

Complexité des méthodes homotopiques pour la résolution des systèmes polynomiaux Journées Nationales de Calcul Formel

Jean-Pierre Dedieu*

Année du Tigre

1 Fondements géométriques

1.1 Systèmes et solutions

Un **système polynomial** réel ou complexe est la donnée d'une application

$$f = (f_1, \dots, f_n) : \mathbb{K}^n \rightarrow \mathbb{K}^n$$

où $f_i = f_i(z_1, \dots, z_n)$ est un polynôme à coefficients dans \mathbb{K} . Ici $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . L'espace des systèmes avec $\deg f_i \leq d_i$ est noté \mathcal{P}_d .

Un **zéro** de ce système est un vecteur en lequel ce système est nul. Ce zéro peut être réel ou complexe.

Il est d'usage de considérer aussi le **système homogène associé** à f . Il est défini par

$$F = (F_1, \dots, F_n) : \mathbb{K}^{n+1} \rightarrow \mathbb{K}^n$$

avec

$$F_i(z_0, z_1, \dots, z_n) = z_0^{d_i} f \left(\frac{z_1}{z_0}, \dots, \frac{z_n}{z_0} \right)$$

lorsque $\deg f_i \leq d_i$. L'espace des systèmes homogènes F avec $\deg F_i = d_i$ est noté \mathcal{H}_d , $d = (d_1, \dots, d_n)$.

Un **zéro** de F est un vecteur non nul $z = (z_0, z_1, \dots, z_n) \in \mathbb{K}^{n+1}$ en lequel ce système est nul. Comme F est homogène, il annule toute la droite $\mathbb{K}(z_0, z_1, \dots, z_n)$ et définit donc un point de l'espace projectif $\mathbb{P}_n(\mathbb{K}) = \mathbb{P}(\mathbb{K}^{n+1})$.

*Institut de Mathématiques de Toulouse, Université Paul Sabatier, 31069 Toulouse cedex 9, France (jean-pierre.dedieu@math.univ-toulouse.fr).

Le rapport entre zéros de f et zéros de F est le suivant :

$$F(z_0, z_1, \dots, z_n) = 0 \text{ et } z_0 \neq 0 \text{ si et seulement si } f\left(\frac{z_1}{z_0}, \dots, \frac{z_n}{z_0}\right) = 0$$

en particulier

$$F(1, z_1, \dots, z_n) = 0 \text{ si et seulement si } f(z_1, \dots, z_n) = 0.$$

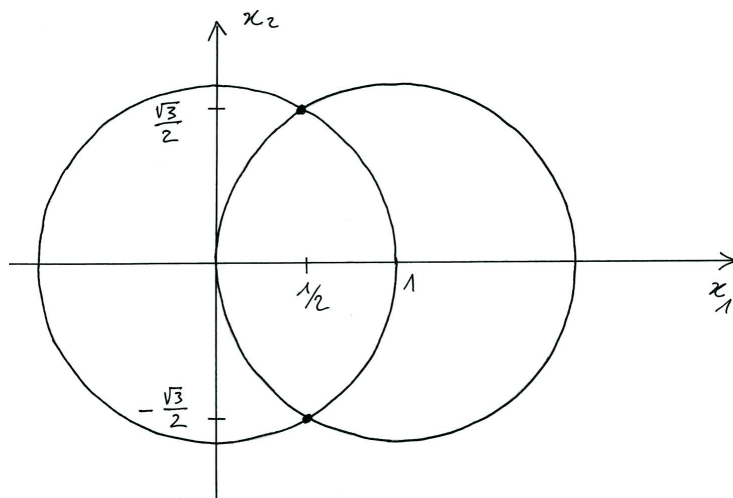
On parle alors de **zéro affine**. Lorsque $F(0, z_1, \dots, z_n) = 0$ on dit que c'est un **zéro à l'infini**.

Par exemple le système

$$f(z_1, z_2) = (z_1^2 + z_2^2 - 1, (z_1 - 1)^2 + z_2^2 - 1)$$

a pour solutions

$$(z_1, z_2) = \left(\frac{1}{2}, \pm \frac{\sqrt{3}}{2}\right).$$



Le système homogène associé est

$$F(z_0, z_1, z_2) = (z_1^2 + z_2^2 - z_0^2, (z_1 - z_0)^2 + z_2^2 - z_0^2).$$

Il a pour solutions

$$\left(1, \frac{1}{2}, \pm \frac{\sqrt{3}}{2}\right) \text{ et } (0, 1, \pm i).$$

Deux solutions à l'infini viennent s'ajouter aux précédentes.

L'espace des solutions sera, suivant les cas considérés :

$$\mathbb{R}^n, \mathbb{C}^n, \mathbb{P}_n(\mathbb{R}), \mathbb{P}_n(\mathbb{C}).$$

L'exemple précédent montre l'intérêt de considérer l'espace projectif $\mathbb{P}_n(\mathbb{C})$ comme espace des solutions ce que nous ferons désormais.

Comme tout multiple d'un système a les mêmes zéros que le système lui-même nous considérerons au lieu de \mathcal{P}_d et \mathcal{H}_d les espaces projectifs associés : $\mathbb{P}(\mathcal{P}_d)$ et $\mathbb{P}(\mathcal{H}_d)$. Encore une fois, au vu de l'exemple précédent, nous préférons le contexte homogène.

Nous introduisons enfin la **variété problèmes-solutions**

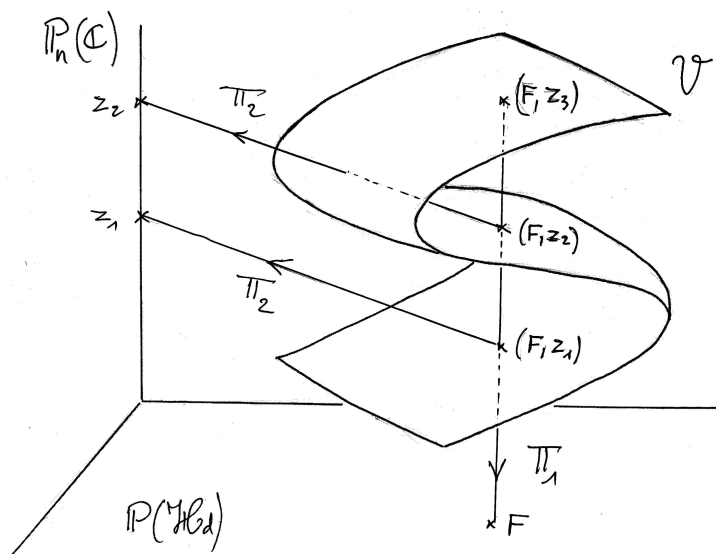
$$\mathcal{V} = \{(F, z) \in \mathbb{P}(\mathcal{H}_d) \times \mathbb{P}_n(\mathbb{C}) : F(z) = 0\}$$

et nous lui associons les restrictions Π_1 et Π_2 des projections sur les espaces de coordonnées :

$$Pr_1 : (F, z) \in \mathbb{P}(\mathcal{H}_d) \times \mathbb{P}_n(\mathbb{C}) \rightarrow F \in \mathbb{P}(\mathcal{H}_d)$$

et

$$Pr_2 : (F, z) \in \mathbb{P}(\mathcal{H}_d) \times \mathbb{P}_n(\mathbb{C}) \rightarrow z \in \mathbb{P}_n(\mathbb{C}).$$



1.2 La structure de variété riemannienne de \mathcal{V}

$\mathbb{P}_n(\mathbb{C})$ est muni d'une structure de **variété différentiable** et l'espace tangent à $z \in \mathbb{P}_n(\mathbb{C})$ est identifié à

$$T_z \mathbb{P}_n(\mathbb{C}) = z^\perp = \{u \in \mathbb{C}^{n+1} : \langle u, z \rangle = 0\}.$$

On munit enfin $\mathbb{P}_n(\mathbb{C})$ de la **structure riemannienne** définie par le produit scalaire

$$\langle \cdot, \cdot \rangle_z = \frac{\langle \cdot, \cdot \rangle}{\langle z, z \rangle}$$

sur l'espace tangent sur $T_z \mathbb{P}_n(\mathbb{C})$. La **longueur d'un arc de courbe** (absolument continue) $\gamma : [0, 1] \rightarrow \mathbb{P}_n(\mathbb{C})$ est donnée par l'intégrale :

$$L(\gamma) = \int_0^1 \left\| \frac{d}{dt} \gamma(t) \right\|_{\gamma(t)} dt.$$

À ce dernier concept, nous associons celui de **distance riemannienne** d_R qui est l'infimum des longueurs des courbes reliant deux points. Dans le cas présent,

$$d_R(u, v) = \arccos \frac{|\langle u, v \rangle|}{\|u\| \|v\|}$$

est l'angle fait par les droites issues de u et v .

Une autre distance, souventfois utilisée, est la **distance projective** définie par

$$d_P(u, v) = \sin d_R(u, v).$$

\mathcal{H}_d est muni du produit scalaire hermitien suivant

$$\langle F, G \rangle_d = \sum_{i=1}^n \langle F_i, G_i \rangle_{d_i}.$$

Le produit hermitien $\langle F_i, G_i \rangle_{d_i}$ est défini de la façon suivante : posons

$$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^{n+1}, \quad |\alpha| = \alpha_0 + \alpha_1 + \dots + \alpha_n, \quad \binom{d_i}{\alpha} = \frac{d_i!}{\alpha_0! \alpha_1! \dots \alpha_n!}$$

ainsi que

$$F_i = \sum_{|\alpha|=d_i} a_\alpha z^\alpha, \quad G_i = \sum_{|\alpha|=d_i} b_\alpha z^\alpha, \quad \text{avec } z^\alpha = z_0^{\alpha_0} z_1^{\alpha_1} \dots z_n^{\alpha_n}.$$

Alors

$$\langle F_i, G_i \rangle_{d_i} = \sum_{|\alpha|=d_i} \binom{d_i}{\alpha}^{-1} a_\alpha \bar{b}_\alpha = \frac{(d_i + n)!}{d_i! 2\pi^{n+1}} \int_{\mathbb{S}^{2n+1}} F_i(z) \overline{G_i(z)} dz.$$

Cette dernière identité montre que

$$\langle F_i \circ U, G_i \circ U \rangle_{d_i} = \langle F_i, G_i \rangle_{d_i}$$

pour toute transformation unitaire $U \in \mathbb{U}_{n+1}$. La structure riemannienne de $\mathbb{P}(\mathcal{H}_d)$ est définie de façon identique à celle de $\mathbb{P}_n(\mathbb{C})$.

\mathcal{V} est une **sous-variété différentiable** du produit $\mathbb{P}(\mathcal{H}_d) \times \mathbb{P}_n(\mathbb{C})$ et sa dimension complexe est

$$\dim_{\mathbb{C}} \mathcal{V} = \dim_{\mathbb{C}} \mathbb{P}(\mathcal{H}_d).$$

L'espace tangent en un couple $(F, z) \in \mathcal{V}$ est décrit par les trois équations suivantes

$$T_{(F,z)}\mathcal{V} = \left\{ (\dot{F}, \dot{z}) \in \mathcal{H}_d \times \mathbb{C}^{n+1} : \langle \dot{F}, F \rangle = 0, \langle \dot{z}, z \rangle = 0, DF(z)\dot{z} + \dot{F}(z) = 0 \right\}.$$

Cet espace tangent est muni du produit scalaire

$$\langle \cdot, \cdot \rangle_{(F,z)} = \langle \cdot, \cdot \rangle_F + \langle \cdot, \cdot \rangle_z$$

de sorte que la longueur d'une courbe $\Gamma(t) = (F_t, z_t) \in \mathcal{V}$, $0 \leq t \leq 1$, est

$$L(\Gamma) = \int_0^1 \sqrt{\left\| \frac{d}{dt} F_t \right\|_{F_t}^2 + \left\| \frac{d}{dt} z_t \right\|_{z_t}^2} dt.$$

1.3 Variété critique et variété discriminante

Les paires $(F, z) \in \mathcal{V}$ sont de deux types suivant que z est une racine simple de F ou bien multiple. Mais qu'est-ce qu'une racine simple ?

Théorème 1.1 *Soit $(F, z) \in \mathcal{V}$ c'est-à-dire que $F(z) = 0$. Il y a équivalence entre :*

1. $DF(z) : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^n$ est de rang n ,
2. $DF(z)|_{z^\perp} : z^\perp \rightarrow \mathbb{C}^n$ est inversible,
3. $D\Pi_1(F, z) : T_{(F,z)}\mathcal{V} \rightarrow T_F\mathbb{P}(\mathcal{H}_d)$ est un isomorphisme.

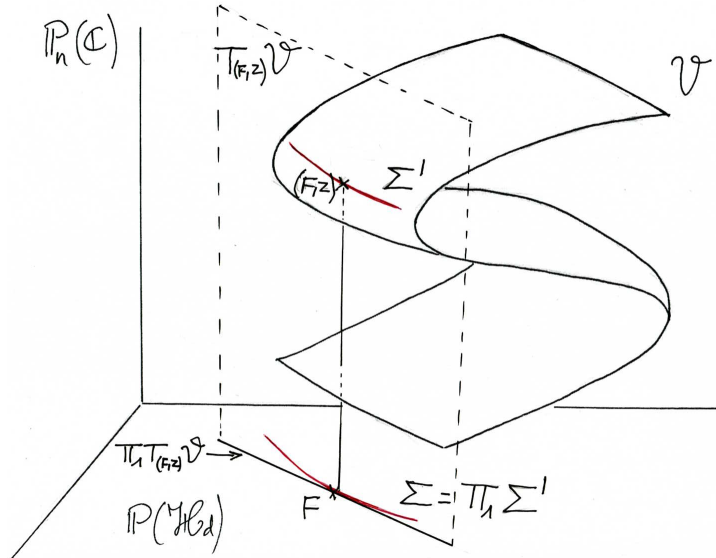
Dans ce cas,

$$\Pi_1 : \mathcal{V} \rightarrow \mathbb{P}(\mathcal{H}_d)$$

est un difféomorphisme d'un voisinage de $(F, z) \in \mathcal{V}$ sur un voisinage de $F \in \mathbb{P}(\mathcal{H}_d)$ et l'on dit que z est une **racine simple** de F ou bien que (F, z) est un **problème bien posé**. Dans le cas contraire on parle de **racine multiple** ou bien de **problème mal posé**.

Ceci conduit à deux nouveaux concepts : la **variété critique** $\Sigma' \subset \mathcal{V}$ est l'ensemble des problèmes mal posés $(F, z) \in \mathcal{V}$. C'est une sous-variété algébrique de \mathcal{V} de codimension complexe 1.

La **variété discriminante** $\Sigma \subset \mathbb{P}(\mathcal{H}_d)$ est la projection de Σ' sur $\mathbb{P}(\mathcal{H}_d)$. C'est l'ensemble des systèmes qui possèdent une racine multiple. La variété discriminante est une sous-variété algébrique de $\mathbb{P}(\mathcal{H}_d)$, elle aussi, de codimension complexe 1.



Cette affaire de **codimension complexe** 1 c'est-à-dire de **codimension réelle** 2 fait que, génériquement, une courbe dans \mathcal{V} reliant deux problèmes bien posés (F_0, z_0) et (F_1, z_1) ne rencontre pas Σ' .

De la même manière, en général, une courbe dans $\mathbb{P}(\mathcal{H}_d)$ reliant deux systèmes F_0 et F_1 sans racines multiples ne rencontre pas Σ . Cette possibilité est à la base des méthodes d'homotopie.

1.4 Le théorème de Bézout

Étienne Bézout naquit à Nemours en 1730 et mourut à Basses-Loges en 1783. On lui doit le théorème suivant :

Théorème 1.2 (Bézout) *Un système polynomial $F \in \mathbb{P}(\mathcal{H}_d) \setminus \Sigma$ possède $B = d_1 \times \cdots \times d_n$ zéros distincts dans $\mathbb{P}_n(\mathbb{C})$; B est appelé le **nombre de Bézout** de $\mathbb{P}(\mathcal{H}_d)$.*

Ce théorème est en défaut si l'on ne considère que les solutions réelles y compris celles à l'infini, ou bien que les solutions affines, réelles ou complexes, comme on peut le voir sur l'exemple des deux cercles.

1.5 Le cas réel

On peut définir les versions réelles des concepts introduits : **systèmes polynomiaux à coefficients réels** : $\mathbb{P}(\mathcal{H}_d(\mathbb{R}))$, **solutions réelles** : $\mathbb{P}_n(\mathbb{R})$, **variété systèmes-solutions** : $\mathcal{V}_{\mathbb{R}} \subset \mathbb{P}(\mathcal{H}_d(\mathbb{R})) \times \mathbb{P}_n(\mathbb{R})$, **variété critique** : $\Sigma'_{\mathbb{R}} \subset \mathcal{V}_{\mathbb{R}}$ et **variété discriminante** : $\Sigma_{\mathbb{R}} \subset \mathbb{P}(\mathcal{H}_d(\mathbb{R}))$.

À la différence du cas complexe $\mathbb{P}(\mathcal{H}_d(\mathbb{R})) \setminus \Sigma_{\mathbb{R}}$ possède au moins B composantes connexes ce qui ne permet pas toujours de connecter deux systèmes réels sans passer par un système singulier. Cette difficulté est amoindrie si l'on se place dans la variété systèmes-solutions $\mathcal{V}_{\mathbb{R}}$. En effet,

Théorème 1.3 (Beltrán-Shub) *$\mathcal{V}_{\mathbb{R}} \setminus \Sigma'_{\mathbb{R}} \subset \mathbb{P}(\mathcal{H}_d(\mathbb{R})) \times \mathbb{P}_n(\mathbb{R})$ est connexe sauf lorsque $n > 1$ et que n et $d_1 + \cdots + d_n - 1$ sont pairs. Dans ce cas $\mathcal{V}_{\mathbb{R}} \setminus \Sigma'_{\mathbb{R}}$ possède 2 composantes connexes.*

Références : Beltrán-Shub [7]

2 Conditionnement

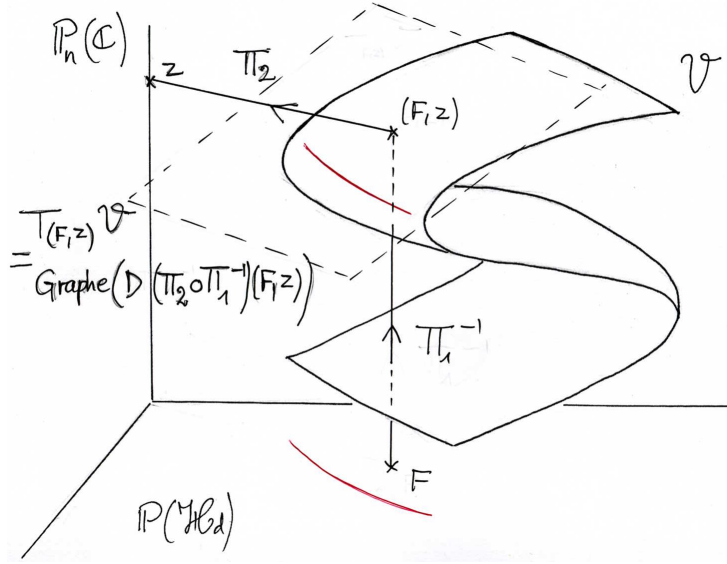
2.1 Variations infinitésimales des zéros d'un système

Soit $(F, z) \in \mathcal{V} \setminus \Sigma'$. Nous avons vu que

$$D\Pi_1(F, z) : T_{(F,z)}V \rightarrow T_F\mathbb{P}(\mathcal{H}_d)$$

est un isomorphisme de sorte que, par le théorème d'inversion locale, on peut inverser localement la projection Π_1 . Par composition avec la projection Π_2 on obtient l'**application solution**

$$\mathcal{S}_{(F,z)} = \Pi_2 \circ \Pi_1^{-1} : V_F \subset \mathbb{P}(\mathcal{H}_d) \rightarrow V_z \subset \mathbb{P}_n(\mathbb{C})$$



qui à tout système G proche de F associe un zéro z_G proche de z . Les **variations au premier ordre** de z en termes des variations de F sont décrites par la dérivée $DS_{(F,z)}(F)$ qui est donnée par

$$DS_{(F,z)}(F) : T_F \mathbb{P}(\mathcal{H}_d) \rightarrow T_z \mathbb{P}_n(\mathbb{C}), \quad \dot{z} = DS_{(F,z)}(F)(\dot{F}) = -(DF(z)|_{z^\perp})^{-1} \dot{F}(z).$$

2.2 Le conditionnement

Les structures riemanniennes introduites sur $\mathbb{P}(\mathcal{H}_d)$ et $\mathbb{P}_n(\mathbb{C})$ induisent une norme d'opérateur sur $\mathcal{L}(T_F \mathbb{P}(\mathcal{H}_d), T_z \mathbb{P}_n(\mathbb{C}))$. Le **conditionnement** de F en z est la norme de l'opérateur $DS_{(F,z)}(F)$:

$$\mu(F, z) = \max_{\dot{F} \in T_F \mathbb{P}(\mathcal{H}_d)} \frac{\|DS_{(F,z)}(F)(\dot{F})\|_z}{\|\dot{F}\|_F} = \|F\| \left\| (DF(z)|_{z^\perp})^{-1} \text{Diag} \left(\|z\|^{d_i-1} \right) \right\|$$

où la dernière norme est la norme d'opérateur associée aux espaces \mathbb{C}^n et \mathbb{C}^{n+1} . On étend cette définition à $(F, z) \in \mathcal{V}$ en posant

$$\mu(F, z) = \infty$$

lorsque (F, z) est mal posé c'est à dire lorsque $DF(z)|_{z^\perp}$ n'est pas inversible.

2.3 Conditionnement et distance aux problèmes mal posés

Le **conditionnement normalisé** est une variante de $\mu(F, z)$ définie par

$$\mu_{norm}(F, z) = \|F\| \left\| (DF(z)|_{z^\perp})^{-1} \text{Diag} \left(\|z\|^{d_i-1} d_i^{1/2} \right) \right\|.$$

Le conditionnement normalisé $\mu_{norm}(F, z)$ s'obtient de façon identique au conditionnement $\mu(F, z)$ mais en utilisant la structure hermitienne

$$\langle F, G \rangle_{(d, norm)} = \sum_{i=1}^n d_i^{-1} \langle F_i, G_i \rangle_{d_i}$$

sur \mathcal{H}_d au lien de la structure produit décrite précédemment. Notons Σ'_z l'**ensemble des problèmes mal posés en z** :

$$\Sigma'_z = \{(G, z) \in \mathcal{V} : (G, z) \in \Sigma'\} = \Sigma' \cap \Pi_2^{-1}(z).$$

Munissons la fibre $\Pi_2^{-1}(z)$ de la distance induite par la distance projective de $\mathbb{P}(\mathcal{H}_d)$: sinus de l'angle fait par deux systèmes mesuré via $\langle \cdot, \cdot \rangle_d$. On pose alors, pour tout $(F, z) \in \mathcal{V}$,

$$\rho(F, z) = \min_{(G, z) \in \Sigma'_z} d_P(F, G).$$

On a

Théorème 2.1 (*Shub-Smale*) *Pour tout $(F, z) \in \mathcal{V}$*

$$\mu_{norm}(F, z) = \frac{1}{\rho(F, z)}.$$

Ce théorème donne une version quantifiée du fait suivant : les problèmes mal conditionnés sont ceux qui sont proches des problèmes mal posés. Et, bien évidemment, les problèmes mal conditionnés sont ceux qui vont poser de gros problèmes numériques.

Références : Blum-Cucker-Shub-Smale [8], Shub-Smale [18], [11].

3 Homotopies

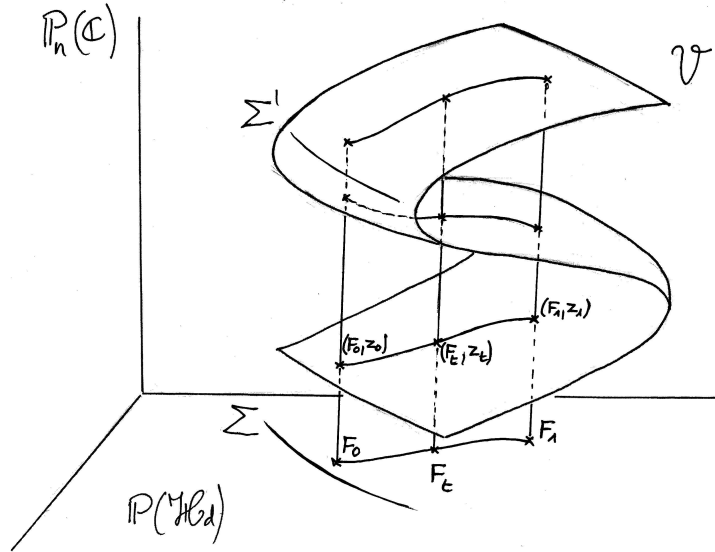
3.1 Principe de base

Étant donné un système $F_1 \in \mathbb{P}(\mathcal{H}_d)$ dont nous souhaitons trouver un zéro $z_1 \in \mathbb{P}_n(\mathbb{C})$, la **méthode d'homotopie** consiste à englober ce problème particulier au sein d'une famille $F_t \in \mathbb{P}(\mathcal{H}_d)$, $0 \leq t \leq 1$, dont nous allons suivre les avatars¹.

Théorème 3.1 *Étant donnés une courbe $F_t \in \mathbb{P}(\mathcal{H}_d) \setminus \Sigma$, $0 \leq t \leq 1$, de classe C^1 et un zéro $z_0 \in \mathbb{P}_n(\mathbb{C})$ de F_0 , il existe une unique courbe $z_t \in \mathbb{P}_n(\mathbb{C})$ qui soit C^1 et qui satisfasse $(F_t, z_t) \in \mathcal{V}$.*

Ainsi, il suffit de suivre la courbe z_t depuis $t = 0$ pour aboutir, en $t = 1$, au z_1 cherché. Ceci pose deux problèmes. Comment transformer cette idée géométrique en un algorithme et quelle est sa complexité ?

¹Stricto sensu, un avatar est une incarnation du dieu Vichnou ou d'une autre divinité de la religion hindoue.



3.2 Exemples

L'homotopie linéaire : $F_t = (1 - t)F_0 + tF_1$, l'homotopie polyédrale qui conserve la structure monomiale donnée c'est-à-dire le polytope de Newton du système que l'on souhaite résoudre, les homotopies multihomogènes. Les géodésiques du conditionnement que nous introduirons plus loin fournissent d'autres exemples.

Références : Allgower-Georg [1], Li [14], Sommese-Wampler [21]

3.3 Prédiction-correction

Au sein du contexte homogène que nous avons choisi, nous voyons que l'équation de base,

$$F_t(z_t) = 0, \quad F_t \text{ et } z_0 \text{ donnés,}$$

est équivalente au problème de condition initiale

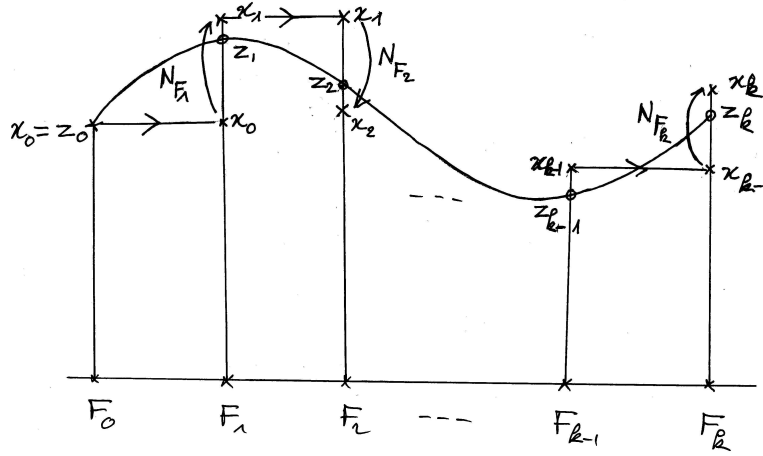
$$\frac{d}{dt} F_t(z_t) = \dot{F}_t(z_t) + DF_t(z_t)(\dot{z}_t) = 0, \quad z_0 \text{ donné,}$$

(nous notons ici \dot{F}_t et \dot{z}_t les dérivées par rapport à t) c'est-à-dire

$$\dot{z}_t = - \left(DF_t(z_t)|_{z_t^\perp} \right)^{-1} \dot{F}_t(z_t), \quad z_0 \text{ donné.}$$

Ainsi z_t peut aussi être décrite comme la solution d'une équation différentielle ordinaire.

Nous discrétisons cette équation en remplaçant l'intervalle $[0, 1]$ par la suite $0 = t_0 < t_1 < \dots < t_k = 1$, les zéros z_{t_i} par des approximations x_i et les dérivées par



rapport à t par des différences divisées. On obtient

$$\frac{x_{i+1} - x_i}{t_{i+1} - t_i} = - (DF_{i+1}(x_i)|_{x_i^\perp})^{-1} \frac{F_{i+1}(x_i) - F_i(x_i)}{t_{i+1} - t_i}$$

et comme $F_i(x_i)$ est sensé être proche de zéro on obtient le schéma

$$x_{i+1} = x_i - (DF_{i+1}(x_i)|_{x_i^\perp})^{-1} F_{i+1}(x_i)$$

que l'on note

$$x_{i+1} = N_{F_{i+1}}(x_i).$$

L'algorithme obtenu est le suivant :

- **Entrée** : F_i , $0 \leq i \leq k$, et x_0 avec $F_0(x_0) = 0$,
- **Itération** : $x_{i+1} = N_{F_{i+1}}(x_i)$, $1 \leq i \leq k - 1$,
- **Sortie** : x_k .

La **complexité** de cet algorithme est mesurée par le nombre k d'étapes nécessaires pour obtenir un **zéro approché** x_k de F_k . Nous allons préciser ce concept.

3.4 La méthode de Newton projective

Pour un système homogène $F(z) = 0$ avec $F : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^n$ la **méthode de Newton projective** (Shub, XX siècle) est définie par

$$N_F(x) = x - (DF(x)|_{x^\perp})^{-1} F(x).$$

Noter que $N_F(z)$ n'est définie que si le rang de $DF(x)$ est maximum, c'est-à-dire n , et que $y = N_F(x)$ n'est autre que la solution de l'équation linéarisée

$$F(x) + DF(x)(y - x) = 0, \quad y - x \in T_x \mathbb{P}_n(\mathbb{C}).$$

On vérifie que c'est une itération bien définie sur $\mathbb{P}_n(\mathbb{C})$.

Nous dirons que $x \in \mathbb{P}_n(\mathbb{C})$ est un **zéro approché de F associé au zéro z** lorsque la suite de Newton projective $x_0 = x$, $x_{i+1} = N_F(x_i)$, est définie et vérifie

$$d_P(x_i, z) \leq \left(\frac{1}{2}\right)^{2^i - 1} d_P(x_0, z).$$

Ici d_P est le sinus de l'angle des droites issues des points considérés.

L'**invariant gamma** est défini pour tout $(F, z) \in \mathcal{V}$ par

$$\gamma(F, z) = \|z\| \sup_{k \geq 2} \left\| \frac{(DF(z)|_{z^\perp})^{-1} D^k F(z)}{k!} \right\|^{\frac{1}{k-1}}$$

ou bien ∞ lorsque $DF(z)$ n'est pas de rang maximum. Il permet de donner la taille d'une boule centrée en z constituée de zéros approchés de F associés à z :

Théorème 3.2 *Soit $(F, z) \in \mathcal{V}$ et soit $x \in \mathbb{P}_n(\mathbb{C})$ tel que*

$$d_P(z, x) \leq \frac{3 - \sqrt{7}}{2\gamma(F, z)}.$$

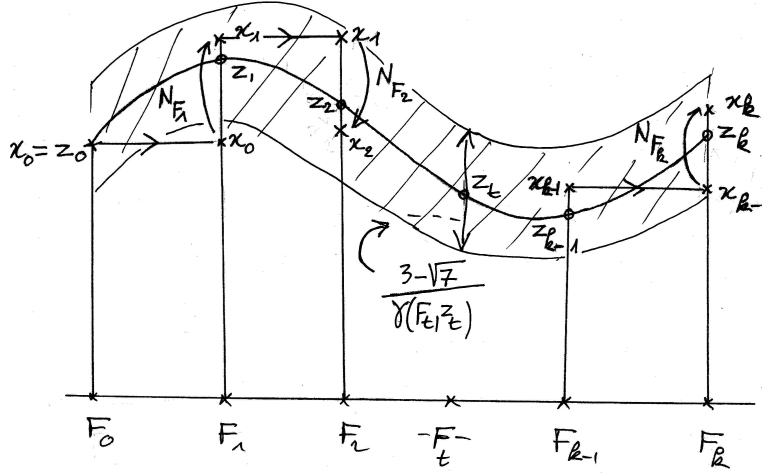
Alors x est un zéro approché de F associé à z .

Un second résultat permet d'estimer l'invariant gamma en termes du conditionnement :

Théorème 3.3 *Soient $(F, z) \in \mathcal{V}$ et $D = \max d_i$. On a :*

$$\gamma(F, z) \leq \frac{D^{3/2}}{2} \mu_{norm}(F, z).$$

Références : Blum-Cucker-Shub-Smale [8], Shub-Smale [18], Dedieu [12], [11]



3.5 Retour aux méthodes de prédiction-correction

Les résultats du paragraphe précédent montrent que la suite construite par la méthode de prédiction-correction ne réalisera notre objectif que si

$$d_P(z_i, x_i) \leq \frac{3 - \sqrt{7}}{2\gamma(F_i, z_i)}, \quad 0 \leq i \leq k,$$

ce qui sera réalisé lorsque

$$d_P(z_i, x_i) \leq \frac{3 - \sqrt{7}}{D^{3/2} \mu_{\text{norm}}(F_i, z_i)} = (3 - \sqrt{7}) D^{-3/2} \rho(F_i, z_i), \quad 0 \leq i \leq k.$$

On voit bien que l'on a tout intérêt à rester éloigné des problèmes mal posés, sans quoi il faudra faire de tout² petits pas et accroître le nombre d'étapes de notre algorithme.

Un résultat essentiel de complexité est alors donné par :

Théorème 3.4 (Shub-Smale) *Étant donné une courbe $F_t \in \mathbb{P}(\mathcal{H}_d) \setminus \Sigma$, $0 \leq t \leq 1$, un zéro z_0 of F_0 et la courbe relevée correspondante $(F_t, z_t) \in \mathcal{V} \setminus \Sigma'$, il existe une subdivision*

$$0 = t_0 < t_1 < \dots < t_k = 1$$

telle que la suite x_i construite par l'algorithme de prédiction-correction ci-dessus soit constituée de zéros approchés des F_i correspondant aux zéros z_i et que

$$k \leq CD^{3/2} \mu_{\text{norm}}(F, z)^2 L_F.$$

²"tout petits pas" et pas "tous petits pas" !

C est une constante absolue, L_F la longueur de la courbe F_t dans $\mathbb{P}(\mathcal{H}_d)$ et

$$\mu_{norm}(F, z) = \sup_{0 \leq t \leq 1} \mu_{norm}(F_t, z_t)$$

est le conditionnement de la courbe relevée.

3.6 Remarques

- $L_F = \int_0^1 \left\| \dot{F}_t \right\|_{F_t} dt$.
- Pour une homotopie linéaire $L_F \leq \pi/2$ le diamètre de $\mathbb{P}(\mathcal{H}_d)$.
- Le terme dominant de cette borne est $\mu_{norm}(F, z)^2$. Son estimation statistique dans le cas d'homotopies linéaires lorsqu'une extrémité est fixée et que l'autre varie a fait l'objet de nombreux travaux.

Références : Blum-Cucker-Shub-Smale [8], Shub-Smale [18], [19], [20], Beltrán-Pardo [4], [5] .

4 Géodésiques du conditionnement

4.1 Motivations

Nous avons vu que la complexité des algorithmes de prédiction-correction était plombée par le passage près des problèmes mal posés. Pour éviter cet inconvénient, nous souhaitons suivre des courbes dans la variété problèmes-solutions qui s'écartent de la variété critique tout en restant aussi courtes que possible !

Un cas modèle est celui de la **géométrie hyperbolique**. Notons

$$\mathbb{H} = \{z \in \mathbb{C} : \Im z > 0\}$$

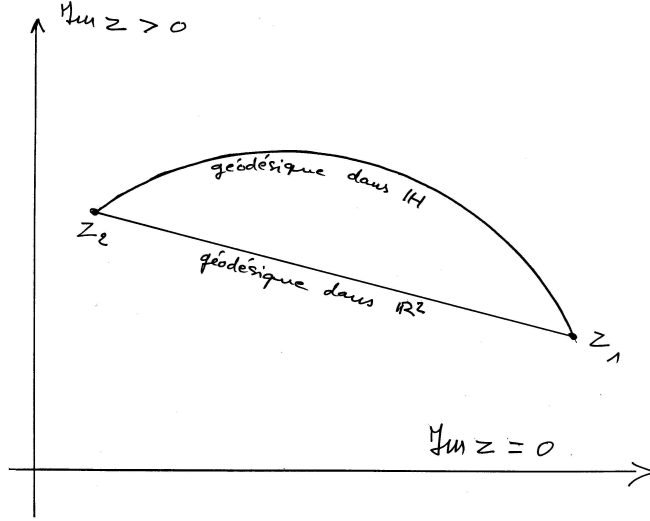
le demi-plan de Poincaré. Un arc de courbe absolument continu $z(t) \in \mathbb{H}$, $0 \leq t \leq 1$, a pour longueur euclidienne

$$L(z) = \int_0^1 |\dot{z}(t)| dt$$

et l'arc ayant la longueur la plus courte parmi ceux qui relient deux points donnés z_0 et $z_1 \in \mathbb{H}$ est le segment reliant ces deux points.

Si l'on rajoute l'exigence **rester loin du lieu critique** défini par $\Im z = 0$, on est amené à considérer l'intégrale

$$L_\kappa(z) = \int_0^1 \frac{|\dot{z}(t)|}{\Im z(t)} dt.$$



La courbe optimale reliant z_0 et $z_1 \in \mathbb{H}$ au sens de cette nouvelle longueur n'est plus le segment qui les relie mais l'arc de cercle passant par ces points et centré sur l'axe réel $\Re z = 0$. C'est le meilleur compromis entre longueur la plus courte et éloignement de l'axe des abscisses.

4.2 La métrique du conditionnement

Nous allons faire exactement la même chose en considérant, dans la variété problèmes-solutions \mathcal{V} , non plus la longueur associée à la structure riemannienne habituelle

$$L(F, z) = \int_0^1 \left\| \frac{d}{dt} (F_t, z_t) \right\|_{(F_t, z_t)} dt$$

avec

$$\left\| (\dot{F}, \dot{z}) \right\|_{(F, z)}^2 = \frac{\|\dot{F}\|^2}{\|F\|^2} + \frac{\|\dot{z}\|^2}{\|z\|^2}$$

où $(\dot{F}, \dot{z}) \in T_{(F, z)}\mathcal{V}$, mais la **longueur du conditionnement**

$$L_\kappa(F, z) = \int_0^1 \left\| \frac{d}{dt} (F_t, z_t) \right\|_{(F_t, z_t)} \mu_{norm}(F_t, z_t) dt.$$

De cette longueur nous déduisons un concept de distance ; la **distance du conditionnement** entre deux paires (F_0, z_0) et $(F_1, z_1) \in \mathcal{V} \setminus \Sigma'$ est l'infimum des longueurs des courbes absolument continues qui relient (F_0, z_0) et (F_1, z_1) :

$$d_\kappa((F_0, z_0), (F_1, z_1)) =$$

$$\inf \{L_\kappa(F, z) : (F(0), z(0)) = (F_0, z_0), (F(1), z(1)) = (F_1, z_1)\}.$$

Une courbe qui réalise ce minimum, reparamétrisée³ par l'abscisse curviligne, est appelée **géodésique du conditionnement**.

4.3 Métrique du conditionnement et complexité

La métrique du conditionnement est un outil naturel pour mesurer la complexité des méthodes de prédiction-correction comme le montre le résultat suivant :

Théorème 4.1 (*Shub*) *Étant donné une courbe C^1 , $(F_t, z_t) \in \mathcal{V} \setminus \Sigma'$, $0 \leq t \leq 1$,*

$$k \leq CD^{3/2}L_\kappa(F, z)$$

étapes de l'algorithme de prédiction-correction sont suffisantes pour réaliser notre programme de calcul de zéros approchés.

Comme

$$L_\kappa(F, z) \leq \mu_{norm}(F, z)L(F, z) \leq \mu_{norm}(F, z)\mu(F, z)L_F \leq \mu_{norm}(F, z)^2L_F$$

la borne obtenue ici améliore celle du théorème 3.4.

$$CD^{3/2}L_\kappa(F) \leq CD^{3/2}\mu_{norm}(F, z)^2L_F.$$

Nous avons donc tout intérêt à suivre les géodésiques du conditionnement !

L'exemple suivant montre que de telles courbes peuvent avoir une longueur très courte. Posons

$$G = (d_1^{1/2}z_0^{d_1-1}z_1, \dots, d_n^{1/2}z_0^{d_n-1}z_n), \text{ and } e_0 = (1, 0, \dots, 0).$$

Théorème 4.2 (*Beltrán-Shub*) *Soit $(F, \zeta) \in \mathcal{V} \setminus \Sigma'$. Il existe une courbe qui connecte (G, e_0) and (F, ζ) et dont la longueur pour la métrique du conditionnement vérifie*

$$L_\kappa \leq 9nD^{3/2} + 2\sqrt{n} \ln \left(\frac{\mu_{norm}(F, \zeta)}{\sqrt{n}} \right).$$

Suivant une telle courbe, nous obtiendrions une complexité bornée par le logarithme du conditionnement du système à atteindre au lieu du carré du conditionnement de la courbe suivie ! Cette construction a hélas un point faible : elle suppose que notre problème initial est déjà résolu !

Références : Beltrán-Shub [17], [6]

³Du verbe "paramétrer". Le verbe "paramétriser" est à proscrire.

4.4 Extension du concept

Considérons, au lieu de \mathcal{V} , une variété riemannienne

$$(\mathcal{M}, \langle \cdot, \cdot \rangle)$$

de classe C^2 . Donnons nous une fonction

$$\alpha : \mathcal{M} \rightarrow \mathbb{R}$$

localement lipschitzienne et à valeurs positives. Notons \mathcal{M}_κ la variété \mathcal{M} équipée de la nouvelle métrique

$$\langle \cdot, \cdot \rangle_{\kappa, x} = \alpha(x) \langle \cdot, \cdot \rangle_x$$

que nous appelons **métrique du conditionnement**.

De même que pour \mathcal{V} on associe à \mathcal{M}_κ un concept de longueur L_κ , une distance d_κ et un concept de géodésique. On obtient alors un **espace de longueur** à la Gromov.

Les exemples que nous avons en vue sont les suivants

1. \mathcal{V} et $\alpha(F, z) = \mu_{norm}(F, z)^2$,
2. \mathbb{H} et $\alpha(z) = (\Im z)^{-2}$,
3. Pour deux entiers $1 \leq n \leq m$, notons par $\mathbb{GL}_{n,m}$ les matrices $X \in \mathbb{K}^{n \times m}$ de rang maximum : $\text{rang } X = n$, $\mathbb{K} = \mathbb{R}$ or \mathbb{C} . Les valeurs singulières de $X \in \mathbb{GL}_{n,m}$ sont rangées par ordre décroissant :

$$\sigma_1(X) \geq \dots \geq \sigma_{n-1}(X) \geq \sigma_n(X) > 0.$$

La métrique du conditionnement dans $\mathbb{GL}_{n,m}$ est définie par

$$\langle \cdot, \cdot \rangle_{\kappa, X} = \langle \cdot, \cdot \rangle_F \sigma_n(X(t))^{-2}$$

où

$$\|M\|_F^2 = \sum |m_{ij}|^2$$

est la norme de Frobenius de la matrice M .

4. Soit \mathcal{N} une sous-variété sans bord de classe C^k , $k \geq 2$, contenue dans \mathbb{R}^j . Posons

$$d(x, \mathcal{N}) = \inf_{y \in \mathcal{N}} \|x - y\|$$

la distance euclidienne de $x \in \mathbb{R}^j$ à \mathcal{N} .

$$\mathcal{U} = \mathbb{R}^j \setminus \mathcal{N}, \quad \alpha(x) = d(x, \mathcal{N})^{-2}$$

fournit un dernier exemple.

4.5 Quelques problèmes

1. Théorème de Hopf-Rinow : dans de tels espaces \mathcal{M}_κ sous quelles conditions existe-t-il une géodésique reliant deux points donnés ? Il suffit que \mathcal{M}_κ soit complet et localement compact.
2. De telles géodésiques sont-elles uniques ?
3. Quelle est leur régularité ?
4. Comment les caractériser ?

5 Propriétés de la métrique du conditionnement

Nous avons équipé $\mathcal{V} \setminus \Sigma'$ de la longueur L_κ qui nous a servi à construire la distance d_κ . On a :

Théorème 5.1 (*Shub*) $\mathcal{V} \setminus \Sigma'$ est, pour la "distance" d_κ , un espace métrique complet, localement compact, et l'infimum définissant d_κ est un minimum.

Une difficulté sérieuse, pour prouver ce théorème, provient du fait la métrique d_κ n'est pas une métrique riemannienne. Cela est dû au fait que μ_{norm} **n'est pas de classe C^1** mais seulement **localement lipschitzienne**.

On a un théorème identique pour \mathbb{H} et pour $\mathbb{GL}_{n,m}$.

Comme dans le cas riemannien, on peut étendre le concept de **géodésique** à des courbes absolument continues qui sont paramétrées par l'abscisse curviligne et qui sont **localement des géodésiques minimisantes**. On démontre alors

Théorème 5.2 (*Beltran-Boito-Dedieu-Malajovich-Shub-Pugh*) Soit \mathcal{M} une variété différentiable de classe C^2 équipée d'une structure Lipschitz-riemannienne. Toute géodésique est de classe $C^1 + Lip = W^{2,\infty}$.

Références : Beltrán-Shub [17], [6], Beltrán-Dedieu-Malajovich-Shub [2], [3], Boito-Dedieu [9], Pugh [16]

6 Comment les calculer ?

Le calcul des géodésiques minimisantes peut être modélisé de la façon suivante :

$$\min \int_0^1 \left\| \frac{d}{dt} (F_t, z_t) \right\|_{(F_t, z_t)} \mu_{norm}(F_t, z_t) dt$$

sous les contraintes

$$\begin{cases} (F, z) \in W^{1,1}([0, 1], \mathcal{V}), \\ (F_0, z_0) \text{ et } (F_1, z_1) \text{ donnés.} \end{cases}$$

Il s'agit là d'une instance du **problème de Bolza**

$$\min \int_0^1 H(X(t), \dot{X}(t)) dt$$

où $X : [0, 1] \rightarrow \mathcal{M} \subset \mathbb{R}^N$ est absolument continue, à valeurs dans une sous-variété différentiable \mathcal{M} de \mathbb{R}^N , et vérifie les conditions aux limites : $X(0)$ et $X(1)$ donnés.

Lorsque l'**intégrande** H est régulier et lorsque $\mathcal{M} = \mathbb{R}^N$, les solutions du problème de Bolza satisfont l'**équation d'Euler-Lagrange** avec conditions aux limites :

$$\begin{cases} -\frac{d}{dt} \frac{\partial}{\partial \dot{X}} H(X, \dot{X}) + \frac{\partial}{\partial X} H(X, \dot{X}) = 0 \\ X(0) = X_0, \quad X(1) = X_1. \end{cases}$$

La longueur du conditionnement échappe à ce contexte puisque $H(X, \dot{X})$ n'est pas différentiable en la variable X mais seulement localement lipschitzienne. On obtient alors, en utilisant le **gradient généralisé de Clarke** une **inclusion différentielle**.

Soit f une application localement lipschitzienne définie sur un ouvert Ω d'un espace euclidien \mathbb{E} . La **dérivée directionnelle de f en x dans la direction $d \in \mathbb{E}$, $d \neq 0$** , est

$$f^\circ(x, d) = \limsup_{\substack{y \rightarrow x \\ t \rightarrow 0_+}} \frac{f(y + td) - f(y)}{t}$$

et le **gradient généralisé** de f en x est le sous-ensemble convexe, compact, non vide de \mathbb{E} donné par

$$\partial f(x) = \{s \in \mathbb{E} : \langle s, d \rangle \leq f^\circ(x, d) \text{ for all } d \in \mathbb{E}\}.$$

Lorsque $f \in C^1(\Omega)$ le gradient généralisé est identique au gradient usuel :

$$\partial f(x) = \{\nabla f(x)\}.$$

Plus généralement (Théorème de Rademacher)

$$\partial f(x) = \text{co} \left\{ \lim_{i \rightarrow \infty} \nabla f(x_i) : x_i \rightarrow x, \nabla f(x_i) \text{ existe et } \lim_{i \rightarrow \infty} \nabla f(x_i) \text{ existe} \right\}.$$

6.1 L'équation d'Euler-Lagrange généralisée

La solution du problème de Bolza décrit ci-dessus est solution d'une inclusion différentielle comme le montre le

Théorème 6.1 (Clarke) *Supposons que $H(X, \dot{X})$ soit localement lipschitzienne, que X soit une solution du problème de Bolza et que $\dot{X} \in L^\infty([0, 1])$. Alors, il existe une application absolument continue $P : [0, 1] \rightarrow \mathbb{R}^N$ telle que*

$$\begin{cases} \dot{P}(t) \in \partial_X H(X(t), \dot{X}(t)) \\ P(t) \in \partial_{\dot{X}} H(X(t), \dot{X}(t)) \text{ a.e.} \end{cases}$$

Les méthodes de calcul de problèmes aux limites sont fondées sur des **méthodes de tir** ce qui nécessite à chaque étape la résolution d'un problème de condition initiale. De telles méthodes sont inapplicables ici parce que le problème de condition initiale associé à une inclusion différentielle n'a pas, en général, de solution unique.

Pour ces raisons, nous avons choisi d'utiliser

1. Une approximation de l'espace

$$\{X \in W^{1,1}([0, 1], \mathcal{M}) : X(0) \text{ et } X(1) \text{ donnés}\}$$

par un espace de dimension finie \mathbb{E} ,

2. Une approximation $S(X)$ de l'intégrale $\int_0^1 H(X(t), \dot{X}(t)) dt$,
3. Une méthode de résolution du problème

$$\min_{X \in \mathbb{E}} S(X).$$

6.2 Le cas linéaire

Donnons nous deux entiers $1 \leq n \leq m$ et notons par $\mathbb{GL}_{n,m}$ les matrices $A \in \mathbb{K}^{n \times m}$ de rang maximum : $\text{rang } A = n$, $\mathbb{K} = \mathbb{R}$ or \mathbb{C} . Les valeurs singulières de $A \in \mathbb{GL}_{n,m}$ sont rangées par ordre décroissant :

$$\sigma_1(A) \geq \dots \geq \sigma_{n-1}(A) \geq \sigma_n(A) > 0.$$

La métrique du conditionnement dans $\mathbb{GL}_{n,m}$ est définie via l'intégrale

$$L_\kappa(X) = \int_0^1 \left\| \frac{d}{dt} X(t) \right\|_F \sigma_n(X(t))^{-1} dt$$

où $X(t) \in \mathbb{GL}_{n,m}$ est une courbe absolument continue et où

$$\|M\|_F^2 = \sum |m_{ij}|^2$$

est la norme de Frobenius de la matrice M .

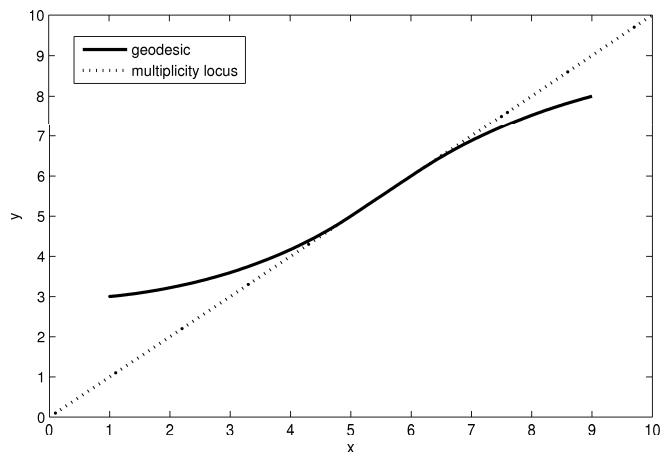
L'approximation de $X(t)$ a été réalisée par des expressions du type

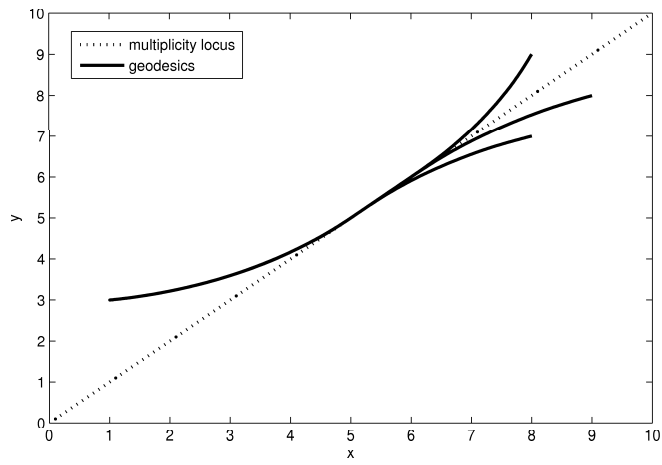
$$X_0 + (X_1 - X_0)t + A_1 \sin(\pi t) + A_2 \sin(2\pi t) + \cdots + A_N \sin(N\pi t)$$

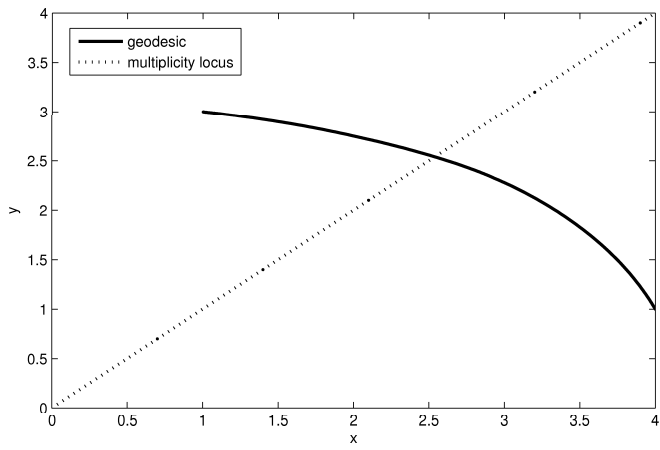
(avec $N = 10$) et le problème d'optimisation correspondant a été résolu par l'algorithme BFGS. Les résultats obtenus ont toujours été en accord avec nos connaissances théoriques.

Les trois dessins qui suivent montrent des exemples de géodésiques du conditionnement dans l'espace des matrices 2×2 symétriques et définies positives. Elles sont constituées d'arcs de cercles et de segments de droites, ces courbes sont de classe C^1 .

Références : Beltrán-Dedieu-Malajovich-Shub [2], [3], Boito-Dedieu [9], Overton [15]







7 Auto-convexité

7.1 Définition

On veut savoir si, le long d'une géodésique pour la métrique du conditionnement, on évite effectivement les problèmes mal conditionnés. On peut reformuler cette question de la façon suivante :

Étant donné une telle géodésique (F_t, z_t) dans $\mathcal{V} \setminus \Sigma'$, est-ce que le maximum de $\mu_{norm}(F_t, z_t)$ est nécessairement atteint aux extrémités de cet arc ? C'est-à-dire : est-ce que $t \rightarrow \mu_{norm}(F_t, z_t)$ est une fonction **quasi-convexe** ?

Une propriété plus forte est la **convexité** de $t \rightarrow \mu_{norm}(F_t, z_t)$.

Une propriété encore plus forte est la **log-convexité** de $t \rightarrow \mu_{norm}(F_t, z_t)$ c'est-à-dire la convexité de $t \rightarrow \log(\mu_{norm}(F_t, z_t))$

Nous appelons **auto-convexité du conditionnement** cette dernière propriété.

7.2 Extension

Le concept d'auto-convexité peut être étendu à des situations plus générales. On se donne une variété riemannienne

$$(\mathcal{M}, \langle \cdot, \cdot \rangle)$$

et une fonction $\alpha : \mathcal{M} \rightarrow]0, \infty[$ localement lipschitzienne. Nous définissons (voir ci-dessus) \mathcal{M}_κ , par

$$(\mathcal{M}, \langle \cdot, \cdot \rangle_m \alpha(m))$$

Nous dirons que α est **auto-convexe** lorsque $\log \circ \alpha$ est convexe le long de toute géodésique de \mathcal{M}_κ .

On caractérise cette propriété dans le cas lisse par :

Proposition 7.1 *Lorsque $\alpha : \mathcal{M} \rightarrow \mathbb{R}$ est de classe C^2 , l'auto-convexité de α est équivalente à*

$$2\alpha(x)D^2\alpha(x)(\dot{x}, \dot{x}) + \|D\alpha(x)\|_x^2 \|\dot{x}\|_x^2 - 4(D\alpha(x)\dot{x})^2 \geq 0$$

pour tout $x \in \mathcal{M}$ et pour tout vecteur $\dot{x} \in T_x\mathcal{M}$, l'espace tangent en x .

Il est facile de voir que $1/\Im z$ est auto-convexe dans le demi-plan de Poincaré \mathbb{H} .

Un autre exemple est donné par $\alpha(x) = d(x, \mathcal{N})^{-2}$ lorsque l'on se restreint à l'ensemble \mathcal{M} des points de $\mathcal{U} = \mathbb{R}^j \setminus \mathcal{N}$ qui ont une unique projection sur \mathcal{N} .

7.3 Auto-convexité dans $\mathbb{GL}_{n,m}$

Rappelons que la métrique du conditionnement dans $\mathbb{GL}_{n,m}$ est donnée par le produit scalaire sur $T_X \mathbb{GL}_{n,m}$ défini par

$$\langle \cdot, \cdot \rangle_{F\sigma_n(X)^{-2}}.$$

Théorème 7.1 (*Beltrán-Dedieu-Malajovich-Shub*) σ_n^{-1} est auto-convexe dans $\mathbb{GL}_{n,m}$.

La preuve est (à l'heure actuelle) particulièrement longue et difficile. Elle utilise trois arguments principaux :

1. Un découpage de $\mathbb{GL}_{n,m}$ en sous-variétés différentiables dans lesquelles les valeurs singulières sont de mêmes multiplicités,
2. L'application de la proposition précédente à ce cas lisse,
3. Une technique de recollement.

Références : Beltrán-Dedieu-Malajovich-Shub [2], [3]

8 Quelques problèmes ouverts

1. Comment étendre au cas polynomial le calcul des géodésiques effectué dans le cas linéaire ? (Travail en cours).
2. La propriété d'auto-convexité, ou une propriété plus faible, est-elle satisfaite dans le cas polynomial ?
3. Les géodésiques du conditionnement dont les extrémités sont données et suffisamment proches sont-elles uniques ?
4. Peut-on étendre ces résultats à des structures LR générales et pas seulement aux structures conformes ?

References

- [1] E. ALLGOWER, K. GEORG, *Numerical continuation methods*. Springer (1990).
- [2] C. BELTRÁN, J.-P. DEDIEU, G. MALAJOVICH, AND M. SHUB, *Convexity properties of the condition number*. SIAM. J. Matrix Anal. Appl. Volume 31, Issue 3, pp. 1491-1506 (2010).

- [3] C. BELTRÁN, J.-P. DEDIEU, G. MALAJOVICH, AND M. SHUB, *Convexity properties of the condition number II*. Preprint (2010).
- [4] C. BELTRÁN, AND L. M. PARDO, *On Smale's 17th Problem: a Probabilistic Positive Solution*. FOCM, (2008) 1-43.
- [5] C. BELTRÁN, AND L. M. PARDO, *Smale's 17th Problem: Average Polynomial Time to Compute Affine and Projective Solutions*. J. AMS, 22 (2009) 363-385.
- [6] C. BELTRÁN, AND M. SHUB, *Complexity of Bézout's Theorem VII: Distances Estimates in the Condition Metric*, FOCM 9 (2009) 179-195.
- [7] C. BELTRÁN, AND M. SHUB, *On the Geometry and Topology of the Solution Variety for Polynomial System Solving* (2008) <https://sites.google.com/site/beltranc/preprints>
- [8] L. BLUM, F. CUCKER, M. SHUB, AND S. SMALE, *Complexity and Real Computation*, Springer, 1998.
- [9] P. BOITO, AND J.-P. DEDIEU, *The condition metric in the space of rectangular full rank matrices*. To appear in SIMAX.
- [10] CLARKE F., *Optimization and Nonsmooth Analysis*. J. Wiley and Sons, 1983.
- [11] J.-P. DEDIEU, *Approximate Solutions of Numerical Problems, Condition Number Analysis and Condition Number Theorems*. In : The Mathematics of Numerical Analysis, J. Renegar, M. Shub, S. Smale editors, Lectures in Applied Mathematics, Vol. 23, American Mathematical Society, 1996.
- [12] J.-P. DEDIEU, *Points fixes, zéros et la méthode de Newton*. Mathématiques et Applications, Springer, 2006.
- [13] GROMOV M., *Metric Structures for Riemannian and Non-Riemannian Spaces*, Birkhauser, third printing 2007.
- [14] T. Y. LI, *Numerical solution of multivariate polynomial systems by homotopy continuation methods*, Acta Numerica 6 (1997), 399-436.
- [15] OVERTON M., An implementation of the BFGS method, <http://cs.nyu.edu/overtton/software/index.html>
- [16] PUGH C., *Lipschitz Riemann Structures*. Private communication, 2007.
- [17] M. SHUB, *Complexity of Bézout's Theorem VI: Geodesics in the Condition Metric*, FOCM 9 (2009) 171-178.

- [18] M. SHUB, AND S. SMALE, *Complexity of Bézout's Theorem I: Geometric Aspects*, J. Am. Math. Soc. (1993) 6 pp. 459-501.
- [19] M. SHUB, AND S. SMALE, *Complexity of Bézout's Theorem II : Volumes and Probabilities*, in : F. Eyssette, A. Galligo Eds. Computational Algebraic Geometry, Progress in Mathematics. Vol. 109, Birkhäuser, (1993).
- [20] M. SHUB, AND S. SMALE, *Complexity of Bézout's Theorem V : Polynomial Time*, Theoretical Computer Science, 133, 141-164 (1994).
- [21] A. SOMMESE, C. WAMPLER, *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, 2005.