My view on the work of Jean–Pierre Dedieu

Luis M. Pardo

Univ. Cantabria

October 22, 2009

Jean–Pierre Dedieu doing Maths...

Preamble

* When the organizers asked me to do a talk like this....

Preamble

* When the organizers asked me to do a talk like this....I couldn't say NO, but...

Preamble

* When the organizers asked me to do a talk like this....I couldn't say NO, but...

* Just fifty minutes is unfair to summarize a life devoted to mathematical science.

Preamble

* When the organizers asked me to do a talk like this....I couldn't say NO, but...

* Just fifty minutes is unfair to summarize a life devoted to mathematical science.

* This is then a portrait done with a broad brush.

* I hope Jean–Pierre, dont feel upset or disappointed: Our friendship is more important...

Introduction

Q.: How can you “classify” Jean–Pierre Dedieu’s work?

Q.: How can you "classify" Jean–Pierre Dedieu's work?
A.1: Not inside the classical boxes of known standard Mathematics Subject Classifications.

Introduction

Q.: How can you "classify" Jean–Pierre Dedieu's work?
A.1: Not inside the classical boxes of known standard Mathematics Subject Classifications.
A.2: His research belongs to the interdisciplinary field of Foundations of Computational Mathematics.

Introduction

Q.: How can you "classify" Jean–Pierre Dedieu's work?
A.1: Not inside the classical boxes of known standard Mathematics Subject Classifications.
A.2: His research belongs to the interdisciplinary field of Foundations of Computational Mathematics.
A.3: Most of the questions and problems are brand new.

Introduction

Q.: How can you "classify" Jean–Pierre Dedieu's work?
A.1: Not inside the classical boxes of known standard Mathematics Subject Classifications.
A.2: His research belongs to the interdisciplinary field of Foundations of Computational Mathematics.
A.3: Most of the questions and problems are brand new.
There is neither an obvious tradition, nor established "great schools" existing "from long ago".
People come from different mathematical origins, technical fields and cultural backgrounds.

Introduction

Q.: How can you "classify" Jean–Pierre Dedieu's work?
A.1: Not inside the classical boxes of known standard Mathematics Subject Classifications.
A.2: His research belongs to the interdisciplinary field of Foundations of Computational Mathematics.
A.3: Most of the questions and problems are brand new.
There is neither an obvious tradition, nor established "great schools" existing "from long ago".
People come from different mathematical origins, technical fields and cultural backgrounds.
Research must face "new" challenges with "old" (and new) knowledge, and this yields original material.

Introduction

Q.: How can you “classify” Jean–Pierre Dedieu’s work?
A.1: Not inside the classical boxes of known standard Mathematics Subject Classifications.
A.2: His research belongs to the interdisciplinary field of Foundations of Computational Mathematics.
A.3: Most of the questions and problems are brand new.
There is neither an obvious tradition, nor established “great schools” existing “from long ago”.
People come from different mathematical origins, technical fields and cultural backgrounds.
Research must face “new” challenges with “old” (and new) knowledge, and this yields original material.
And Jean–Pierre’s research belong to this.

For instance...

* Jean–Pierre was "born" in a Functional Analysis "neighborhood" (with M. Atteia in the late seventies)
* He is interested in Applied Mathematics (as a main motivation)
* He found some Lower Complexity Bounds (as in his work with S. Smale)
* He worked on Models of Computation (as the work on Decision machines and round off with F. Cucker)
* He went ahead with (Quantitative) Numerical Algebraic Geometry (we see later)
* He is one of the main supporters of Numerical Analysis in Riemannian Geometry (this needs some more detail)
* He produced deep results in Linear Optimization (I will devote a while to details)
* ...

For instance...

* Jean–Pierre was "born" in a Functional Analysis "neighborhood" (with M. Atteia in the late seventies)
* He is interested in Applied Mathematics (as a main motivation)
* He produced works on Lower Complexity Bounds (as in his work with S. Smale)
* He worked on Models of Computation (as the work on Decision machines and round off with F. Cucker)
* He pushed ahead (Quantitative) Numerical Algebraic Geometry (we see later)
* He is one of the main supported of Numerical Analysis in Riemannian Geometry (this needs some more detail)
* He produced deep results in Linear Optimization (I will devote a while to details)
* ...

Too many topics for just fifty minutes!!

Long list of collaborators (I hope none's been forgotten)

* Adler, Roy L.
* Armentano, Diego
* Atteia, Marc
* Bellido, Anne-Mercedes
* Beltrán, Carlos
* Boito, Paula
* Cucker, Felipe
* Darracq-Calmettes, Marie-Cécile
* Favardin, Ch.
* Gourdon, Xavier
* Gregorac, Robert J.
* Kim, Myong-Hi
* Li, Chong
* Malajovich, Gregorio
* Margulies, Joseph Y.
* Martens, Marco
* Nowicki, Dmitry
* Piétrus, Alain
* Priouret, Pierre
* Roy, Marie-Françoise
* Shub, Michael
* Smale, Stephen J.
* Tisseur, Françoise
* Wang, Jin Hua
* Yakoubsohn, Jean-Claude

The Talk

1. Introduction
2. "Early" Works: Functional Analysis and Non–Convex Optimization
2. Back to research: Miscellanea
2. 1995–2000: Quantitative aspects in Numerical Algebraic Geometry
3. Extending $\alpha-$ and $\gamma-$Theories to Riemannian Manifolds and Lie Groups.
4. Optimization and Linear Programming
5. Some Lower Complexity Bounds

2.- "Early" Works: Functional Analysis and Non–Convex Optimization

70's So far away... "Cône assymptote"

A functional analysis approach to Optimization in non–convex domains.
Let E be a t. v. s., $A \subseteq E$. The Asymptotic Cone of A at $x_0 \in E$:

$$A_\infty := \bigcup_{\lambda > 0} \lambda . (A - x_0) .$$

Theorem (Dedieu, 77–79)

Let $C(A)$ be the cone in $E \times \mathbb{R}$ with vertex 0 and generated by $A \times \{1\}$. Namely,

$$C(A) := \bigcup_{1 \geq \lambda \geq 0} \lambda (A \times \{1\}) .$$

Then, the closure $\overline{C(A)}$ can be determined by:

$$\overline{C(A)} := C(A) \bigcup (A_\infty \times \{0\}) .$$

Other works in the same period with M. Atteia [Atteia-Dedieu, 79–81]

Miscellanea

Several co–authors: A. Bellido, Ch. Favardin, R.J. Gregorac, A. Pietrus, M.F. Roy, J.C.Yakoubsohn...

1988-1995: Miscellanea

The term "Miscellanea" simply means that this speaker does not how to classify several works at that time treating different topics and problems.

- *Quintic spline approximations* (computation, error analysis, related linear algebra, heptic splines, interpolation, convexity...[Dedieu, 88a, 88b, 92]).
- *Univariate polynomial solving* ("Dandelin–Graeffe" in [D89], roots of maximum absolute value [Dedieu–Roy, 89],).
- *Hyperbolic polynomials* ([D91], [D92], [Dedieu-Gregorac, 94]).
- *Computational Geometry* ([D-Favardin, 94], [Bellido-D- Yakoubsohn, 91–92]).
- *Optimization* ("penatlty" optimization in [D92], local minima in [D95]).
- *Weyl's Exclusion Method* [Dedieu–Yakoubsohn, 91,92,93].

Exclusion Method, Dedieu-Yakoubsohn, 92

Problem

Given a Polynomial $P \in \mathbb{R}[X_1, \ldots, X_n]$ and a compact subset $F \subseteq \mathbb{R}^n$, "localize" the hyper–surface $V(P) \cap F = \{x \in F \ : \ P(x) = 0\}$.

The Method: Define:

$$M(x,t) := P(x) - \sum_{t=1}^{k} b_k t^k,$$

where

$$b_k := \frac{1}{k!} \sum_{1 \leq i_1, \ldots, i_k \leq n} \left| \frac{\partial^k P(x)}{\partial x_{i_1} \cdots \partial x_{i_k}} \right| .$$

Exclusion Method, Dedieu-Yakoubsohn, 92

Proposition

The Polynomial $M(x,t)$ *has degree* $\deg(p)$ *with respect to the variable* t*. For each* x *there ia a unique positive real root* $m(x)$ *(i.e.* $M(x, m(x)) = 0$ *such that*

1. *The function* $x \longmapsto m(x)$ *is continuous semi–algebraic and satisfies that for every compact set* $F \subseteq \mathbb{K}^n$*, there exist constants* c_1, c_2 *such that*

$$c_1|P(x)| \leq m(x) \leq c_2|P(x)|^{1/d}, \ \forall x \in F.$$

2. *For every compact semi–algebraic subset* $F \subseteq \mathbb{K}^n$*, there exists* $n_1 \in \mathbb{Z}$ *and there exists* $a_1 \in \mathbb{R}_+$ *such that*

$$a_1 d(x, V)^{n_1} \leq m(x) \leq dist(x, V), \ \forall x \in F.$$

Moreover, we can choose $n_1 = 1$ *if* $F \cap V$ *is smooth.*

Exclusion: The Algorithmn, D–Yakoubsohn, 92

Input P and F.
Pick two sequences converging to zero (r_p) and (ε_p) (for instance, $r_p = 1/p = \varepsilon_p$).
Define a sequence of compact $1/2-$algebraic subsets F_p by:

$$F_0 := F$$

while $F_{p-1} \neq \emptyset$ **do**
find a covering $\{B(x_i^{(p)}, r_p)\}_p$ of F_{p-1}.
Choose $s_i^{(p)} \in \mathbb{R}_+$ such that $m(x_i^{(p)}) - \varepsilon_p \leq s_i^{(p)} \leq m(x_i^{(p)})$.

$$B_i^p := \begin{cases} B(x_i^p, s_i^p) & \textit{if} \quad P(x_i^p) \neq 0 \\ \emptyset & \qquad \textit{otherwise} \end{cases}$$

$$F_p := F_{p-1} \setminus \bigcup B_i^p.$$

od

Exclusion algorithm, D–Yakousohn, 92

The algorithm stops (and they gave estimates on the number of steps) if $F \cap V = \emptyset$ or you get a close picture (using the last computed F_p) of the hyper–surface.

1995-2000: Numerical Algebraic Geometry: Quantitative Aspects

Quantitative Numerical Algebraic Geometry (QNAG)

- *Multi–variate "Dandelin–Graeffe Method* [Dedieu–Gourdon–Yakoubsohn,96].
- *Bounds on the Separation of Zeros of Polynomial Equations* [Dedieu, 97].
- *Condition Number for Sparse Polynomial Systems* [Dedieu, 97]
- *Multi–Homogeneous Systems of Equations* [Dedieu–Shub, 2000a].
- *Over and Under–determined System* [Dedieu–Shub, 2000b, 2000c], [Dedieu, 2000].
- *On Simple Double Zeros* [Dedieu–Shub, 2001].
- *Implicit* $\gamma-$ *Theorem* [Dedieu–Kim–Shub–Tisseur, 2003].

General Context of QNAG

It goes back to [Smale, 81], and were strongly established in a long cooperation between S. Smale and M.Shub since Middle eighties (Univariate Case) till early nineties (the impressive series [Shub–Smale, Béz. I to V]).
Preliminary Notations: $\mathbb{K} = \mathbb{R} \vee \mathbb{C}$
An analytic (resp. polynomial) mapping $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$,($\mathbb{K} = \mathbb{R} \vee \mathbb{C}$). Assume that $0 \in \mathbb{C}^m$ is a regular value and, then, the fiber

$$V(f) := f^{-1}(\{0\}),$$

is a sub–manifold of co–dimension m.
A Goal: Solving Non–Linear Equations

Notations on relevant quantities

- $\gamma(f,\zeta)$ Related to the convergence radius of the inverse mapping (case $m=n$) or to the "safe" radius (around $\zeta \in V(f)$) such that quadratic convergence of Newton's operator is granted.
- $\beta(f,x) := dist(x, N_f(x))$ Measures the distance between a point and its image under Newton's operator, when defined.
- $\alpha(f,x) := \gamma(f,x)\beta(f,x)$ Yields a proximity test (withy respect to a zero) without any knowledge of the zero!!!!!.
- $\mu(f,\zeta)$ "Non–Linear Condition Number": linked to γ, β and α. It helps for error analysis, and (most important aspect) helps to get upper bounds for the complexity of path following methods.

The impressive series Béz. I to V

- M.Shub and S. Smale did a complete study of these quantities for exploring zeros in affine spaces $\mathbb{K}^n$, ($\mathbb{K} = \mathbb{R} \vee \mathbb{C}$) or projective spaces $\mathbb{P}_n(\mathbb{R}), \mathbb{P}_n(\mathbb{C})$
- J. P. Dedieu continued these studies of the quantities $\alpha, \beta, \gamma, \mu$ in a several different extensions that I will discuss with some detail.

[Dedieu–Shub, 2000a]: Multi–Homogeneous, Motivations

Example (Generalized Eigenvalue Problem)

Given matrices $A, B \in \mathcal{M}_n(\mathbb{C})$, *complex numbers* $\alpha, \beta \in \mathbb{C}^n$ *and a point* $x \in \mathbb{C}^n$, *we have the equation*

$$(\alpha B - \beta A)x = 0.$$

This equations is defined by multi–homogeneous polynomial equations.

Example (Evaluation Map)

Homogeneous polynomials $f \in \mathcal{H}_{(d)}$ *(bounded degree) and points* $x \in \mathbb{C}^{n+1}$ *and the equation*

$$eval(f, x) = 0, (f(x) = 0)$$

is also a multi–homogeneous system of equations.

Other motivations: Nash Equilibria....

[Dedieu–Shub, 2000a]: Multi–Homogeneous, Notations

Decompose $\mathbb{C}^{n+1} = \prod_{i=1}^{k} E_i$, where $E_i := \mathbb{C}^{n_i+1}$.
$f \in \mathbb{C}[X_0, \ldots, X_n]$ multi–homogeneous of degree $(d) = (d_1, \ldots, d_k)$ iff f is homogeneous of degree d_i with respect to the group of variables $x_i \in E_i$.

Multi–homogeneous mapping of degree $((d)) := ((d_1), \ldots, (d_m))$

$$f := (f_1, \ldots, f_m) : E \longrightarrow \mathbb{C}^m$$

s.t. f_i is multi–homogeneous of multi–degree (d_i).

The product of projective spaces $\mathbb{P}^{(k)} := \prod_{i=1}^{k} \mathbb{P}(E_i)$. and its tangent space $T_x\mathbb{P}^{(k)} := \prod_{l=1}^{k} T_{x_i}\mathbb{P}(E_i)$,

[Dedieu–Shub, 2000a]: Multi–Hom., Notations, II

$$Df(x)\mid_{T_x}: T_x\mathbb{P}^{(k)} \longrightarrow \mathbb{C}^m.$$

Newton's operator given by:

$$N_f(x) := x - (Df(x)\mid_{T_x})^\dagger f(x).$$

Multi–homogeneous quantities:

$$\gamma(f,x) := \max\left\{1, \sup_{k\geq 2}\left\|(Df(x)\mid_{T_x})^\dagger \frac{D^k f(x)}{k!}\right\|_x^{1/k-1}\right\}.$$

$$\beta(f,x) := \|\,(Df(x)\mid_{T_x})^\dagger f(x)\| = dist(x, N_f(x)).$$

$$\alpha(f,x) := \beta(f,x)\gamma(f,x).$$

[Dedieu–Shub, 00a]: Multi–Homogeneous, α–Theorem

Theorem (α–Theorem, multi-homogeneous case, Dedieu–Shub, 2000a)

There is a universal constant $\alpha_u = \frac{1}{137}$ *such that:*
Given f *and* x *s.t* $\alpha(f,x) \leq \alpha_u$ *and* $Df(x) \mid_{T_x}$ *is onto, then,*
the Newton sequence $x_k := N_f^k(x)$ *is well–defined and satisfies:*

$$dist(x_{k+1}, x_k) \leq \left(\frac{1}{2}\right)^{2^k - 1} \beta(f,x).$$

And there is a zero $\zeta \in E$ *such that*

$$dist(\zeta, x_k) \leq 2\left(\frac{1}{2}\right)^{2^k - 1} \beta(f,x).$$

[Dedieu–Shub, 00a]: Multi–Homogeneous, γ–Theorem

Theorem (γ–Theorem multi-homogeneous case)

There is a universal constant δ_u *such that the following holds:*
Given f *and* $\zeta \in E$ *such that* $f(\zeta) = 0$ *and* $Df(\zeta)\mid_{T_\zeta}$ *is onto, then for every* $z \in E$ *such that*

$$dist(z, \zeta) \leq \frac{\delta_u}{\gamma(f, \zeta)},$$

the following properties hold:
The sequence $z_k := N_f^k(z)$ *is well–defined and there is some zero* $\zeta' \in E$ *of* f *such that*

$$d(\zeta', z_k) \leq 2\left(\frac{1}{2}\right)^{2^k - 1} \beta(f, x).$$

Moreover, δ_u *may be chosen such that:*

$$dist(\zeta', z_k) \leq \left(\frac{1}{2}\right)^{2^k} dist(\zeta, x).$$

Assume $m = dimF = dimT_x$. A path of equations

$$G := \{f_t : \mathbb{C}^{n+1} \longrightarrow F \ : \ t \in [0,1]\}.$$

Assume $m = dimF = dimT_x$. A path of equations

$$G := \{f_t : \mathbb{C}^{n+1} \longrightarrow F \ : \ t \in [0,1]\}.$$

Assume they do not intersect the "discriminant" variety (i.e. 0 is regular for f_t, for every $t \in [0,1]$.).

Assume $m = dimF = dimT_x$. A path of equations

$$G := \{f_t : \mathbb{C}^{n+1} \longrightarrow F \ : \ t \in [0,1]\}.$$

Assume they do not intersect the "discriminant" variety (i.e. 0 is regular for f_t, for every $t \in [0,1]$.).
Then you may lift the curve G of equations to get a curve Γ of equations/solutions:

$$\Gamma := \{(f_t, \zeta_t) \in \mathcal{H}^m_{((d))} \times \mathbb{P}^{(k)} \ : \ f_t(\zeta_t) = 0, \ t \in [0,1]\}.$$

Assume $m = dimF = dimT_x$. A path of equations

$$G := \{f_t : \mathbb{C}^{n+1} \longrightarrow F \ : \ t \in [0,1]\}.$$

Assume they do not intersect the "discriminant" variety (i.e. 0 is regular for f_t, for every $t \in [0,1]$.).
Then you may lift the curve G of equations to get a curve Γ of equations/solutions:

$$\Gamma := \{(f_t, \zeta_t) \in \mathcal{H}^m_{((d))} \times \mathbb{P}^{(k)} \ : \ f_t(\zeta_t) = 0, \ t \in [0,1]\}.$$

Goal: From a zero ζ_0 of f_0 compute a zero ζ_1 of f_1.

Succint Introd. to Path Following, II

Pick a “good” partition of the interval [0,1]:

$$0 = t_0 < t_1 < t_2 < \cdots < t_p = 1.$$

Succint Introd. to Path Following, II

Pick a “god” partition of the interval [0,1]:

$$0 = t_0 < t_1 < t_2 < \cdots < t_p = 1.$$

Starting at $z_0 = \zeta_0$, compute (provided that it is possible):

$$z_{i+1} := N_{f_{t_{i+1}}}(z_i), \;\; 0 \leq i \leq p-1$$

Succint Introd. to Path Following, II

Pick a “god” partition of the interval [0,1]:

$$0 = t_0 < t_1 < t_2 < \cdots < t_p = 1.$$

Starting at $z_0 = \zeta_0$, compute (provided that it is possible):

$$z_{i+1} := N_{f_{t_{i+1}}}(z_i), \quad 0 \leq i \leq p-1$$

Output $z_p = N_{f_{t_p}}(z_{p-1})$

Pick a "god" partition of the interval [0,1]:

$$0 = t_0 < t_1 < t_2 < \cdots < t_p = 1.$$

Starting at $z_0 = \zeta_0$, compute (provided that it is possible):

$$z_{i+1} := N_{f_{t_{i+1}}}(z_i), \;\; 0 \leq i \leq p-1$$

Output $z_p = N_{f_{t_p}}(z_{p-1})$
In the case "everything was right", the output z_p will be an approximate zero of f_1 with associated zero ζ_1.

Succint Introd. to Path Following, II

Pick a "god" partition of the interval [0,1]:

$$0 = t_0 < t_1 < t_2 < \cdots < t_p = 1.$$

Starting at $z_0 = \zeta_0$, compute (provided that it is possible):

$$z_{i+1} := N_{f_{t_{i+1}}}(z_i), \;\; 0 \leq i \leq p-1$$

Output $z_p = N_{f_{t_p}}(z_{p-1})$
In the case "everything was right", the output z_p will be an approximate zero of f_1 with associated zero ζ_1.

The time complexity of the "procedure" essentially depends on p.

Finding sharp bounds for p

Define

$$\gamma(\Gamma) := \sup_{t\in[0,1]} \gamma(f_t, \zeta_t).$$

and

$$\mu(\Gamma) := \sup_{t\in[0,1]} \left\| (Df(x) \mid_{T_x})^{\dagger} \right\|_{\zeta_t}.$$

[Dedieu–Shub, 00a]: Multi–Homogeneous

Theorem (Path Following in the multi–homogeneous case)

There is a partition of the interval $[0,1]$*:*

$$0 = t_0 < t_1 < t_2 < \cdots < t_p = 1,$$

where

$$p := \left\lfloor \frac{2}{\delta_u} \gamma(\Gamma)\mu(\Gamma) L \right\rfloor + 1,$$

and L *is the length of the curve* Γ*, such that the following holds: the sequence* $x_0 = \zeta_0$ *and*

$$x_{i+1} := N_{f_{t_{i+1}}}(x_i)$$

is well–defined and satisfies:

$$dist(x_i, \zeta_{t_i}) \leq \frac{\delta_u}{\gamma(f_{t_i}, \zeta_{t_i})}.$$

Under–Determined and Over–Determined Systems

When the number of equations differs from the number of solutions.

Under–Determined and Over–Determined Systems

$f : \mathbb{C}^n \longrightarrow \mathbb{C}^m$ a mapping, $f := (f_1, \ldots, f_m)$, where $f_i \in \mathbb{C}[X_1, \ldots, X_n]$ of degree at most d_i. Affine zeros:

$$V_{\mathbb{A}}(f) := \{x \in \mathbb{C}^n \ : \ f(x) = 0\}.$$

Assume 0 is a regular value of f (i.e. $V_{\mathbb{A}}(f)$ smooth variety if non–empty). Three cases:

- Geometric Case (also Under–determined): $m < n$ In this case $V_{\mathbb{A}}(f)$ is smooth of co–dimension m.
- Solving Problem : $m = n$ In this case $V_{\mathbb{A}}(f)$ is a zero–dimensional variety.
- Consistency Problem (also Over–determined case) : $m > n$ In this case $V_{\mathbb{A}}(f)$ is (generically in terms of the f's) empty and, if non–empty, it consists of a single point with probability one.

Also treated by [Ben–Israel, 66], [Allgower–Georg, 90], [Shub–Smale, Bez IV], [Sommese–Wampler, NAG]...
Main contributions from [Dedieu-Kim, 02], [Dedieu–Shub, 00b] and [Dedieu, 00].

Under–Determined Geometric Case

Newton's operator as:

$$N_f(x) := x - (Df(x))^{\dagger} f(x).$$

We have α, γ Theories and Theorems:

$$\gamma(f,x) := \sup_{k\geq 2} \left\| (Df(x))^{\dagger} \frac{D^k f(x)}{k!} \right\|^{1/k-1}.$$

$$\beta(f,x) := \| (Df(x))^{\dagger} f(x)\| = \|x - N_f(x)\|.$$

$$\alpha(f,x) := \beta(f,x)\gamma(f,x).$$

Under–Determined $\alpha-$ Theorem

Theorem (Under–determined $\alpha-$Theorem, Shub–Smale)

There is a universal constant $\alpha_0 \approx 0.13071...$ *such that if* $\alpha(f,x) \leq \alpha_0$, *and* $Df(x)$ *is onto, then the sequence* $x_k := N_f^k(x)$ *is well–defined and:*

1. *The sequence* $\{x_k\}$ *is a Cauchy sequence such that*

$$dist(x_{k+1}, x_k) = \beta(f, x_k) \leq \left(\frac{1}{2}\right)^{2^k - 1} \beta(f, x).$$

2. *There is some zero* $\zeta \in V_{\mathbb{A}}(f)$ *such that*

$$dist(x_k, V_{\mathbb{A}}(f)) \leq dist(x_k, \zeta) \leq 2\left(\frac{1}{2}\right)^{2^k - 1} \beta(f, x).$$

Under–Determined $\gamma-$ Theorem, Shub–Smale

Theorem (Under–determined $\gamma-$Theorem)

There is a universal constant $u_0 \approx 0.0599...$ *such that if* $Df(x)$ *is onto and*

$$dist(x, V(f)) := \min_{z \in V(f)} dist(x, z) \leq \frac{u_0}{\gamma(f, x)},$$

then the sequence $x_k := N_f^k(x)$ *is well–defined and:*
The sequence $\{x_k\}$ *is a Cauchy sequence that satisfies*

$$dist(x_k, V(f)) \leq 2\left(\frac{1}{2}\right)^{2^k-1} dist(x, V(f)).$$

Moreover, the limit of $M_f(x) = \lim_{k\to\infty} N_f^k(x) \in V(f)$, *satisfies*

$$dist(x_k, M_f(x)) \leq 2\left(\frac{1}{2}\right)^{2^k-1} dist(x, V(f)).$$

The limit operator M_f, Dedieu, 00

A neighborhood $\mathcal{T}$ of $V_{\mathbb{A}}(f)$ given by:

$$\mathcal{T} := \{x \in \mathbb{K}^n \ : \ \exists \zeta \in V_{\mathbb{A}}(f), \ dist(x - \zeta) \leq \frac{u_0}{\gamma(f, x)}\}.$$

Theorem (Neighborhood of the Zero Set)

We have

1. *The map* $M_f : \mathcal{T} \longrightarrow V(f)$ *is continuous in* $\mathcal{T}$ *and differentiable in the interior of* $\mathcal{T}$*.*
2. *For every* $x \in \mathcal{T}$ *we have:*

$$dist(x, V(f)) \leq dist(x, M_f(x)) \leq 2dist(x, V(f)).$$

The mapping M_f "looks like" an (almost) "orthogonal" projection onto the variety.

[Dedieu–Shub, 2000b]: Over–determined

Least–Squares solutions. Ancestors in [Dennis–Schnabel, 83], [Seber–Wild, 89]
From $f := (f_1, \ldots, f_m) : \mathbb{C}^n \longrightarrow \mathbb{C}^m$, we have residue function

$$F(x) := \frac{1}{2}\sum_{i=1}^{m} \|f_i(x)\|^2.$$

Theorem

Invariant points of Newton's operator (i.e. $N_f(x) = x$) are exactly stationary points of the residue function F, i.e. those points x such that $DF(x) = 0$.

Definition (Least Square Solution)

A least square solution of $f : U \longrightarrow V$ is an invariant point of N_f.

[Dedieu–Shub, 2000b]: Over–determined II

$$\gamma_1(f,x) := \sup_{k\geq 2}\left(\|Df(x)^{\dagger}\|\left\|\frac{D^{(k)}f(x)}{k!}\right\|\right)^{\frac{1}{k-1}},$$
$$\beta_1(f,x) := \|Df(x)^{\dagger}\|\|f(x)\|,$$
$$\alpha_1(f,x) := \beta_1(f,x)\gamma_1(f,x).$$

Theorem ($\gamma-$Theorem in the Over–determined Case:solutions)

Let $\zeta \in E$ be such that $f(\zeta)=0$ and $Df(\zeta)$ is injective. Then, for every $z \in E$ such that

$$\|\zeta - z\| \leq \frac{3-\sqrt{7}}{2\gamma_1(f,\zeta)},$$

Newton's sequence $z_k := N_f^k(z)$ is well–defined and converges to ζ satisfying:

$$\|z_k - \zeta\| \leq \left(\frac{1}{2}\right)^{2^k-1}\|z-\zeta\|.$$

Theorem ($(\alpha, \gamma)-$Theorem in the Over–determined Case: LSS)

Let $\zeta \in E$ be a least-squares solution and $Df(\zeta)$ is injective. Then, for every $z \in E$ such that

$$\|\zeta - z\| \leq \frac{2-\sqrt{2}}{2\gamma_1(f,\zeta)},$$

If

$$\alpha_1(f,\zeta) \leq \frac{1}{2\sqrt{2}},$$

then Newton's sequence $z_k := N_f^k(z)$ is well–defined and there exists a constant $\lambda < 1$ (depending on α_1 and γ_1) such that:

$$\|z_k - \zeta\| \leq \lambda^k \|z - \zeta\|.$$

[Dedieu–Shub, 01]: On Simple Double Zeros

$f : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ analytic, $x \in \mathbb{C}^n$, a zero,
$\mathbb{C}\{X_1, \ldots, X_n\}_x :=$ local ring of germs of analytic functions,
$I_{f,x}$ the ideal generated by the coordinates of f.

$$mult(f, x) := \dim_{\mathbb{C}} \mathbb{C}\{X_1, \ldots, X_n\}_x / I_{f,x}.$$

"simple double zero"

- $mult(f, x) = 2$, $\dim ImDf(x) = n - 1$ (i.e. it is a singularity of co–rank 1).
- For all $v \in \mathbb{C}^n$, $\|v\| = 1$, then

$$D^2 f(x)(v, v) \notin ImDf(x).$$

f bad conditioned near x (according to Smale's α and γ Theories): $Df(x)$ is a singular matrix.

Remark

For higher multiplicities, continuations by J.C. Yakoubsohn, M. Giusti, G. Lecerf and B. Salvy.

[Dedieu–Shub, 01]: On Simple Double Zeros, The Quantities

They (slightly perturbed) define a linear operator

$$A(f, xv) = A(f, x) := Df(x) + \frac{1}{2}D^2 f(x)(v, \Pi_v),$$

where v is any unit vector in $kerDf(x)$, and Π_v is the orthogonal projection onto the subspace $\langle v \rangle$ spanned by v (Namely, $kerDf(x)$).

Definition (Generalizing γ to singular cases)

If $A(f, x)$ is invertible, define:

$$\gamma_2(f, x) := \max\left(1, \sup_{k\geq 2}\left\|A(f, x)^{-1}\frac{D^k f(x)}{k!}\right\|^{1/(k-1)}\right).$$

[Dedieu–Shub, 01]: On Simple Double Zeros, Separation and Quantitative Rouché's

* [Separation of Zeros, [Dedieu-Shub,01]: *If x is a simple double zero and $f(y) = 0$, then*

$$\|y - x\| \geq \frac{c}{\gamma_2(f, x)}.$$

Sup distance between holomorphic functions as:

$$d_R(f, g) := \sup_{\|y-x\| \leq R} \|f(y) - g(y)\|.$$

Theorem (Quantitative Rouché's Theorem near a simple double zero, Dedieu-Shub, 01)

Let x be a simple double zero of f and $0 < R \leq \frac{c}{2\gamma_2(f,x)}$. Then, if

$$d_R(f, g) < \frac{cR^2}{\|A(f, x)^{-1}\|},$$

then the sum of the multiplicities of the zeros of g in $B(x, R)$ is 2.

[Dedieu–Shub, 01]: On Simple Double Zeros, an answer to the problem

* $L : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ linear, s. t. $L\mid_{v^\perp} = 0$ and $L(v) = Df(x)v$.
* $B(f, x) = A(f, x) - L$.

$$\gamma_2(f, x, L) := \max\left(1, \sup_{k\geq 2}\left\|B(f, x, L)^{-1}\frac{D^k f(x)}{k!}\right\|^{1/(k-1)}\right).$$

Theorem (Dedieu–Shub, 01)

If

$$\|f(x)\| + \|Df(x)\|\frac{c}{2\gamma_2(f, x, L)} < \frac{c^3}{4\|B(f, x, L)^{-1}\|\gamma_2(f, x, L)^4},$$

then f has two zeros (counting multiplicities) in the ball around x of radius

$$\frac{c}{2\gamma_2(f, x, L)}.$$

Extending $\alpha-$ and $\gamma-$Theories to Riemannian Manifolds and Lie Groups.

Generalizing α and γ Theories to Riemann manifolds

A relevant ancestor is [Shub, 86]

- Human Spine modeling [Adler–Dedieu– Margulies, Martens, Shub]
- Intrinsic (Covariant) α and γ Theories for Riemann manidols [Dedieu–Proouret–Malajovich, 03]
- Computing exponential mapping for submanifolds of $\mathbb{R}^n$ [Dedieu–Nowicki, 05].
- Newton's method in Lie Groups [Dedieu–Li–Wang, 09].

Motivations I: math problems

Example (The eigenvalue problem)

The problem is to compute $\lambda \in \mathbb{C}$ *and* $x \in \mathbb{P}_{n-1}(\mathbb{C})$ *such that*

$$Ax = \lambda x.$$

Example (k–Invariant subspace computations)

The input is again a matrix $A \in \mathcal{M}_n(\mathbb{C})$ *and the goal is to compute a subspace* $V \in \mathbb{G}_{n,k}(\mathbb{C})$ *(Grassmannian of* $k-$*dimensional subspaces of* $\mathbb{C}^n$*) such that* $AV \subseteq V$.

Remark (Other examples:)

Symmetric eigenvalue problem, minimization problems with orthogonality constraints, optimization problems with equality constraints.

Motivations II: Human Spine and Scoliosis

Scoliosis(from Greek: skoliosis meaning "crooked")is a medical condition in which a person's spine is curved from side to side, shaped like an "s", and may also be rotated [Wiki *dixit*].

[Adler–Dedieu–Margulis–Martens–Shub]
They do not include cervical vertebrae and identify Sacral vertebrae as a single one. This yields a geometric model for the human spine with 18 vertebrae.

The *position* of each vertebra is given by an $SO(3)$ orthogonal matrix m_i (which defines a frame of orthogonal vectors).

After some reductions, the total alignment discrepancies between spinal elements are determined by a function $\phi(m_2, \ldots, m_{17})$.

Thus, the problem of the relative position of a Human Spine becomes an optimization problem of the kind:

Problem (Perfect alignment of Human Spine)

Given $(m_2, \ldots, m_{17}) \in SO(3)^{16}$ *subject to some additional constraints* $h_1 = 0, h_2 = 0$, *find a minimum of a quadratic polynomial function* $\phi : SO(3)^{16} \longrightarrow \mathbb{R}_+$.

Newton's operator: the case of mappings

M is a manifold of dimension n and $f : M \longrightarrow \mathbb{R}^m$ a differentiable mapping.

Definition (Newton's operator)

Newton's operator is a mapping $N_f : M \longrightarrow M$, given by

$$N_f(x) := exp_x\left(-Df(x)^{\dagger} f(x)\right).$$

As in the QNAG over–determined case, these authors introduce the residue function $F : M \longrightarrow \mathbb{R}_+$ given by

$$F(x) := \frac{1}{2}\|f(x)\|^2.$$

Definition (Least Square Solution)

A a least square solution of $f : M \longrightarrow \mathbb{R}^m$ is a point $x \in M$ such that the residue function F satisfies $DF(x) = 0$.

Newton's operator: the case of mappings

Some of the main contributions in this work are the following ones:

Proposition (Adler–Dedieu–Margulies–Martens–Shub, 02)

Fixed points of Newton's operator N_f correspond to least square solutions of f.

Proposition (Adler–Dedieu–Margulies–Martens–Shub, 02)

- *Attractive fixed points of N_f are strict local minima of the residue function F.*
- *Local maxima of the residue function F are repelling points for N_f.*
- *When $dim(M) = m$, and $Df(x)$ is invertible, then the fixed points of N_f are indeed zeros of f and the convergence is quadratic.*

As for the under–determined case (i.e. $m < dim(M)$) they obtained similar results to those already discussed in the Numerical Algebraic Geometry case.

Newton's operator: the case of vector fields

Again, they took the stream of [Shub, 86]: M be a $n-$dimensional Riemannian real geodesically complete analytic manifold.
For every vector field $X \in \mathfrak{X}(M)$ *(i.e.* $X : M \longrightarrow TM$*), they also consider the problem of computing the zeros* $x \in M$ *such that* $X(x) = 0_x \in T_xM$*.*
Newton's operator is given by:

$$N_X(x) := exp_x(-DX(x)^{-1}X(x)),$$

where $DX(x)$ is a notation to represent the following mapping:
Levi–Civita connection ∇ defines a linear endomorphism
$\nabla_{-}X(x) : T_xM \longrightarrow T_xM$

$$\nabla_{-}X(x)(\omega) = (\nabla_Y X)(x) \in T_xM,$$

where $Y \in \mathfrak{X}(M)$ is any vector field such that $Y(x) = \omega$.

$$DX(x) := \nabla_{-}X(x)$$

and by $DX(x)^{-1}$ its inverse, provided that it exists.

They obtained similar results for vector fields

Proposition (Adler–Dedieu *et al.*, 02)

With these notations, if $x \in M$ *is a fixed point for* N_X*, then* $X(x) = 0$ *and* $DN_X(x) = 0$.

The Final Outcomes

In [Adler–Dedieu *et al.*, 02] they did a terrific work on these ideas to produce explicit formulae for Newton's method associated to their geometric model of Human Spine.

Intrinsic Quantitative Numerics in Riemannian Geometry

From [Dedieu–Priouret–Malajovich, 03].
M is an analytic complete Riemannian manifold of dimension n.
The idea was to develop α and γ Theories for computing zeros both for analytic mappings

$$f : M \longrightarrow \mathbb{R}^m$$

and analytic vector fields $X \in \mathfrak{X}(M)$:

$$X : M \longrightarrow TM$$

α and γ quantities, analytic mappings

Definition (α and γ quantities for mappings)

Let $f : M \longrightarrow \mathbb{R}^n$ *be an analytic mapping. Define the following quantities:*

$$\gamma(f, x) := \sup_{k \geq 2} \| \frac{Df(x)^{-1} D^k f(x)}{k!} \|^{1/k-1},$$

and

$$\beta(f, x) := \|Df(x)^{-1} f(x)\|.$$

Finally, define α *quantity by:*

$$\alpha(f, x) := \beta(f, x)\gamma(f, x).$$

α and γ quantities, analytic vector fields

Multilinear mapping

$$D^k X(x) : (T_x M)^k \longrightarrow T_x M,$$

given by:

$$D^k X(x)(u_1, \ldots, u_k) := \nabla_{X_1, \ldots, X_k} X(x) \in T_x M,$$

where for every i, $1 \leq i \leq k$, X_i is any vector field such that $X_i(x) = u_i$.

Definition (α and γ quantities for vector fields)

Let $X \in \mathfrak{X}(M)$ be an analytic vector field. We define the following quantities:

$$\gamma(X, x) := \sup_{k \geq 2} \| \frac{DX(x)^{-1} D^k X(x)}{k!} \|^{1/k-1},$$

and

$$\beta(X, x) := \|DX(x)^{-1} X(x)\|.$$

Finally, we define the α quantity by:

$$\alpha(X, x) := \beta(X, x)\gamma(X, x).$$

α and γ quantities, analytic vector fields

Definition

Let (M, g) *be a Riemannian manifold and* $x \in M$. *We define the quantity:*

$$K_x := \sup \left\{ \frac{d(exp_z(u), exp_z(v))}{\|u - v\|_z} \ : \ z \in B_M(x, r_x), \ u, v \in T_z M \right\}.$$

Proposition

With the same notations, $K_x \geq 1$ *and, if* M *has non–negative sectional curvature,* $K_z = 1$.

The authors use parallel transport to state a Taylor's formula and a radius of convergence of Taylor's series both for analytic mappings and vector fields. The radius of convergence is respectively given by

$$\frac{1}{\gamma(f, x)}, \frac{1}{\gamma(X, x)}.$$

γ Theorem, analytic mappings

Theorem ($\gamma-$Theorem, Dedieu–Priouret–Malajovich, 03)

Let $f : M \longrightarrow \mathbb{R}^n$ *be an analytic mapping. Let* $\zeta \in M$ *be such that* $f(\zeta) = 0$ *and* $Df(\zeta) \in GL(n, \mathbb{R})$. *Let us define the quantity:*

$$R(f,\zeta) := \min\left\{ r_\zeta, \frac{K_\zeta + 2 - \sqrt{K_\zeta^2 + 4K_\zeta + 2}}{2\gamma(f,\zeta)} \right\}.$$

Then, for every $z \in M$, *if* $d(z,\zeta) \leq R(f,\zeta)$, *then,*

- *The sequence* $z_k := N_f^k(z)$ *is well–defined and*
- *for each* $k \geq 0$

$$d(z_k,\zeta) \leq \left(\frac{1}{2}\right)^{2^k-1} d(z,\zeta).$$

In the case M has non–negative sectional curvature

$$R(f, zeta) := \min\left\{ r_\zeta, \frac{3-\sqrt{7}}{2\gamma(f,\zeta)} \right\}.$$

α Theorem, analytic mappings

Theorem ($\alpha-$Theorem, Dedieu–Priouret–Malajovic, 03)

With the same notations as above, there are two universal quantities:

$$s_0 := 0.103621842...$$

$$\alpha_0 := 0.130716944...$$

such that the following holds:
If $\beta(f,z) := \|Df(z)^{-1}f(z)\|_z \leq s_0 r_z$ and $\alpha(f,z) = \beta(f,z)\gamma(f,z) \leq \alpha_0$, then z is an approximate zero of f with some associated zero $\zeta \in M$.

γ Theorem, analytic vector fields

Theorem ($\gamma-$Theorem, Dedieu–Priouret–Malajovich, 03)

Let $X \in \mathfrak{X}(M)$ *be an analytic vector field. Let* $\zeta \in M$ *be such that* $X(\zeta) = 0$ *and* $DX(\zeta)$ *is an isomorphism. Let us define the quantity:*

$$R(X,\zeta) := \min\left\{ r_\zeta, \frac{K_\zeta + 2 - \sqrt{K_\zeta^2 + 4K_\zeta + 2}}{2\gamma(X,\zeta)} \right\}.$$

Then, for every $z \in M$*, if* $d(z,\zeta) \leq R(X,\zeta)$*, then*

- *The sequence* $z_k := N_X^k(z)$ *is well–defined and*
- *for each* $k \geq 0$

$$d(z_k,\zeta) \leq \left(\frac{1}{2}\right)^{2^k-1} d(z,\zeta).$$

γ Theorem, analytic vector fields

Theorem ($\alpha-$Theorem, Dedieu–Priouret–Malajovich, 03)

With the same notations as above, there are two universal quantities:

$$s_0 := 0.103621842...$$

$$\alpha_0 := 0.130716944...$$

such that the following holds:
If $\beta(X, x) := \|DX(x)^{-1}X(x)\|_x \leq s_0 r_x$ *and* $\alpha(X, x) = \beta(X, x)\gamma(X, x) \leq \alpha_0$, *then* z *is an approximate zero of* X *with some associated zero* $\zeta \in M$.

Evaluating the exponential function, Dedieu–Nowicki, 05

Newton's method in Riemannian manifolds requires efficient evaluation of the exponential function

Evaluating the exponential function, Dedieu–Nowicki, 05

Evaluating Newton's method in either case (vector fields or mappings) requires evaluating the exponential mapping:

$$N_X(x) := exp_x(-DX(x)^{-1}X(x)), \quad N_f(x) := exp_x(-Df(x)^{\dagger}f(x))$$

and this amounts to evaluate geodesics $x(1)$ which given in local coordinates by the usual equations:

$$\ddot{x}_i = \sum_{k,j} \Gamma^i_{k,j}\dot{x}_k\dot{x}_j,$$

In [Dedieu–Nowicki, 95] they show how to circumvent local charts.

Evaluating the exponential function, Lagrangian

$F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ a smooth mapping such that 0 is a regular value. Let $M := F^{-1}(0) \subseteq \mathbb{R}^n$ be the smooth Riemannian sub-manifold of dimension $n - m$ of $\mathbb{R}^n$.
They introduce a Lagrange multiplier $\lambda(t)$ and observe that geodesics are the solutions of some Euler–Lagrange equation

$$\begin{cases} F(x(t)) = 0 \\ \ddot{x}(t) = -DF(x)^* \lambda(t) \\ x(0) = x, \quad \dot{x}(0) = u \end{cases}$$

associated to the following Lagrangian:

$$\mathcal{L}(x, t, \dot{x}) := \frac{1}{2} \|\dot{x}\|^2 - \sum_{i=1}^{m} \lambda_i F_i(x).$$

Evaluating the exponential function, Lagrangian

Then, they proceed as in classical mechanics, introducing two new groups of variables $p \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^m$ and a Hamiltonian:

$$\mathcal{H}(x, p, \mu) := \langle p, \dot{x} \rangle - L(x, \dot{x}, t) = \langle p, \dot{x} \rangle - \frac{1}{2}\|\dot{x}\|^2 + \sum_{i=1}^{m} \mu_i DF_i(x)\dot{x},$$

where

$$p = \dot{x} - DF(x)^* \mu = \frac{\partial L}{\partial \dot{x}}.$$

According to Pontryagin's principle, they obtain the corresponding Hamiltonian equations:

$$\begin{cases} \dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial x}(x(t), p(t), \mu(t)), \\ p(t) = \dot{x}(t) - DF(x(t))^* \mu(t) \\ \dot{\mu}(t) = -\lambda(t), \quad \mu(0) = 0 \end{cases}$$

Evaluating the exponential function, apply symplectic methods

Theorem

Geodesics may be rewritten as the solutions of:

$$\begin{cases} \dot{p}(t) = -\sum_{i=1}^{m} \mu_i D^2 F_i(x)\dot{x}(t), \\ \dot{x}(t) = \left(Id - DF(x(t))^{\dagger} DF(x(t))\right) p(t), \\ \mu(t) = -\left(DF(x(t))^{*}\right)^{\dagger} p(t), \\ x(0) = x, \ \ \dot{x}(0) = u \end{cases}$$

Finally, they apply symplectic Runge–Kutta methods as in [Hairer, 03] or [Sanz–Serna–Calvo, 94] to make a backwards error analysis, they implemented their procedure and gave an estimate for the "computational complexity" in numerical terms.

In practical implementations, it seems to run fast.

This topic has been discussed again in [Boito–Dedieu, 09] for geodesics of the general linear group with respect to the condition metric.

The Case of Lie Groups

Another α and γ Theories, specific for Lie groups

After the works in [Dedieu–Priouret–Malajovich], several authors discussed quantitative aspects on Newton's methods in Riemannian manifolds and Lie groups.

In cite[Owren–Welfert, 00] discussed Newton's method on Lie groups.

In [Li–Wang, 06], the authors developed a completely different $\gamma-$condition on Riemannian manifolds, whereas in [Alvarez–Bolte–Munier, 08] the authors introduced a unifying criterion for the convergence of Newton's method.

In [Dedieu–Li–Wang, 09], they introduced a new γ Theory for Lie Groups which improves the one in [Li–Wang, 06] because it does not depend on the curvature of the manifold.

Basic notations

G a Lie group, $\mathcal{G} = T_e G$ its Lie algebra.

$L_g : G \longrightarrow G$ the left translation defined by g For every $u \in \mathcal{G}$, let

$X_u : G \longrightarrow TG$ be the vector field $X_u : G \longrightarrow TG$ given by:

$$g \longmapsto T_e L_g(u) \in T_g G = g\mathcal{G}.$$

Let $\sigma_u : \mathbb{R} \longrightarrow G$ be the one–parameter subgroup given by :

$$\begin{cases} \dot{\sigma}_u(0) = u, \;\; \sigma_u(0) = e \\ \dot{\sigma}_u = X_u(\sigma_u) = T_e L_{\sigma_u(t)}(u). \end{cases}$$

Then the exponential mapping $exp_e : \mathcal{G} \longrightarrow G$ is given by

$$exp_e(u) := \sigma_u(1).$$

Piece–wise one–parameter subgroup

Definition (Piece–wise one–parameter subgroup)

Given $x, y \in G$*, a piece–wise one–parameter subgroup connecting* x *and* y *is a mapping* $c : [0, m+1] \longrightarrow G$ *such that for every* i*,* $0 \leq i \leq m$*, there exists* $u_i \in \mathcal{G}$*, such that for every* $t \in [i, i+1]$ *we have:*

$$c(t) := c(i) exp_e((t-i)u_i).$$

Note that the length of a one parameter subgroup c is given by

$$\ell(c) := \int_0^{m+1} \|\dot{c}(t)\| dt = \sum_{i=1}^{m} \|u_i\|,$$

and

$$d(x, y) \leq \sum_{i=1}^{m} \|u_i\|.$$

Let $f : G \longrightarrow \mathcal{G}$ be a smooth mapping. They define

$$df_x : \mathcal{G} \longrightarrow \mathcal{G}$$

by

$$df_x := T_x f \circ T_e L_x.$$

Definition (Newton's operator)

Given $f : G \longrightarrow \mathcal{G}$, we define Newton's operator by

$$N_f(x) := x \cdot exp_e(-df_x^{-1} f(x)).$$

Pieces one–parameter subgroup γ−condition

Definition

With the same notations as above, let $f : G \longrightarrow \mathcal{G}$ be a smooth mapping and $\gamma > 0$, $r > 0$ two positive real numbers such that $\gamma r \leq 1$. Let $x_0 \in G$ be a point such that $df_{x_0}^{-1}$ exists.
We say that f satisfies the pieces one–parameter subgroup γ−condition at x_0 in $B_G(x_0, r)$ if for every $x \in B_G(x_0, r)$ and for every piece–wise one–parameter subgroup c connecting x_0 and x with arc-length at most r the following holds:

$$\|df_{x_0}^{-1} d^2 f_x\| \leq \frac{2\gamma}{(1 - \gamma\ell(c))^3}$$

Then, they exhibited generalized α and γ Theories for smooth mappings.

The Analytic Case, pre-amble

They define the "usual" quantities:

$$\gamma(f,x) := \sup_{k\geq 2} \left\| \frac{df_x^{-1} d^k f_x}{k!} \right\|^{1/(k-1)}, \quad \beta(f,x) := \|df_x^{-1} f(x)\|,$$

and

$$\alpha(f,x) := \beta(f,x)\gamma(f,x).$$

Proposition

For analytic mappings $f : G \longrightarrow \mathcal{G}$, taking $\gamma := \gamma(f,x)$ and $r := \frac{2-\sqrt{2}}{2\gamma}$, then f satisfies pieces γ-condition at x on $B_G(x,r)$.

Some Universal Constants

Some universal constants: Let $\psi(s)$ be the univariate polynomial

$$\psi(s) := 1 - 4s + 2s^2.$$

Let $a_0 \approx 0.808..$ be the smallest positive root of the equation:

$$\frac{t}{\psi(t)} = 3 - 2\sqrt{2}.$$

Let s_0 be the real number

$$s_0 := \frac{2\psi(a_0)}{(2+\sqrt{2})(1-a_0) + 2\psi(a_0)}.$$

The Analytic Case, $\alpha-$Theorem

Theorem ($\alpha-$Theorem)

Assume that

$$\alpha(f,x) \leq \frac{13 - 3\sqrt{17}}{4}.$$

Then Newton's method with initial point x *is well–defined and converges to some real zero* ζ *of* f *in* $\overline{B_G(x, r_1(\alpha))}$. *Moreover,*

$$d_G\left(N_f^{n+1}(x), N_f^n(x)\right) \leq \left(\frac{1}{2}\right)^{2^n - 1} \beta(f,x).$$

the constant K_z is no more in the statement.

Theorem ($\gamma-$Theorem)

Let $\zeta \in G$ such that $f(\zeta) = 0$ and df_ζ is invertible.
Let $\rho > 0$ be the largest real number such that

$$B(e, \rho) \subseteq exp\left(B_{\mathcal{G}}(0, \frac{2-\sqrt{2}}{2\gamma(f,\zeta)}) \right),$$

and let $r(f,\zeta) \in \mathbb{R}$ be the positive real number given by

$$r(f,\zeta) := \min\left\{ \frac{a_0}{\gamma(f,\zeta)}, s_0\rho \right\}.$$

Then, for every point $x \in G$ such that

$$x \in N(\zeta, r) := \zeta \cdot exp\left(B_{\mathcal{G}}(0, r) \right).$$

Newton's sequence $N_f^k(x)$ with initial point x is well–defined and converges quadratically to ζ.

Linear Programming, Newton Flow and Curvature

Dealing with central paths in Linear Optimization

- Newton Flow Interior Point Methods in LP [Dedieu–Shub, 2005]
- Curvature of the Central Path [Dedieu-Malajovich–Shub, 2005]
- On the number of local Minima [Dedieu-Malajovich,2009]

Newton Flow Interior Point M. in LP

[Dedieu–Shub, 05] *"...the aim is to give a global picture of the central paths even for degenerate problems as solution curves of the Newton vector field of the logarithmic barrier function..."*

Problem (Primal Problem)

Minimize:

$$\min_{Ax \geq b} \langle c, x \rangle,$$

where $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, $c \in \mathbb{R}^n$ *and* $b \in \mathbb{R}^m$ *are the instances of the problem.*

Problem (Dual Problem)

Maximize

$$\max_{A^T y = c, y \geq 0} \langle b, y \rangle,$$

and same instances.

Fast Definit. of the central path

Let A_i be the $i-$th row of the matrix A, and define the barrier function:

$$f(x) := \sum_{i=1}^{m} \ln(A_i x - b_i).$$

Define the objective function ($t > 0$) as

$$f_t(x) := t\langle c, x\rangle - f(x).$$

A family of linear optimization problems $LP(t)$ given by:

$$\min_{x\in\mathbb{R}^n} f_t(x).$$

Problem $LP(t)$ has a unique solution $c(t)$ and the curve

$$c : (0, \infty) \longrightarrow \mathcal{P}$$

is called the central path of the primal problem.

A Conjecture

Conjecture

The curvature of the central path is linearly bounded by the dimension n of the polytope.

In their own words: *"Our point in studying the total curvature is that curves with small total curvature may be easy to approximate with straight lines. So, small total curvature may contribute to the understanding of why long step interior point methods are seen to be efficient in practice".*

Primal/Dual Central Paths, [Dedieu–Malajovich–Shub, 05]

Primal:

$$\min_{Ax-s=b,\ s\geq 0} \langle c, x\rangle,$$

Dual:

$$\max_{A^T y=c, y\geq 0} \langle b, y\rangle,$$

Primal/Dual central Path :the curve $(x(\mu), s(\mu), y(\mu))$, $\mu \in (0, \infty)$ satisfying:

$$\begin{cases} Ax - s = b, \\ A^T y = c, \\ sy = \mu(1, \ldots, 1), \\ y > 0, \quad s > 0 \end{cases}$$

- The curve $(x(\mu), s(\mu))$ is the central path of the primal problem and minimizes $-\mu \sum_{i=1}^{m} \ln(s_j) + \langle c, x\rangle$.
- The curve $y(\mu)$ is the central path of the dual problem and maximizes $\mu \sum_{i=1}^{m} \ln(y_i) + \langle b, y\rangle$ restricted to the dual polytope.

Total curvature, [Dedieu–Malajovich–Shub, 05]

For every central paths $c(\mu)\mathbb{R}^N$, let $\dot{c}(s) \in S^{N-1}$ be its natural, arc-length or unit speed parametrization (also Gauss curve). Let $\kappa(s)$ and K be resp. its curvature and total curvature

$$\kappa(s) := \frac{d\dot{c}(s)}{d\ell}(s), \quad K := \int_0^L \|\kappa(s)\| ds.$$

Theorem (A Poincaré formula, Dedieu–Malajovich–Shub, 05)

$$L(\gamma) := \int_a^b \|\dot{\gamma}(t)\| \, dt = \int_{G_{n,n-1}} \sharp(H \cap \gamma) \, dG(H).$$

Moreover, if there is some constant $\mathcal{B}$ *such that* $\sharp(H \cap \gamma) \leq \mathcal{B}$ *we then conclude:*

$$L(\gamma) \leq \pi \mathcal{B}.$$

Multi–Homogeneous degree bound, [Dedieu–Malajovich–Shub, 05]

They proved that the Gauss curve of Primal/Dual central paths may be described as the solution of a system of multi–homogeneous equations:

- $m - n$ equations of multi–degree $(1, 0)$,
- $m - 1$ equations of multi–degree $(1, 1)$, and
- an additional equation of multi–degree $(1, 2n + 1)$.

Similar equations arise in the case of the primal (the last equation has multi–degree $(0, 2n - 2)$) and dual (the last equation with multi–degree $(0, 2n + 1)$) cases.

Total Curvature Upper Bound, [Dedieu–Malajovich–Shub, 05]

Apply then the multi–homogeneous degree bound to Poincaré's Formula

Theorem

Let $m > n \geq 1$*. Let* A *be an* $m \times n$ *matrix,* $b \in \mathbb{R}^m$ *and* $c \in \mathbb{R}^n$*, satisfying:*

$$b \notin Im(A), \; c \neq 0.$$

The sum over all 2^m *sign conditions*[a]

$$s_i \varepsilon_i 0, \; \varepsilon_i \in \{\geq, \leq\}.$$

of the total curvature of the primal/dual central path is at most:

$$2\pi n \binom{m-1}{n}.$$

[a] $\min_{Ax-s=b, \; s\varepsilon 0} \langle c, x \rangle$,

And..

Theorem

In the case of the primal central path, this bound becomes:

$$2\pi(n-1)\binom{m-1}{n}.$$

In the case of the dual

$$2\pi n\binom{m-1}{n}.$$

Averaging over all poly–topes, [Dedieu–Malajovich–Shub, 05]

Theorem

With the same notations as above, the average total curvature is at most:

1. *In the Primal/Dual case* $\leq 2\pi n$,
2. *in the Primal case* $\leq 2\pi(n-1)$, *and*
3. *in the Dual case* $\leq 2\pi n$.

Then answering their original conjecture.

Lower Complexity Bounds for Homotopy Methods

Lower Complexity Bounds I

Some works (as the lower bound for the separation of roots in[Dedieu, 97]) may be used to show lower complexity bounds...

I like the main outcome in [Dedieu-Smale, 98] on lower complexity bounds for homotopy (path following) methods based on the $\alpha-$Theorem.

* $f_t : \mathbb{C}^{n+1} \longrightarrow \mathbb{C}^m$, $t \in [0, 1]$, is a path of equations.
* $g = f_0$, $f = f_1$.
* Newton Continuation Method Sequence is a sequence of real numbers

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

and a sequence of equations/solutions

$$(f_i, \zeta_i) = (f_{i_1}, \zeta_i), \;\; 0 \leq i \leq k,$$

satisfying α condition:

$$f_i(\zeta_i) = 0, \;\; \alpha(f_{i+1}, \zeta_i) \leq \alpha_0 \text{ with assoc. zero } \zeta_{i+1}.$$

Lower Complexity Bounds, II

* The homotopy polygonal:

$$(f_i, z_i), \text{ with } z_0 = \zeta_0,\ z_{i+1} := N_{f_{i+1}}(z_i).$$

with initial data (f_0, ζ_0) yields an approximate zero of f with associated zero ζ_k.

Theorem (Dedieu-Smale, 98)

If $(f_i, \zeta_i), 1 \leq i \leq k$ is a Newton Continuation Method Sequence, then

$$k \geq c \max\left\{1, \frac{D-1}{2}\right\} dist(\zeta_0, \zeta_k).$$

Moreover, assume that $d_R(\zeta_i, \Sigma_{f_i}) \leq \varepsilon$, $0 \leq i \leq k$, where Σ_{f_i} is the set of points $x \in \mathbb{C}^{n+1}$ such that rank $Df(x)$ *is not maximal. Then,*

$$k \geq c\varepsilon^{-1} dist(\zeta_0, \zeta_k).$$

The future

Jean–Pierre is maintaining rich activities nowadays (works with Armentano, Beltrán, Boito, Malajovich, Shub...some of them in this conference). I'm sure that Jean–Pierre will continue providing good challenges and good papers to work with...

Joyeux Anniversaire, Jean–Pierre!!