on se propose de donner une condition suffisante d'existence d'un extrema, utilisable dans un cadre relativement large :

Propriété 3. *L'image de tout compact de E par une fonction continue est un compact. En particulier, toute fonction continue sur un compact est bornée et atteint ses bornes.*

Exemple 4. Si K_1, K_2 sont deux compacts de E , alors il existe $(x_1, x_2) \in K_1 \times K_2$ tels que $d(x_1, x_2) = d(K_1, K_2)$.

Corollaire 5. *En dimension finie, les normes sont équivalentes.*

1. Nous avons initialement omis le cadre général d'un ouvert U , ce qui est dommageable car en pratique (et dans certains théorèmes importants comme celui des extrema liés) c'est dans ce cadre que l'on se place. C'est également indispensable pour pouvoir parler d'extremum "local", notamment pour des fonctions qui ne seraient pas définies sur tout E .

Ce corollaire montre bien que le cadre dans lequel la propriété 3 s'applique est très général (pas d'hypothèse de régularité requise). En pratique, ce théorème sert notamment de point de départ pour beaucoup de problèmes d'optimisation car il permet de s'assurer de l'existence d'extrema à peu de frais.

3.2 Extrema locaux et différentiabilité

On donne maintenant une première manière de chercher les extrema d'une fonction moyennant des hypothèses de régularité.

Définition 6. Soient $f : U \rightarrow \mathbb{R}$ et $a \in U$. On dit que $f(a)$ est un **minimum** (resp. **maximum**) **local** pour f si il existe $r > 0$ tel que pour tout $x \in \mathcal{B}(a, r)$, $f(x) \geq f(a)$ (resp $f(x) \leq f(a)$).

Exemple 7. L'application $x \mapsto x^3 - x$ admet un **maximum** (resp. **minimum**) local mais non global en $-1/\sqrt{3}$ (resp. $1/\sqrt{3}$). (c.f figure [A.1](#))

Définition 8. Soient $f : U \rightarrow \mathbb{R}$ différentiable en $a \in U$. Alors a est un **point critique** de f si $Df(a) = 0$.

Propriété 9. Si f est différentiable en $a \in U$ et a est un **extremum local** de f , alors $Df(a) = 0$.

Preuve. Ce résultat est fondamental, et sa preuve est relativement simple. On remarque en effet que pour tout $h \in E$, $f(a+h) = f(a) + Df(a).h + o(h)$ et $f(a-h) = f(a) - Df(a).h + o(h)$ par linéarité de $Df(a)$. De plus, puisque a est un **extremum**, $f(a)$ est ou bien plus grand que ces deux termes, ou bien plus petit. Dans les deux cas on en déduit $0 \leq Df(a).h \leq 0$ pour tout h , soit $Df(a) = 0_{\mathcal{L}(E, \mathbb{R})}$. \square

Exemple 10. L'application $(x, y) \in \mathbb{R}^2 \mapsto x^2 + y^2 \in \mathbb{R}$ admet un **point critique** en $(0, 0)$ qui est un **minimum**. (c.f figure [A.2](#))

Contre-exemple 11. L'application $(x, y) \in \mathbb{R}^2 \mapsto x^2 - y^2 \in \mathbb{R}$ admet un **point critique** en $(0, 0)$ mais qui n'est pas un **extremum** (c.f figure [A.3](#)).

L'exemple précédent est un exemple de ce qu'on appelle un *point col*. On en donne une première définition :

Définition 12. On dit que $f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) \in \mathbb{R}$ admet un **point col** en (x_0, y_0) si $x \mapsto f(x, y_0)$ admet un **minimum local** en x_0 et $y \mapsto f(x_0, y)$ admet un **maximum local** en y_0 , ou bien si $x \mapsto f(x, y_0)$ admet un **maximum local** en x_0 et $y \mapsto f(x_0, y)$ admet un **minimum local** en y_0 .

Remarque 13. De manière plus générale, on peut parler de point col lorsqu'il existe deux directions telles que f admette un minimum suivant l'une et un maximum suivant l'autre. Le théorème qui suit permet de s'affranchir d'un tel formalisme et permet de définir la notion de point col via la hessienne dans le cas où cette dernière est bien définie.²

Théorème 14. Soient f deux fois différentiable en $a \in U$. On suppose que a est un point critique de f . On suppose $\dim(E) < +\infty$. Alors :

1. Si f admet un maximum (resp. minimum) local, alors la hessienne $D^2f(a)$ est négative (resp. positive) en tant que forme quadratique.
2. Si $D^2f(a)$ est définie négative (resp. définie positive), alors a est un maximum (resp. minimum) local strict de f .
3. Si $D^2f(a)$ est non dégénérée et a deux valeurs propres de signes différents, alors a est un **point col**.

Remarque 15. Si $D^2f(a)$ est dégénérée, alors on ne peut conclure quant à la nature de a .

On peut d'ores et déjà donner un plan (naïf) d'étude des extrema d'une fonction $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 :

1. On cherche les points critiques de f
2. On calcule la matrice hessienne de f en ces points.
 - (a) En les points en lesquels la hessienne est définie positive ou définie négative, on conclut directement à l'existence d'extrema locaux stricts.
 - (b) Si la hessienne est non dégénérée mais ni positive ni négative, on a un point col. Bien que ce ne soit pas un extrema, un point col peut avoir une utilité, par exemple dans l'étude du lagrangien d'un problème d'optimisation sous contrainte.
 - (c) Dans tous les autres cas (hessienne dégénérée), on doit utiliser d'autres outils comme par exemple un développement de Taylor au voisinage des points critiques.

Exemple 16. L'application $(x, y) \in \mathbb{R}^2 \mapsto x^3 + y^3$ admet un point critique en $(0, 0)$ qui n'est ni un extremum local ni un point col, et sa matrice hessienne en $(0, 0)$ est $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ ce qui correspond à une matrice de forme bilinéaire dégénérée. (c.f figure [A.4](#))

2. La notion de point col n'était pas présente dans le plan initial.

Exemple 17. Toute fonction constante a une hessienne nulle (donc dégénérée) en $(0,0)$ et admet un extremum en ce point.

4 Convexité

On s'intéresse dans cette partie à la notion de convexité, qui offre un cadre assez agréable pour l'étude des extrema.

4.1 Généralités sur les fonctions convexes

Définition 18. Soit U un ouvert convexe de E et $f : U \rightarrow \mathbb{R}$. On dit que f est **convexe** si pour tous $x, y \in U$ et $t \in [0, 1]$, $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$. On dit que f est **concave** si $-f$ est convexe.

Propriété 19. Si E est de dimension finie, toute fonction convexe est continue.

Propriété 20. Soit $f : U \rightarrow \mathbb{R}$ deux fois différentiable. Alors f convexe $\iff \forall a \in U, D^2 f(a)$ est positive.

Contre-exemple 21. La valeur absolue $|\cdot|$ est convexe mais non différentiable en 0.

Propriété 22. Soient $f : U \rightarrow \mathbb{R}$ convexe et a minimum local de f . Alors a est un minimum global de f . De même, si $f : U \rightarrow \mathbb{R}$ est concave et a un maximum local de f , alors a est un maximum global de f .

4.2 Espaces de Hilbert

On se donne un espace de Hilbert $(H, \langle \cdot, \cdot \rangle, \|\cdot\|)$.

Théorème 23 (Projection sur un convexe fermé). Soit C un convexe non vide fermé de H . Alors $\forall x \in H, \exists! p_C(x) \in C$; $d(x, C) = \|x - p_C(x)\|$. De plus, $p_C(x)$ est caractérisé par $p_C(x) \in C$ et $\forall y \in C, \operatorname{Re}(\langle x - p_C(x), y - p_C(x) \rangle) \leq 0$. On a également p_C 1-Lipschitzienne sur H .

Remarque 24. Les hypothèses ne sont pas optimales, car H préhilbertien et C convexe fermé non vide complet dans H suffisent.

Remarque 25 (Théorème de projection orthogonale). Si $C = F$ est un sous-espace vectoriel fermé de H , p_F est le projecteur orthogonal sur F parallèlement à F^\perp .

Exemple 26. $\inf_{a,b \in \mathbb{R}} \left(\int_0^1 (x^2 - ax - b)^2 dx \right) = d(X^2, R_1[X])^2 = \|X^2 - p_{R_1[X]}(X^2)\|_2^2 = \frac{1}{180}$

Contre-exemple 27. On considère $C := \{f \in \mathcal{C}^0([0, 1], [0, 1]); f(0) = 0\}$ qui n'est pas un fermé pour la norme $\|\cdot\|_\infty$, et $f \equiv 1$. Alors $d_\infty(f, C) = 1$ mais $\forall g \in C, \|f - g\|_\infty = 1$: pas unicité.

Contre-exemple 28. Soient $H := (\mathcal{C}^0([0, 1], \mathbb{R}), \langle \cdot, \cdot \rangle_{L^2(\mathbb{R})})$ et $F := \left\{ f \in L^2; \int_0^{1/2} f = 0 \right\} \cap H$. Alors $\forall f \in H \setminus F, \forall g \in F, \|f - g\|_{L^2} > d(f, F)$.

Théorème 29 (Théorème de représentation de Riesz). Soit $\phi \in H'$. Alors il existe un unique $f \in H$ tel que $\forall g \in H, \phi(g) = \langle f, g \rangle$, et $\|\phi\|_{H'} = \|f\|_H$.

Théorème 30 (Banach-Alaoglu, DEV 1). Soient H un espace de Hilbert et (f_n) une suite de H .

Si (f_n) est bornée, alors il existe φ une extractrice telle que $(f_{\varphi(n)})$ converge faiblement :

$$\exists f \in H, \forall g \in H \langle f_{\varphi(n)}, g \rangle \xrightarrow{n \rightarrow +\infty} \langle f, g \rangle$$

Preuve. On notera $M > 0$ une borne uniforme pour la suite (f_n) , c'est à dire vérifiant $\forall n \in \mathbb{N}, \|f_n\| \leq M$.

On considèrera dans un premier temps le cas où H est séparable, c'est à dire possédant une partie $D := \{x_n, n \in \mathbb{N}\}$ dénombrable dense.

La première étape consiste à réaliser une extraction diagonale. On remarque tout d'abord que, par l'inégalité de Cauchy-Schwarz, la suite $(\langle f_n, x_0 \rangle)_{n \in \mathbb{N}}$ est une suite de réels bornée par $M \|x_0\|$. Par le théorème de Bolzano-Weierstrass, on peut donc en extraire une sous-suite convergente. Notons φ_0 une extractrice associée.

Supposons maintenant construites les extractrices φ_i rendant les $(\langle f_{\varphi_i(n)}, x_i \rangle)_{n \in \mathbb{N}}$ convergentes pour tous $i \leq k$, pour un certain $k \in \mathbb{N}$. On remarque tout d'abord que pour tout $i \leq k$, la suite $(\langle f_{(\varphi_0 \circ \dots \circ \varphi_i)(n)}, x_k \rangle)_{n \in \mathbb{N}}$ converge.

Ensuite, on remarque là encore que la suite $(\langle f_{(\varphi_0 \circ \dots \circ \varphi_k)(n)}, x_k \rangle)_{n \in \mathbb{N}}$ est une suite de réels bornée, on construit donc une extractrice φ_{k+1} comme précédemment.

En posant $\varphi : n \mapsto (\varphi_0 \circ \dots \circ \varphi_n)(n)$, on obtient une extractrice vérifiant que, pour tout $k \in \mathbb{N}$, $(\langle f_{\varphi(n)}, x_k \rangle)_{n \in \mathbb{N}}$ converge.

La seconde étape consiste à appliquer un critère de Cauchy. Soient $g \in H$ et $n, p \in \mathbb{N}$. On fixe $\varepsilon > 0$.

Par densité de la partie D dans H , il existe $k_0 \in \mathbb{N}$ tel que $\forall k \geq k_0, \|g - x_k\| \leq \frac{\varepsilon}{4M}$.

De plus, $(\langle f_{\varphi(n)}, x_{k_0} \rangle)_{n \in \mathbb{N}}$ converge donc est de Cauchy : il existe $n_0 \in \mathbb{N}$ tel que $\forall n \geq n_0, |\langle f_{\varphi(n+p)} - f_{\varphi(n)}, x_{k_0} \rangle| \leq$

$\frac{\varepsilon}{2}$. Il vient alors

$$\begin{aligned} |\langle f_{\varphi(n+p)} - f_{\varphi(n)}, g \rangle| &= |\langle f_{\varphi(n+p)} - f_{\varphi(n)}, g - x_{k_0} + x_{k_0} \rangle| \\ &\leq |\langle f_{\varphi(n+p)} - f_{\varphi(n)}, g - x_{k_0} \rangle| + |\langle f_{\varphi(n+p)} - f_{\varphi(n)}, x_{k_0} \rangle| \\ &\leq 2M \frac{\varepsilon}{4M} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

On en déduit le résultat en remarquant que la suite est de Cauchy dans H qui est complet, donc convergente. Pour le cas non séparable, il suffit d'appliquer le résultat sur le sous-Hilbert séparable $H_0 = \overline{\{f_n, n \in \mathbb{N}\}}$ et d'écrire $H = H_0 \oplus H_0^\perp$. Il vient alors que $(f_{\varphi(n)})$ converge faiblement vers $f \in H_0$ sur H_0 et pour tout $g = g_0 + g_\perp \in H_0 \oplus H_0^\perp$ on a

$$\langle f_{\varphi(n)}, g \rangle = \langle f_{\varphi(n)}, g_0 \rangle + 0 \xrightarrow[n \rightarrow \infty]{} \langle f, g_0 \rangle + 0 = \langle f, g_0 + g_\perp \rangle.$$

□

Théorème 31 (DEV 1, suite). *Soient H un Hilbert, $C \subset H$ un convexe fermé et $J : C \rightarrow \mathbb{R}$ convexe, coercive et continue. Alors J admet un minimum x^* sur C .*

Preuve. Soit $(x_n)_{n \in \mathbb{N}} \in C^\mathbb{N}$ une suite minimisante pour J (obtenue par application de la caractérisation (séquentielle) de la borne inférieure). On remarque tout d'abord que cette suite est bornée. En effet, si ce n'était pas le cas, on aurait $\|x_n\| \rightarrow +\infty$ et donc $J(x_n) \rightarrow +\infty$ par coercivité de J , ce qui est absurde car on a ici supposé C non vide et donc nécessairement, $\inf_C J < +\infty$. On a donc une suite bornée dans C , et C est un fermé de l'espace de Hilbert H et donc est lui-même un espace de Hilbert. Donc d'après le théorème de Banach-Alaoglu, cette suite converge faiblement dans C . Notons x^* sa limite.

Il s'agit maintenant de montrer que le candidat x^* est bien un minimum de J sur C . Fixons alors un $a > \inf_C J$ et considérons l'ensemble $C_a := \{x \in C, J(x) \leq a\}$. On remarque tout d'abord que C_a est non vide, car x_n est minimisante. De plus, cet ensemble est fermé comme image réciproque du fermé $] -\infty, a]$ par l'application continue J . Il est enfin convexe : si $x, y \in C_a$ et $t \in [0, 1]$ alors par convexité de J il vient

$$J(tx + (1-t)y) \leq tJ(x) + (1-t)J(y) \leq ta + (1-t)a = a.$$

Ainsi, d'après le théorème de projection sur le convexe fermé C_a , il existe un unique $p_a(x^*)$ tel que

$$\forall y \in C_a, \langle x^*, p_a(x^*), y - p_a(x^*) \rangle \leq 0.$$

En particulier pour $y = x^*$ on obtient $\|x^* - p_a(x^*)\|^2 \leq 0$ et donc $x^* \in C_a$. On obtient ainsi le résultat pour tout $a > \inf_C J$, donc il vient

$$\forall a > \inf_C J, J(x^*) \leq a$$

d'où $J(x^*) = \inf_C J = \min_C J$. □

Application 32. Pour $f \in L^2(0,1)$ et $p \geq 1$, l'équation $-u'' + |u|^{p-1}u = f$ d'inconnue $u \in H^2(0,1) \cap H_0^1(0,1)$ admet une unique solution.

5 Optimisation sous contrainte et extrema liés

Définition 33 (Définition par carte locale). Soient $k \geq 1$ et $d \in \mathbb{N}$. Soit $M \subset \mathbb{R}^n$. On dit que M est une *sous-variété* de \mathbb{R}^n si pour tout $x \in M$, il existe U un voisinage de x dans \mathbb{R}^n , V un voisinage de 0 dans \mathbb{R}^n et $\varphi : U \rightarrow V$ un C^k -difféomorphisme, tel que $\varphi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$. On appelle alors le couple (U, φ) une *carte locale* de M en x .

Remarque 34. L'entier d apparaissant dans la définition est unique. On l'appelle *dimension* de M .

Théorème 35. Soit $M \subset \mathbb{R}^n$. Les assertions suivantes sont équivalentes :

- M est une sous-variété de \mathbb{R}^n
- Pour tout $x \in M$, il existe U un voisinage de x dans \mathbb{R}^n et $f : U \rightarrow \mathbb{R}^{n-d} \in C^k$, tel que $d_x f$ est surjective et $M \cap U = f^{-1}(\{0\})$
- Pour tout $x \in M$, il existe U un voisinage de x dans \mathbb{R}^n , V un voisinage de 0 dans \mathbb{R}^n et $j : V \rightarrow \mathbb{R}^n \in C^k$, tel que $d_x j_0$ est injective et $j|_V$ soit un homéomorphisme de V dans $U \cap M$
- Pour tout $x \in M$, il existe U un voisinage de x dans \mathbb{R}^n , V un voisinage de 0 dans \mathbb{R}^n , $u : V \rightarrow \mathbb{R}^{n-p} \in C^k$ et un changement de coordonnées $L \in \text{GL}_n(\mathbb{R})$, tel que $U \cap M = \{L(a, u(a)), a \in V\}$

Définition 36. Soient M une sous variété de \mathbb{R}^n de classe C^k et de dimension d et $x \in M$. L'*espace tangent* à M en x est

$$T_x M = \{\gamma'(0), \gamma : I \rightarrow M \in C^1, 0 \in I, \gamma(0) = x\}$$

Théorème 37 (Théorème des extrema liés, DEV 2). Soient U un ouvert de \mathbb{R}^n , $x \in U$ et f, g_1, \dots, g_p des applications de classe \mathcal{C}^1 de U dans \mathbb{R} . On pose $\Gamma := \{x \in U, g_1(x) = \dots = g_p(x) = 0\}$.

Si $(d_x g_1, \dots, d_x g_p)$ forme une famille libre et si x est un minimum de $f|_\Gamma$, alors il existe $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$, tel que

$$d_x f = \sum_{i=1}^p \lambda_i d_x g_i$$

Preuve. Étape 1 : dimension de l'espace tangent.

D'après le Théorème 35, sous les hypothèses du théorème des extrema liés, Γ est une sous-variété \mathcal{C}^1 de \mathbb{R}^n . On commence par démontrer le lemme suivant qui donne une correspondance entre $T_x \Gamma$ et \mathbb{R}^p :

Lemme 38. Pour tout $x \in \Gamma$, $T_x \Gamma$ est un sous espace vectoriel de \mathbb{R}^n de dimension p .

Preuve du lemme. Soit $x \in \Gamma$. Par définition d'une sous-variété par carte locale, il existe U un voisinage de x dans \mathbb{R}^n , V un voisinage de 0 dans \mathbb{R}^p et $\varphi : U \rightarrow V$ un \mathcal{C}^k -difféomorphisme, tel que $\varphi(U \cap \Gamma) = V \cap (\mathbb{R}^p \times \{0\})$.

Soit $v \in T_x \Gamma$. Par définition, il existe $\gamma : I \rightarrow \Gamma \in \mathcal{C}^1$, avec $0 \in \overset{\circ}{I}$, tel que $\gamma(0) = x$ et $\gamma'(0) = v$. On a donc $\varphi \circ \gamma \in \mathcal{C}^1$ et

$$(\varphi \circ \gamma)'(0) = d_{\gamma(0)} \varphi \cdot \gamma'(0) = d_x \varphi(v) \in \mathbb{R}^p \times \{0\}$$

Ainsi, $T_x \Gamma \in (d_x \varphi)^{-1}(\mathbb{R}^p \times \{0\})$.

Réciproquement, supposons $v \in (d_x \varphi)^{-1}(\mathbb{R}^p \times \{0\})$. On considère $\gamma : I \rightarrow \mathbb{R}^p \times \{0\}$ un chemin de classe \mathcal{C}^1 passant par $0 \in \overset{\circ}{I}$ tel que $\gamma(0) = 0$ et $\gamma'(0) = d_x \varphi.v$. Ceci est légitime car $V \cap \mathbb{R}^p \times \{0\}$ est un ouvert de l'espace vectoriel $\mathbb{R}^p \times \{0\}$. En particulier il existe une boule ouverte $\mathcal{B}(d_x \varphi.v, r) \subset \mathbb{R}^p \times \{0\}$ autour de $d_x \varphi.v$ qui soit incluse dans $\varphi(U \cap \Gamma)$. On considère alors $\gamma : t \in I \mapsto x + t d_x \varphi.v \in \mathbb{R}^p \times \{0\}$ avec $I :=]-r, r[$. Ainsi, $\varphi^{-1} \circ \gamma$ est un chemin de classe \mathcal{C}^1 de Γ et

$$(\varphi^{-1} \circ \gamma)'(0) = d_{\gamma(0)} \varphi^{-1} \cdot (\gamma'(0)) = (d_x \varphi)^{-1} \cdot (d_x \varphi.v) = v.$$

Ainsi, $T_x \Gamma = (d_x \varphi)^{-1}(\mathbb{R}^p \times \{0\}) \simeq \mathbb{R}^p$ est bien de dimension p . □

Étape 2 : montrer $T_x \Gamma = \bigcap_{i=1}^p \text{Ker}(d_x g_i)$.

Comme la famille $(d_x g_1, \dots, d_x g_p)$ est libre, il suffit de montrer que $T_x \Gamma \subset \bigcap_{i=1}^p \text{Ker}(d_x g_i)$ et on peut conclure par égalité des dimensions d'après le lemme précédent.

Soit $v \in T_x \Gamma$. Il existe $\gamma : I \rightarrow \Gamma$ un chemin \mathcal{C}^1 , avec $0 \in \overset{\circ}{I}$, tel que $\gamma(0) = x$ et $\gamma'(0) = v$. Pour tout $i \in \llbracket 1, p \rrbracket$, on a $g_i \circ \gamma = 0$, car $\gamma(I) \subset \Gamma$, et donc

$$d_x g_i \cdot v = d_{\gamma(0)} g_i \cdot \gamma'(0) = (g_i \circ \gamma)'(0) = 0$$

d'où le résultat annoncé.

Étape 3 : montrer que la famille $(d_x f, d_x g_1, \dots, d_x g_p)$ est liée.³

Pour montrer cela, on commence par montrer que $\bigcap_{i=1}^p \text{Ker}(d_x g_i) \subset \text{Ker}(d_x f)$.

Soit $v \in T_x \Gamma = \bigcap_{i=1}^p \text{Ker}(d_x g_i)$. Il existe $\gamma : I \rightarrow \Gamma \in \mathcal{C}^1$, avec $0 \in \overset{\circ}{I}$, tel que $\gamma(0) = x$ et $\gamma'(0) = v$. Comme x est un minimum de f sur Γ , que $\gamma(I) \subset \Gamma$ et que $\gamma(0) = x$, on en déduit que 0 est un minimum de $f \circ \gamma$ sur I . Ainsi,

$$0 = (f \circ \gamma)'(0) = df_{\gamma(0)} \cdot \gamma'(0) = d_x f \cdot v$$

d'où le résultat. Montrons maintenant que la famille $(d_x f, d_x g_1, \dots, d_x g_p)$ est liée.

D'après le Théorème de représentation de Riesz, il existe u, v_1, \dots, v_p des vecteurs de \mathbb{R}^n , tel que, pour tout $w \in \mathbb{R}^n$, $d_x f \cdot w = \langle u, w \rangle$ et pour tout $i \in \llbracket 1, p \rrbracket$, $d_x g_i \cdot w = \langle v_i, w \rangle$. On pose alors

$$M = \begin{pmatrix} u \\ v_1 \\ \vdots \\ v_p \end{pmatrix}$$

On remarque alors que M est la matrice de l'application $F : \begin{pmatrix} \mathbb{R}^n & \rightarrow & \mathbb{R}^{p+1} \\ w & \mapsto & (d_x f \cdot w, d_x g_1 \cdot w, \dots, d_x g_p \cdot w) \end{pmatrix}$

D'après le théorème du rang, on a $rg(M) + \dim \text{Ker}(M) = n$.

Comme $\bigcap_{i=1}^p \text{Ker}(d_x g_i) \subset \text{Ker}(d_x f)$, on a $\bigcap_{i=1}^p \text{Ker}(d_x g_i) \subset \text{Ker}(M)$. De plus, comme $(d_x g_1, \dots, d_x g_p)$ est une

famille libre, on a $\dim(\text{Ker}(M)) \geq \dim \bigcap_{i=1}^p \text{Ker}(d_x g_i) = n - p$.

Ainsi, $rg(M) \leq p$ et les $p + 1$ vecteurs u, v_1, \dots, v_p sont liés. Il existe donc $(\mu_0, \dots, \mu_p) \in \mathbb{R}^{p+1}$ non tous

3. La leçon se concentrant sur le calcul différentiel plutôt que l'algèbre linéaire, il peut être judicieux d'abrégé cette étape à l'oral.

nuls, tel que

$$\sum_{i=1}^p \mu_i d_x g_i + \mu_0 d_x f = 0$$

En particulier, comme $(d_x g_1, \dots, d_x g_p)$ est une famille libre, on a $\mu_0 \neq 0$ et, en posant $\lambda_i = -\frac{\mu_i}{\mu_0}$ pour tout $i \in \llbracket 1, p \rrbracket$, on a

$$d_x f = \sum_{i=1}^p \lambda_i d_x g_i$$

□

Application 39. Si $u \in \mathcal{L}(\mathbb{R}^n)$ autoadjoint : l'étude des extrema de $f : x \in \mathbb{R}^n \mapsto \langle u(x), x \rangle$ sur la sphère $\mathcal{S}(0, 1) \subset \mathbb{R}^n$ permet de montrer le théorème spectral.⁴

6 Méthodes numériques

Théorème 40 (Dichotomie). Soient $I = [a, b]$ un intervalle de \mathbb{R} et $f \in \mathcal{C}^1(I, \mathbb{R})$, tel que $f'(a)f'(b) < 0$. Soit $x_0 \in [a, b]$, tel que $f'(x_0) = 0$. On considère $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$ et $(c_n)_{n \in \mathbb{N}}$ les suites définies par :

- $a_0 = a, b_0 = b, c_0 = \frac{b+a}{2}$.
- Pour $n \in \mathbb{N}$:

★ Si $f'(c_n) = 0$ on arrête l'algorithme.

★ Si $f'(c_n)f'(b_n) < 0$, alors on pose $a_{n+1} = c_n, b_{n+1} = b_n$ et $c_{n+1} = \frac{b_{n+1} + a_{n+1}}{2}$.

★ Si $f'(a_n)f'(c_n) < 0$, alors on pose $a_{n+1} = a_n, b_{n+1} = c_n$ et $c_{n+1} = \frac{b_{n+1} + a_{n+1}}{2}$.

On a alors pour tout $n \in \mathbb{N}$, $|c_n - x_0| \leq \frac{b+a}{2^{n+1}}$.

Remarque 41. En particulier $(c_n)_{n \in \mathbb{N}}$ converge vers le point critique x_0 . Cela permet de trouver les points critiques de f et donc ses extrema potentiels.

Dans la suite de cette partie, A désigne une matrice inversible et on désigne par λ_i ses valeurs propres, rangées par module croissant : $0 < |\lambda_1| \leq \dots \leq |\lambda_n|$.

4. Initialement partie intégrante du développement, il paraît après réflexion plus judicieux de l'enlever afin notamment de passer plus de temps sur les extrema liés pour développer un maximum le point de vue des sous-variétés.

Définition 42. Soit $A \in \text{GL}_n(\mathbb{R})$ et $M, N \in \text{GL}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R})$ telles que $A = M - N$. Une **méthode itérative** associée à A est la donnée d'une suite récurrente de la forme

$$\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N}, Mx_{k+1} = Nx_k + b \end{cases}$$

Soit $e_k := x_k - x$ le vecteur d'erreur, où $x := A^{-1}b$. On dit qu'une telle méthode **converge** si $e_k \xrightarrow[k \rightarrow +\infty]{} 0$.

Propriété 43. Une telle méthode itérative est convergente si et seulement si $\rho(M^{-1}N) < 1$, où ρ désigne le rayon spectral.

Preuve. On a, en utilisant $x_{k+1} = M^{-1}Nx_k + M^{-1}b$ ainsi que l'égalité $b = Ax = (M - N)x$,

$$e_{k+1} = M^{-1}Nx_k + M^{-1}b - x = M^{-1}N(x_k - x) + M^{-1}Nx + M^{-1}Ax - x = M^{-1}Ne_k$$

d'où $e_k = (M^{-1}N)^k e_0$. Soit une norme subordonnée $\|\cdot\|$ telle que $\forall A \in \mathcal{M}_n(\mathbb{R}), \|A\| \leq \rho(A) + \varepsilon$ avec $\varepsilon = \frac{1 - \rho(M^{-1}N)}{2}$. De cette manière on a en particulier $\|M^{-1}N\| \leq \frac{1 + \rho(M^{-1}N)}{2} < 1$. On a alors

$$\|e_k\| \leq \left\| (M^{-1}N)^k \right\| \|e_0\| \leq \|M^{-1}N\|^k \|e_0\| \leq \left(\frac{1 + \rho(M^{-1}N)}{2} \right)^k \|e_0\| \xrightarrow[k \rightarrow +\infty]{} 0$$

d'où le résultat annoncé. □

Théorème 44 (Méthode du gradient à pas fixe). Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$ une matrice symétrique définie positive et b un vecteur de \mathbb{R}^n . On considère $J : x \in \mathbb{R}^n \mapsto \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle$. Si $\alpha \in \left] 0, \frac{2}{\lambda_n} \right[$, alors pour tout choix de $x_0 \in \mathbb{R}^n$, la suite récurrente définie par $\forall n \in \mathbb{N}, x_{k+1} = x_k - \alpha \nabla J(x_k)$ converge. C'est la méthode itérative avec $M = \frac{1}{\alpha} I_n$ et $N = M - A$.

Preuve. On pose $x := A^{-1}b$. Montrons que $e_k = x_k - x \xrightarrow[k \rightarrow +\infty]{} 0$. On a, pour tout $k \in \mathbb{N}$,

$$e_{k+1} = x_{k+1} - x = (I_n - \alpha A)x_k + \alpha b - x = (I_n - \alpha A)(x_k - x) - \alpha Ax + \alpha b = (I_n - \alpha A)e_k$$

d'où, par récurrence immédiate, $e_k = (I_n - \alpha A)^k e_0$. Or ici $\rho(I_n - \alpha A) < 1$. En effet, si x est un vecteur propre associé à une valeur propre $\lambda \in \mathbb{R}$ de $I_n - \alpha A$ alors

$$(I_n - \alpha A)x = \lambda x \text{ et donc } Ax = \frac{1 - \lambda}{\alpha} x$$

car $\alpha > 0$. Donc il existe $j \in \llbracket 1, n \rrbracket$ tel que $\frac{1-\lambda}{\alpha} = \lambda_j$. En particulier $\lambda = 1 - \lambda_j \alpha$ et donc puisque $0 < \lambda_j \leq \lambda_n$ et que $\alpha < \frac{2}{\lambda_n}$ on en déduit $-1 < \lambda_j \alpha - 1 < 1$ c'est à dire $\lambda < 1$. D'où $\rho(I_n - \alpha A) < 1$.

Enfin, ici on a $M^{-1}N = \left(\frac{1}{\alpha}I_n\right)^{-1} \left(\frac{1}{\alpha}I_n - A\right) = I_n - \alpha A$ et donc par la propriété 43 le schéma est bien convergent. \square

Remarque 45. *L'intérêt de considérer la suite définie par $x_{k+1} = x_k - \alpha \nabla J(x_k)$ est que cette suite converge vers l'unique $x \in \mathbb{R}^n$ tel que $Ax = b$. En effet ici on a par symétrie de A l'égalité*

$$J(y+h) = \frac{1}{2} (\langle Ay, h \rangle + \langle Ay, h \rangle + \langle Ah, y \rangle + \langle Ah, h \rangle) - \langle b, y \rangle - \langle b, h \rangle = J(y) + \langle (Ay - b), h \rangle + o(h)$$

et donc on a $\nabla_y J = Ay - b$, et donc l'unique point critique de J est x . En fait, on a $J(x+h) = J(x) + 0 + \langle Ah, h \rangle \geq J(x)$ pour tout h car $A \in \mathcal{S}_n^{++}(\mathbb{R})$. On en déduit donc que x est un minimum local pour J . On a par ailleurs J convexe d'après la propriété 20 car $D_y^2 J = A$. Donc d'après x est le minimum global de J d'après la propriété 22.

Remarque 46. *La vitesse de convergence est optimale pour $\alpha = \frac{2}{\lambda_n + \lambda_1}$.*

Un des avantages de la méthode du gradient à pas fixe est que le pas est justement fixe donc il n'y a besoin de le calculer qu'une seule fois. Cependant, il existe une méthode dite du gradient conjugué qui permet, au prix d'une modification du pas, d'aboutir à la solution exacte du système en un nombre fini d'étapes.

Les résultats qui suivent, traitant du gradient conjugué, font partie des ajouts qui n'étaient pas présents sur le plan initial. Seul l'énoncé du gradient à pas fixe était présent, ce qui rendait pauvre cette partie sur l'analyse numérique. Il semble pourtant naturel de présenter des méthodes effectives de recherche d'extrema, et on aurait pu également discuter d'autres méthodes comme par exemple la méthode du gradient à pas optimal qui est une autre amélioration de la méthode de point fixe. Il aurait également été possible de présenter des méthodes numériques de recherche de valeurs propres qui apparaissent, par exemple dans l'énoncé du théorème 44, comme des valeurs remarquables dans les problèmes d'optimisation.

Définition 47. *Soit $r_0 \in \mathbb{R}^n$. On appelle le m -ième espace de Krylov associé à A et r_0 l'espace*

$$\mathcal{K}_m := \text{vect}(r_0, \dots, A^m r_0) \subset \mathbb{R}^n$$

Propriété 48. La suite $(\mathcal{K}_m)_{m \in \mathbb{N}}$ est stationnaire. De plus, le plus petit rang k_0 à partir duquel elle stationne vérifie $1 \leq k_0 \leq n - 1$.

Théorème 49 (Méthode du gradient conjugué). Soient $A \in \mathcal{S}_n^{++}(\mathbb{R})$ une matrice symétrique définie positive et b un vecteur de \mathbb{R}^n . Soit $x_0 \in \mathbb{R}^n$. On pose $r_0 := b - Ax_0$ et on considère la suite $(\mathcal{K}_m)_{m \in \mathbb{N}}$ des espaces de Krylov associés. Alors pour tout $k \geq 1$ il existe un unique vecteur x_k vérifiant

$$x_k \in x_0 + \mathcal{K}_{k-1} \text{ et } r_k = b - Ax_k \in \mathcal{K}_{k-1}^\perp.$$

Cette méthode, appelée méthode du gradient conjugué, stationne à l'unique solution x du système $Ax = b$ en au plus n itérations : il existe un rang $n_0 \leq n$ à partir duquel $x_n = x$.

Preuve. Montrons tout d'abord l'existence et l'unicité de x_k pour tout $k \geq 1$. On considère le produit scalaire $\langle A \cdot, \cdot \rangle$ que l'on note $\langle \cdot, \cdot \rangle_A$. C'est bien un produit scalaire puisque A est symétrique définie positive. Soit $y_k \in \mathcal{K}_{k-1}$ et $x := x_0 + y_k$. La condition d'orthogonalité $r_k \in \mathcal{K}_{k-1}^\perp$ est vérifiée si et seulement si pour tout $y \in \mathcal{K}_{k-1}$,

$$\langle r_k, y \rangle = 0 \iff \langle b - A(x_0 + y_k), y \rangle = 0 \iff \langle r_0 - Ay_k, y \rangle = 0 \iff \langle A^{-1}r_0 - y_k, y \rangle_A = 0.$$

D'après le théorème de la projection orthogonale du vecteur $A^{-1}r_0$ sur le sous-espace vectoriel \mathcal{K}_{k-1} dans l'espace euclidien $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_A)$, il existe un unique vecteur y_k vérifiant ceci, d'où l'existence et l'unicité de $x_k = x_0 + y_k$. Soit alors k_0 comme en définition 47. On a alors $A\mathcal{K}_{k_0} \subset \mathcal{K}_{k_0+1} = \mathcal{K}_{k_0}$ donc $r_{k_0+1} \in \mathcal{K}_{k_0} \cap \mathcal{K}_{k_0}^\perp = \{0\}$. On en déduit que $r_{k_0+1} = b_0 - Ax_{k_0+1} = 0$ et donc que x_{k_0+1} est l'unique solution de $Ax = b$. Puisque $k_0 \leq n$ cela conclut. \square

Remarque 50. Pour ce qui est du calcul des r_k et des x_k en pratique, on peut considérer l'algorithme suivant : on choisit $x_0 \in \mathbb{R}^n$, on pose $p_0 = r_0 = b - Ax_0$ et, pour $k \geq 0$:

$$\begin{cases} x_{k+1} = x_k + \alpha_k p_k \\ r_{k+1} = r_k - \alpha_k A p_k \\ p_{k+1} = r_{k+1} + \beta_k p_k \end{cases}$$

où les α_k et β_k sont définis pour tout $k \geq 0$ par $\alpha_k := \frac{\|r_k\|^2}{\langle p_k, p_k \rangle_A}$ et $\beta_k := \frac{\|r_{k+1}\|^2}{\|r_k\|^2}$.

On a donc établi une méthode itérative qui converge vers une solution exacte en un nombre fini d'itérations. En pratique cependant il faut prendre en compte les erreurs d'arrondis provenant de la machine, et le vecteur x_{k_0+1} obtenu est rarement exactement égal à x . Il est par exemple possible d'implémenter l'algorithme précédent avec un critère d'arrêt du type $\frac{\|r_k\|}{\|r_0\|} \leq \varepsilon$ avec une tolérance ε à choisir. On pourra se référer à [ALL02] pour plus de détails.

7 Etude du module des fonctions holomorphes

Dans cette partie, Ω est un ouvert de \mathbb{C} et $f : \Omega \rightarrow \mathbb{C}$ désigne une fonction à variable complexe.

Propriété 51 (Formule de Cauchy sur les disques). *Soit $f : \Omega \rightarrow \mathbb{C}$ holomorphe. Alors*

$$\forall b \in \Omega, f(b) = \frac{1}{2i\pi} \int_{\partial D(b,r)} \frac{f(z)}{z-b} dz \text{ avec } \overline{D(b,r)} \subset \Omega$$

Propriété 52 (Principe du maximum). *Soit $f : \Omega \rightarrow \mathbb{C}$ holomorphe. Alors*

$$\forall b \in \Omega, |f(b)| \leq \max_{z \in \partial D(b,r)} |f(z)| \text{ avec } \overline{D(b,r)} \subset \Omega$$

Corollaire 53. *Les fonctions constantes sont les seules fonctions holomorphes admettant un maximum local non global sur \mathbb{C} .*

Corollaire 54. *Soit $f : \Omega \rightarrow \mathbb{C}$ holomorphe.*

Si $|f|$ admet un minimum local sur Ω , alors ce minimum est nul.

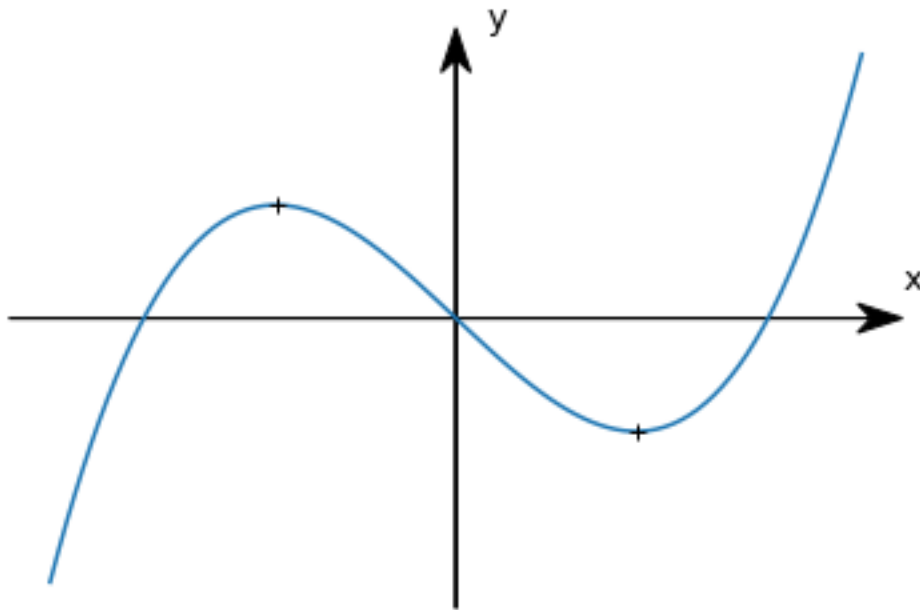
Application 55 (Théorème de d'Alembert Gauss). *\mathbb{C} est algébriquement clos.*

Preuve. Le raisonnement peut être mené de manière similaire, quoique plus simple, que le théorème d'optimisation dans un Hilbert. Soit P un polynôme non constant à coefficients complexes. On remarque que $|P|$ est coercif, au sens où $|P(z)| \xrightarrow{|z| \rightarrow +\infty} +\infty$. En particulier, si on prend une suite minimisante (z_n) pour $|P|$ on obtient une suite bornée. En effet si cette suite minimisante n'était pas bornée, on pourrait en extraire une sous-suite $z_{\varphi(n)}$ qui vérifie $|z_{\varphi(n)}| \xrightarrow{n \rightarrow +\infty} +\infty$. En particulier $|P(z_{\varphi(n)})| \xrightarrow{n \rightarrow +\infty} +\infty$ et donc par définition d'une suite minimisante on a alors $|P(z_{\varphi(n)})| \xrightarrow{n \rightarrow +\infty} \inf_{\mathbb{C}} |P|$, d'où $\inf_{\mathbb{C}} |P| = +\infty$ par unicité de la limite. Ceci est contradictoire car P est une fonction bien définie sur \mathbb{C} tout entier donc nécessairement $\inf_{\mathbb{C}} |P| < +\infty$.

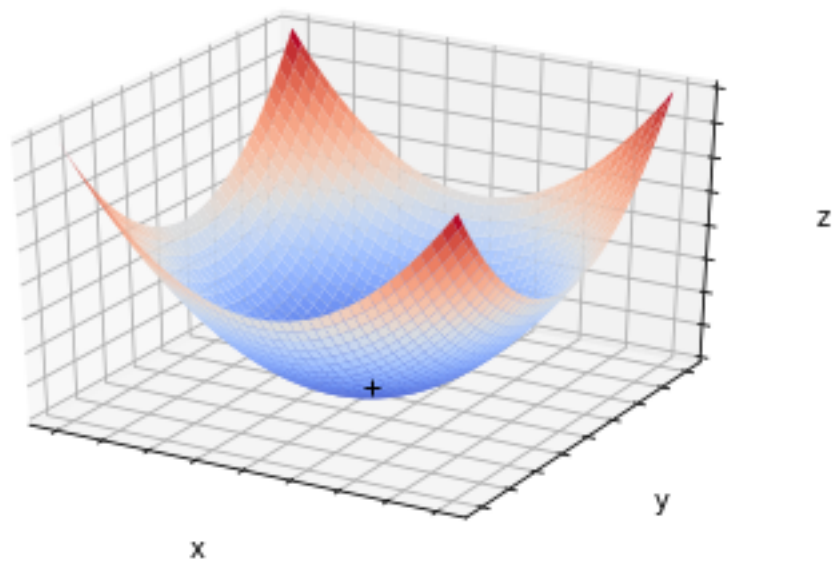
On en extrait par Bolzano-Weierstrass une suite convergente : on obtient un $z \in \mathbb{C}$ tel que $|P(z)| = \min_{\mathbb{C}} |P|$. En vertu de l'holomorphie de P , on obtient par le corollaire précédent $z \in \mathbb{C}$ tel que $|P(z)| = 0$. Ainsi, tout polynôme complexe non constant a une racine, ce qui se traduit par le fait que \mathbb{C} est algébriquement clos. \square

A Figures

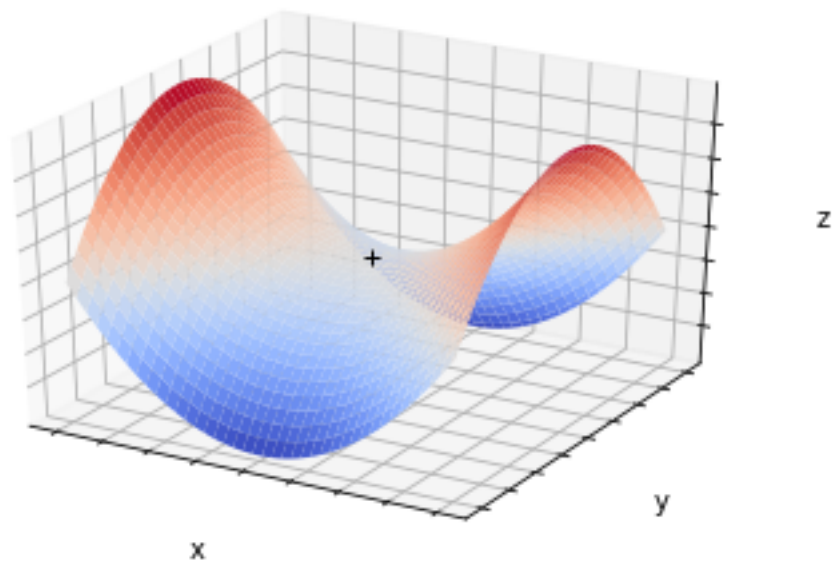
A.1 Graphe de $f : x \mapsto x^3 - x$



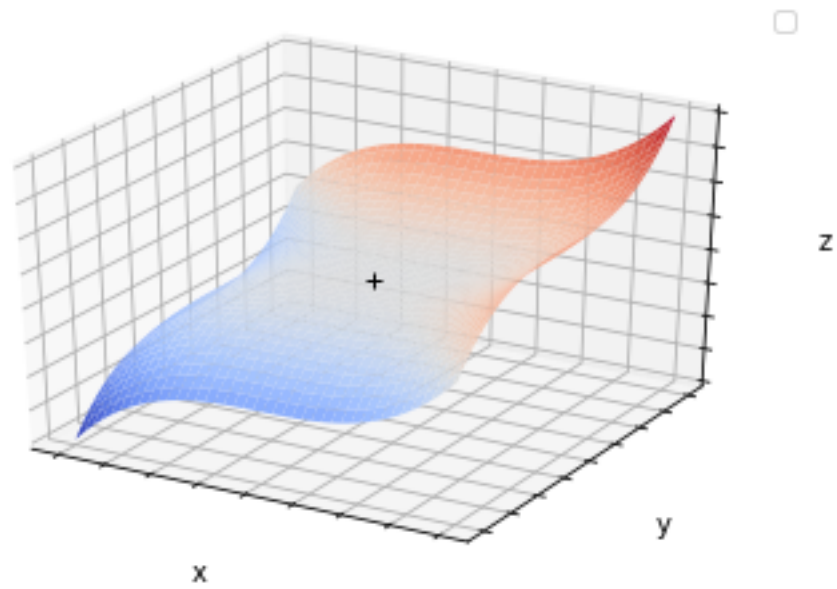
A.2 Graphe de $f : x \mapsto x^2 + y^2$



A.3 Graphe de $f : x \mapsto x^2 - y^2$



A.4 Graphe de $f : x \mapsto x^3 + y^3$



Références

- [ALL02] Grégoire Allaire, Analyse numérique et optimisation
- [AVE93] Avez André, Calcul différentiel
- [ROU99] François Rouvière, Petit guide de calcul différentiel