# SELF-CONVEXITY

GREGORIO MALAJOVICH

Numerical analysis has a rich story of interplay between two definitions of the condition number: analytic and geometric. The analytic condition number is the norm of the derivative of the solution with respect to the coefficients. The geometric condition number is the reciprocal distance to the set of ill-posed instances. For instance, the Eckart-Young theorem says that for affine equation solving or least squares, those definitions are the same.

Self-convexity is an attempt to encompass some of the convex geometric properties of condition numbers in general. It was motivated by an attack to Smale's 17-th problem, which deals with systems of polynomial equations. However, the results obtained so far are still limited to affine systems.

The objective of this talk is to present the basic ideas, results and motivations. Much of this subject is still wide open, so there may be more questions than theorems.

## 1. What is self-convexity

Through this talk, $(\mathcal{M}, \langle\cdot,\cdot\rangle_x)$ is always a **smooth** Riemannian manifold and $\alpha : \mathcal{M} \to \mathbb{R}_{>0}$ is a **Lipschitz** function.

We can endow the manifold $M$ with a new metric, namely

$$\langle\cdot,\cdot\rangle'_x = \alpha(x)\langle\cdot,\cdot\rangle_x$$

which is conformally equivalent to the previous one. This new norm will be called the **$\alpha$-metric** and sometimes the **condition metric**. It defines a **Lipschitz Riemannian** structure on $\mathcal{M}$.

**Definition 1.1.** We say that $\alpha$ is **self-convex** if and only if, for any **geodesic** $\gamma$ in the $\alpha$-structure, $t \mapsto \log \alpha(\gamma(t))$ is a convex function.

This definition makes sense when $\alpha$ is of class $\mathcal{C}^1$ so that the geodesic differential equation has a solution. When $\alpha$ is merely Lipschitz, a **geodesic** is a locally minimizing absolutely continuous ($C^{1+\mathrm{Lip}} = W^{2,\infty}$) path, parametrized by arc length a.e. (For a discussion, see Boito and Dedieu (2010) and Beltrán, Dedieu, Malajovich, and Shub (2012)).

## 2. Known examples of self-convexity

**Theorem 2.1.** *Let $C \subset \mathbb{R}^n$ a (closed) convex body. Let $\mathcal{M} = (\mathbb{R}^n \setminus C)$ and let $\alpha : \mathbf{x} \mapsto d(\mathbf{x}, C)^{-2}$. Then $\alpha$ is self-convex.*

**Theorem 2.2** (Beltrán et al. (2010), Th.2)**.** *Let $N \subset \mathbb{R}^m$ be an embedded submanifold (without border of course). Let $\mathcal{M}$ be the largest open set in $\mathbb{R}^m \setminus N$ such that every point of $\mathcal{M}$ has a unique closest point in $N$. Let $\alpha : \mathbf{x} \mapsto d(\mathbf{x}, \mathcal{N})^{-2}$. Then $\alpha$ is self-convex.*

Theorem 2.1 follows immediately from Li and Nirenberg (2005) and the result above.

**Theorem 2.3** (Beltrán et al. (2012), Th.1)**.** *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $L(m,n) = \mathbb{K}^{m\times n}$ where we assume that $m \geq n$, endowed with the trace inner product, and let $\mathcal{M} = L(m,n) \setminus \{A : \mathrm{Rank}(A) < n\}$. Let $\alpha : A \mapsto \|(A^*A)^{-1}\|$. Then $\alpha$ is self-convex.*

More examples are known, and also some counterexamples (Beltrán et al., 2010; Beltrán et al., 2012).

Since proofs can get extremely technical, I will not attempt to sketch any argument. Instead, I intend to explain in the rest of the talk why are we investigating such issues.

Our main motivation is Smale's 17-th problem. This is a long story, that started with the **Bézout saga** (Shub and Smale, 1993a; 1993b; 1993c; 1996; 1994; Shub, 2009; Beltrán and Shub, 2009).

## 3. The algebra

Let $\mathcal{H}_d$ be the space of complex homogeneous polynomials of degree $d$, in $n$ variables. There are many standard ways to represent polynomials, here are two:

$$F(\mathbf{X}) = \sum_{\sum \mathbf{a}_i = d} F_{\mathbf{a}} X_1^{a_1} X_2^{a_2} \cdots X_n^{a_n} = \sum_{0 \leq j_1,\dots,j_d \leq n} S_{\mathbf{j}} X_{j_1} X_{j_2} \dots X_{j_n}.$$

In the last representation, we assume that the $S_{\mathbf{j}}$ are coefficients of a symmetric $d$-contravariant tensor $S(\mathbf{X}, \mathbf{Y}, \dots, \mathbf{Z})$ and $F(\mathbf{X}) = S(\mathbf{X}, \dots, \mathbf{X})$.

The quantity

$$\|\mathbf{S}\|^2 = \sum_{0 \leq j_1,\dots,j_d \leq n} |S_{\mathbf{j}}|^2$$

is invariant by unitary rotations. This is actually an exercise in my book (Malajovich, 2011).

The corresponding norm in the polynomial representation

$$\|\mathbf{F}\|^2 = \sum_{\sum a_i = d} |F_{\mathbf{a}}|^2 \frac{a_1! \dots a_n!}{d!}$$

is known as the **Weyl** norm or sometimes **Bombieri** norm. Both come with an inner product. This is the inner product $\mathcal{H}_d$ is endowed with.

Moreover, $\mathcal{H}_d$ is a **reproducing kernel space**. If we set $K_d(\mathbf{X}, \mathbf{Y}) = \langle \mathbf{X}, \mathbf{Y}\rangle^d$ then

$$F(\mathbf{Y}) = \langle F(\cdot), K_d(\cdot, \mathbf{Y})\rangle.$$

---

*Date:* July 12, 2013

Talk given at the SIAM Annual Meeting, San Diego CA, 2013.

This is joint work with Carlos Beltrán, Jean-Pierre Dedieu and Mike Shub.

The author was partially supported by CNPq and CAPES (Brazil), and by MathAmSud grant *Complexity.* Part of this work was done during while visiting the Fields institute, Toronto, Canada during the Thematic Program in the Foundations of Computational Mathematics, and while visiting the Institut de Matemàtica de la Universitat de Barcelona.

If $\mathbf{d} = (d_1, \dots, d_n)$, the space of systems of polynomials

$$\mathcal{H}_{\mathbf{d}} = \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$$

is also endowed with the unitarily invariant, product space inner product.

## 4. The algebraic geometry

The *solution variety* is the set of pairs (problem, solution). Formally,

$$\mathcal{V} = \{(\mathbf{f}, \mathbf{x}) \in \mathbb{P}(\mathcal{H}_{\mathbf{d}}) \times \mathbb{P}^n : \mathbf{f}(\mathbf{x}) = 0\}.$$

This compactification is not **always** necessary, but it is extremely convenient. Through this talk I follow the convention that vectors are upper case ($\mathbf{X}$) and the corresponding projective points are lower case ($\mathbf{x}$).

Let $\mathrm{ev}(\mathbf{F}, \mathbf{X})$ denote the evaluation of $\mathbf{F}$ at $\mathbf{X}$,

$$\mathrm{ev}(\mathbf{F}, \mathbf{X}) = \begin{bmatrix} F_1(\mathbf{X}) \\ \vdots \\ F_n(\mathbf{X}) \end{bmatrix} = \begin{bmatrix} \langle F_1(\cdot), K_{d_1}(\cdot, \mathbf{X})\rangle \\ \vdots \\ \langle F_n(\cdot), K_{d_n}(\cdot, \mathbf{X})\rangle \end{bmatrix}$$

The $i$-th coordinate of the evaluation function is a polynomial in $F \in \mathcal{H}_{d_i}$ and $\mathbf{X} \in \mathbb{C}^n$, and it is an easy exercise to show that $D\mathrm{ev}(\mathbf{F}, \mathbf{X})$ is surjective. Thus $\mathcal{V}$ is a smooth algebraic variety. Consider now the two canonical projections

$$\pi_1 : \mathcal{V} \to \mathbb{P}(\mathcal{H}_{\mathbf{d}}) \qquad \text{and} \qquad \pi_2 : \mathcal{V} \to \mathbb{P}^n$$

Let $\Sigma$ be the set of critical values of $\pi_1$. It follows from Sard's theorem that $\Sigma$ has measure zero, and from elimination theory that $\Sigma$ is an algebraic set.

Moreover, $\pi_1$ is onto.

Therefore, for **generic** $\mathbf{F}_0$ and $\mathbf{F}_1$, the complex line

$$(1-t)\mathbf{F}_0 + t\mathbf{F}_1$$

cuts $\Sigma$ in finitely many (complex) values of $t$. Therefore if we require $t \in [0,1]$, the event of $(\mathbf{F}_t)_{t\in[0,1]}$ hitting $\Sigma$ has probability zero. Therefore the lifting theorem applies and can be used to solve polynomial systems. This is where the Bézout saga begins.

## 5. The calculus

Assume that $(\mathbf{F}_0, \mathbf{X}_0) \in \mathcal{V}$, $\mathbf{F}_0 \notin \Sigma$. Then we are under the hypotheses of the implicit function theorem: there are $\delta > 0$ and a function $G : B(f_0, \delta) \to \mathbb{P}^n$ such that

$$\begin{aligned} \mathrm{ev}(\mathbf{F}, G(\mathbf{F})) &\equiv 0 \\ G(\mathbf{F}_0) &= \mathbf{X}_0 \end{aligned}$$

In order to design path-following algorithms, it is important to give bounds for $\delta$. In the early Bézout saga, this was ultimately done in terms of condition numbers.

There are two current definitions of the condition number. The **unnormalized condition number** measures the sensitivity of the (projectivized) solution $\mathbf{x}$ to the (projectivized) input $\mathbf{f}$. It is defined as $\|DG(\mathbf{f}, \mathbf{x})\|$, where the operator norm of $DG(\mathbf{f}, \mathbf{x}) : T_f\mathbb{P}(\mathcal{H}_{\mathbf{d}}) \to T_x\mathbb{P}^n$ is assumed.

**Lemma 5.1.** *In the context above, let $F \in \mathcal{H}_{\mathbf{d}}$, $X \in \mathbb{C}^{n+1}$ be representatives of $(\mathbf{f}, \mathbf{x}) \in \mathcal{V}$.*

$$(1) \qquad \|DG(\mathbf{f}, \mathbf{x})\| = \|\mathbf{F}\| \left\| \left( \begin{bmatrix} \|\mathbf{X}\|^{-d_1+1} & & \\ & \ddots & \\ & & \|\mathbf{X}\|^{-d_n+1} \end{bmatrix} D\mathbf{F}(\mathbf{X})_{\mathbf{X}^\perp} \right)^{-1} \right\|$$

*(Again, operator norm is assumed).*

*Proof.* We first differentiate $G$. Let $(\mathbf{F}_t, \mathbf{X}_t)$ be a smooth path. Differentiating $\mathbf{F}_t(\mathbf{X}_t) \equiv 0$, one gets

$$D\mathbf{F}_t(\mathbf{X}_t)\dot{\mathbf{X}}_t + \dot{\mathbf{F}}_t(\mathbf{X}_t) = 0$$

Therefore,

$$DG(\mathbf{X}_t) : \dot{\mathbf{F}} \to -D\mathbf{F}_t(\mathbf{X}_t)^{-1} \begin{bmatrix} K_{d_1}(\cdot, \mathbf{X}_t)^* & & \\ & \ddots & \\ & & K_{d_n}(\cdot, \mathbf{X}_t)^* \end{bmatrix} \dot{\mathbf{F}}$$

The condition number and the right hand side of (1) are invariant by scalings in $\mathcal{H}_{\mathbf{d}}$, in $\mathbb{C}^{n+1}$ and also by unitary action $(\mathbf{f}, \mathbf{x}) \mapsto (\mathbf{f} \circ U^*, U\mathbf{x})$. Therefore we can assume without loss of generality that $\|\mathbf{F}\| = 1$ and that $\mathbf{X} = \mathrm{e}_0$. Calculations are immediate. □

Shub and Smale introduced the **normalized condition number**

$$\mu(\mathbf{f}, \mathbf{x}) = \|\mathbf{F}\| \left\| \left( \begin{bmatrix} \frac{\|\mathbf{X}\|^{-d_1+1}}{\sqrt{d_1}} & & \\ & \ddots & \\ & & \frac{\|\mathbf{X}\|^{-d_n+1}}{\sqrt{d_n}} \end{bmatrix} D\mathbf{F}(\mathbf{X})_{\mathbf{X}^\perp} \right)^{-1} \right\|.$$

The operator $\mathcal{H}_{\mathbf{d}} \mapsto \mathcal{H}_{(1,\cdots,1)}$ given by

$$\mathbf{F} \mapsto \begin{bmatrix} d_1^{-1/2}\|\mathbf{X}\|^{-d_1+1} & & \\ & \ddots & \\ & & d_n^{-1/2}\|\mathbf{X}\|^{-d_n+1} \end{bmatrix} D\mathbf{F}(\mathbf{X})_{\mathbf{X}^\perp}$$

is an isometric projection. This definition makes the condition theorem true:

**Theorem 5.2.** (Shub and Smale, 1993a) *The condition number $\mu(f, x)$ equal to the reciprocal of the distance of $f$ to the discriminant variety $\Sigma$ along the fiber of systems vanishing at $x$.*

(See Shub and Smale (1993a) or Blum et al. (1998) for the original version and Malajovich (2011) for generalizations).

Notice that

$$\|DG(\mathbf{f}, \mathbf{x})\| \le \mu(\mathbf{f}, \mathbf{x}) \le \sqrt{\max d_i}\|DG(\mathbf{f}, \mathbf{x})\|$$

## 6. The numerical analysis

One of the main results of the early Bézout saga was a family of path-following methods, with number of homotopy steps of

$$\mathcal{O}\left( d(\mathbf{f}_0, \mathbf{f}_1) \max_{t\in[0,1]} \mu(\mathbf{f}_t, \mathbf{x}_t)^2 \; dt \right).$$

The general procedure was of the form:

$$(2) \qquad \mathbf{x}_{i+1} = N(\mathbf{f}_{t_{i+1}}, \mathbf{x}_{t_i})$$

where $N$ denotes certain Newton iteration in projective space.

I must say now what is an approximate zero. Let $d(\mathbf{x}, \mathbf{y}) = \min \|\mathbf{X} - \lambda\mathbf{Y}\|/\|\mathbf{X}\|$ be the projective metric in projective space, that is the sine of the Riemannian distance.

**Definition 6.1** (Smale)**.** An **approximate zero** for $\mathbf{F} \in \mathcal{H}_{\mathbf{d}}$ is a point $0 \ne \mathbf{Y} \in \mathbb{C}^{n+1}$ so that the sequence $\mathbf{y}_{i+1} = N(\mathbf{f}, \mathbf{y}_i)$ satisfies

$$d(\mathbf{y}_i, \mathbf{y}_{i+1}) \le 2^{-2^i+1} d(\mathbf{y}_0, \mathbf{y}_1).$$

It turns out that approximate zeros exist (Smale, 1986) and can be numerically certified through Smale' s **alpha theory**. (Again see my book, Malajovich (2011)).

The outcome of Shub and Smale (1994) is that for every $\mathcal{H}_{\mathbf{d}}$, there is a pair $(\mathbf{F}_0, \mathbf{X}_0)$ such that, for every $\mathbf{F}_1$, there is a sequence $t_i$, so that (2) produces $t_N = 1$ and $\mathbf{X}_N$ so that $\mathbf{X}_N$ is an **approximate zero** for $\mathbf{F}_1$.

The following problems where left mostly open.

(1) How to find a good starting pair $(\mathbf{F}_0, \mathbf{X}_0)$.
(2) How to generate the sequence of $t_i$'s.

## 7. Smale's 17-th problem

**Open Problem 7.1.** (Smale, 1998) Can a zero of $n$ complex polynomial equations in $n$ unknowns be found approximately , on the average, in polynomial time with a uniform algorithm?

All the terms above are technical. Here is my translation:

*Does there exist a deterministic algorithm $M$ (BSS machine over $\mathbb{R}$ or a similar model) with input $(n \in \mathbb{N}, d_1 \in \mathbb{N}, \dots, d_n \in \mathbb{N}, \mathbf{F} \in \mathcal{H}_{\mathbf{d}})$ producing $\mathbf{X} \in \mathbb{C}^{n+1} \setminus \{0\}$ so that*

(1) $\mathbf{x}$ *is an approximate zero for* $\mathbf{f}$*, and*
(2) *There is a polynomial $p$ such that for any fixed* $\mathbf{d}$*,*

$$\mathrm{AVG}_{\mathbf{F} \in N(0,1;\mathcal{H}_{\mathbf{d}})} R(\mathbf{d}, \mathbf{F}) \le p(\dim(\mathcal{H}_{\mathbf{d}}))$$

*where $R(\mathbf{d}, \mathbf{F})$ is the running time of $M$ with input* $\mathbf{F}$*?*

Two major advances in this subject are a polynomial time **randomized** algorithm (Beltrán and Pardo, 2011) and a deterministic algorithm (Bürgisser and Cucker, 2011) that runs in time

$$(\dim \mathcal{H}_{\mathbf{d}})^{\log \log \dim \mathcal{H}_{\mathbf{d}}}.$$

## 8. The fast homotopy

Most homotopy algorithms until now prescribed an upper bound for the time mesh $t_{i+1} - t_i$. In Dedieu, Malajovich, and Shub (2013), we constructed an algorithm with a lower bound for the time mesh, in terms of a certain integral. (The existence of that time mesh appeared in Shub (2009), and another algorithm can be found in Beltrán (2011).

The algorithm in Dedieu et al. (2013) performs path-lifting in at most

$$1 + 0.65 (\max d_i)^{3/2} \epsilon^{-2} \mathcal{L}(\mathbf{f}_t, \mathbf{x}_t; 0, 1)$$

homotopy steps, where

$$\mathcal{L}(\mathbf{f}_t, \mathbf{x}_t; 0, 1) = \int_0^1 \mu(\mathbf{f}_t, \mathbf{z}_t) \left( \|\dot{\mathbf{f}}_t\|_{\mathbf{f}_t} + \|\dot{\mathbf{x}}_t\|_{\mathbf{x}_t} \right) \tag{3}$$

The algorithm is robust, and the accuracy parameter $\epsilon$ allows for approximate computations. This is where the idea of the $\alpha$-structure comes from.

## 9. Geometric forms of Smale's 17-th problem

Let $\alpha(\mathbf{f}, \mathbf{x}) = \mu(\mathbf{f}, \mathbf{x})^2$ and let $M = \{(\mathbf{f}, \mathbf{x}) \in V : \alpha(\mathbf{f}, \mathbf{x}) < \infty\}$.

**Conjecture 9.1.** *$\alpha$ is self-convex in $M$.*

In particular, $\mu$ would be convex along the geodesics of the $\alpha$-structure. The maximum of $\mu$ along a geodesic would be found at an extremity.

Moreover, a short geodesic (in the condition-structure) between an arbitrary $(\mathbf{f}, \mathbf{x})$ and a global minimum for $\mu$ is guaranteed to exist (Beltrán and Shub, 2009). At this time we do not know how to approximate such a geodesic in polynomial time.

Here is a possible approach for Smale's 17-th problem. *For every $\mathbf{f} \in \mathbb{P}(\mathcal{H}_{\mathbf{d}})$, produce a* **path** *$\mathbf{f}_t \in \mathbb{P}(\mathcal{H}_{\mathbf{d}})$ with $\mathbf{f}_1 = \mathbf{f}$, and produce $\mathbf{z}_0 \in \mathbb{P}^n$ so that the $\alpha$-length of the lifting of $\mathbf{f}_t$ passing through $(\mathbf{f}_0, \mathbf{z}_0)$ is $\le \dim(\mathcal{H}_{\mathbf{d}})^k$ (where $k$ must be a universal constant).*

The most technical algorithmic issues are gone. What we have above is a geometrical or variational problem.

## 10. More on convexity of the condition number

Last year, Wiesel (2012) pointed out another context where the geodesic convexity of the condition number plays a major role.

Assume that $\mathbf{x} = \nu \mathbf{u} \in \mathbb{R}^p$ is a random variable, where $\nu$ is a random scalar and $\mathbf{u}$ a Gaussian random vector with zero average and unknown covariance $A$. Actually, the covariance is defined up to a constant scaling, so it is an element of $\mathbb{RP}(\mathbb{S}^p_{++})$ where $\mathbb{S}^p_{++}$ is the cone of symmetric positive definite matrices $p \times p$.

Estimating $A$ from a sample $\mathbf{x}_i$ of $\mathbf{x}$, $i = 1, \dots, n$ is a relevant problem in applications such as radar or wireless communications (Wiesel, 2012). A possible approach is to normalize

$$\mathbf{s}_i = \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}$$

and then minimize the **negative log-likelihood**

$$l(\{\mathbf{s}_i\}; A) = \frac{p}{n} \sum_{i=1}^n \log(\mathbf{s}_i^T A^{-1} \mathbf{s}_i) + \log\det(A)$$

The **Tyler estimator** is an iteration to solve the problem above by repeatedly solving a certain minimization problem. It is known to be efficient for $n \gg p$. Wiesel (2012) suggests the following class of estimators for the case $n \gtrsim p$. Define a **penalty function** $h = h(A)$ to be scale invariant and convex on geodesics. Then at each step $k$, define

$$A_{k+1} = \arg \min_{\substack{A \succ 0 \\ h(A) \le h_0}} Q(A, A_k)$$

with

$$Q(A, B) = \frac{p}{n} \sum_{i=1}^n \left( \log(\mathbf{s}_i^T B^{-1} \mathbf{s}_i) + \frac{\mathbf{s}_i^T A^{-1} \mathbf{s}_i}{\mathbf{s}_i^T B^{-1} \mathbf{s}_i} \right) - p + \log\det(A)$$

where $h_0$ is a parameter. This function $Q(A, B)$ has the property that $Q(A, B) \ge l(\{\mathbf{s}_i\}; A)$ with equality when $B = A$.

His final suggestion for the penalty function is the condition number $h(A) = \|A\|_2 \|A^{-1}\|_2$. However, both $h(A)$ and

$Q(A,B)$ are geodesically convex with respect to the **Natural Covariance Metric** described by Smith (2005).

$$d_{\mathrm{cov}}(A_1, A_2) = \sqrt{\sum_k (\log \lambda_k)^2}$$

where the $\lambda_i$ are generalized eigenvalues of the matrix pencil $A_1 - \lambda A_2$. In general,

$$\langle B, C \rangle_A^{\mathrm{cov}} = \mathrm{tr}(B^* A^{-2} C).$$

This metric makes $h$ and $A \mapsto Q(A,B)$ geodesically convex.

## 11. Discussion

It is important to point out that covariance matrices are symmetric, so they admit a convenient metric where the condition number is convex.

From the point of view of the condition metric, the set of symmetric matrices is a geodesically complete manifold.

**Question 11.1.** *Is $\lambda_i^{-1}$ geodesically convex in the condition metric, for all $i$?*

An affirmative answer implies the geodesic convexity of $A \mapsto Q(A,B)$, and a regularization using condition metric would then be possible in theory. The current open problem to produce condition geodesics efficiently remains.

We may also invert the question. Consider the following metric in $\mathcal{V}$. As before, $(\mathbf{g},\mathbf{y}), (\mathbf{h},\mathbf{z}) \in T_{\mathbf{f},\mathbf{x}}$ are all represented by $(\mathbf{G},\mathbf{Y})$, $(\mathbf{H},\mathbf{Z}) \in \mathcal{H}_{\mathbf{d}} \times \mathbb{C}^{n+1} \setminus \{0\}$ correctly scaled with respect to the representative $(\mathbf{F},\mathbf{X})$ of $(\mathbf{f},\mathbf{x})$. Then we set

$$\begin{aligned}\langle (\mathbf{g},\mathbf{y}), (\mathbf{h},\mathbf{z}) \rangle_{\mathbf{f},\mathbf{x}}^{\mathrm{cov}} &= \langle M(\mathbf{F},\mathbf{X})\mathbf{G}, M(\mathbf{F},\mathbf{X})\mathbf{H} \rangle \\ &\quad + \langle M(\mathbf{F},\mathbf{X})^T \mathbf{Y}, M(\mathbf{F},\mathbf{X})^T \mathbf{Z} \rangle\end{aligned}$$

with

$$M(\mathbf{F},\mathbf{X}) = \|\mathbf{F}\| \left( \begin{bmatrix} \frac{\|\mathbf{X}\|^{-d_1+1}}{\sqrt{d_1}} & & \\ & \ddots & \\ & & \frac{\|\mathbf{X}\|^{-d_n+1}}{\sqrt{d_n}} \end{bmatrix} D\mathbf{F}(\mathbf{X})_{|\mathbf{X}^\perp} \right)^{-1}$$

We will have always

$$\|(\mathbf{g},\mathbf{y})\|_{(\mathbf{f},\mathbf{x})}^{\mathrm{cov}} \le \mu(\mathbf{f},\mathbf{x}) \|(\mathbf{g},\mathbf{y})\| = \|(\mathbf{g},\mathbf{y})\|'_{(\mathbf{f},\mathbf{x})}$$

**Question 11.2.** *Can the complexity of the algorithm by Dedieu, Malajovich, and Shub (2013) be estimated in terms of $\langle \cdot, \cdot \rangle^{\mathrm{cov}}$?*

**Question 11.3.** *Is $\log \mu$ geodesically convex wrt $\langle \cdot, \cdot \rangle^{\mathrm{cov}}$?*

## References


Beltrán, C. 2011. *A continuation method to solve polynomial systems and its complexity*, Numer. Math. **117**, no. 1, 89–113, DOI 10.1007/s00211-010-0334-3.

Beltrán, C., J.-P. Dedieu, G. Malajovich, and M. Shub. 2010. *Convexity properties of the condition number*, SIAM Journal on Matrix Analysis and Applications **31**, no. 3, 1491-1506, DOI 10.1137/080718681.

Beltrán, C., J.-P. Dedieu, G. Malajovich, and M. Shub. 2012. *Convexity properties of the condition number II*, SIAM J. on Matrix Analysis **33**, no. 3, 905-939.

Beltrán, C. and A. Leykin. 2013. *Robust certified numerical homotopy tracking*, Found. Comput. Math. **13**, no. 2, 253–295, DOI 10.1007/s10208-013-9143-2.

Beltrán, C. and L. M. Pardo. 2011. *Fast linear homotopy to find approximate zeros of polynomial systems*, Foundations of Computational Mathematics **11**, 95–129.

Beltrán, C. and M. Shub. 2009. *Complexity of Bezout's theorem. VII. Distance estimates in the condition metric*, Found. Comput. Math. **9**, no. 2, 179–195, DOI 10.1007/s10208-007-9018-5.

———. *On the geometry and topology of the solution variety for polynomial system solving.* to appear.

Blum, L., F. Cucker, M. Shub, and S. Smale. 1998. *Complexity and real computation*, Springer-Verlag, New York. With a foreword by Richard M. Karp.

Boito, P. and J.-P. Dedieu. 2010. *The condition metric in the space of rectangular full rank matrices*, SIAM J. Matrix Anal. Appl. **31**, no. 5, 2580–2602, DOI 10.1137/08073874X.

Bürgisser, P. and F. Cucker. 2011. *On a problem posed by Steve Smale*, Ann. of Math. (2) **174**, no. 3, 1785–1836, DOI 10.4007/annals.2011.174.3.8.

Dedieu, J.-P., G. Malajovich, and M. Shub. 2013. *Adaptative Step Size Selection for Homotopy Methods to Solve Polynomial Equations*, IMA J. on Numerical Analysis **33**, 1–29.

Li, Y. and L. Nirenberg. 2005. *Regularity of the distance function to the boundary*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5) **29**, 257–264.

Malajovich, G. 2011. *Nonlinear equations*, Publicações Matemáticas do IMPA. [IMPA Mathematical Publications], Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro. With an appendix by Carlos Beltrán, Jean-Pierre Dedieu, Luis Miguel Pardo and Mike Shub; 28º Colóquio Brasileiro de Matemática. [28th Brazilian Mathematics Colloquium].

Shub, M. 2009. *Complexity of Bezout's theorem. VI. Geodesics in the condition (number) metric*, Found. Comput. Math. **9**, no. 2, 171–178, DOI 10.1007/s10208-007-9017-6.

Shub, M. and S. Smale. 1993a. *Complexity of Bézout's theorem. I. Geometric aspects*, J. Amer. Math. Soc. **6**, no. 2, 459–501, DOI 10.2307/2152805.

Shub, M. and S. Smale. 1993b. *Complexity of Bezout's theorem. II. Volumes and probabilities*, Computational algebraic geometry (Nice, 1992), Progr. Math., vol. 109, Birkhäuser Boston, Boston, MA, pp. 267–285.

Shub, M. and S. Smale. 1993c. *Complexity of Bezout's theorem. III. Condition number and packing*, J. Complexity **9**, no. 1, 4–14, DOI 10.1006/jcom.1993.1002. Festschrift for Joseph F. Traub, Part I.

———. 1996. *Complexity of Bezout's theorem. IV. Probability of success; extensions*, SIAM J. Numer. Anal. **33**, no. 1, 128–148, DOI 10.1137/0733008.

Shub, M. and S. Smale. 1994. *Complexity of Bezout's theorem. V. Polynomial time*, Theoret. Comput. Sci. **133**, no. 1, 141–164, DOI 10.1016/0304-3975(94)90122-8. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993).

Smale, S. 1986. *Newton's method estimates from data at one point*, computational mathematics (Laramie, Wyo., 1985), Springer, New York, pp. 185–196.

———. 1998. *Mathematical problems for the next century*, Math. Intelligencer **20**, no. 2, 7–15, DOI 10.1007/BF03025291.

Smith, S. T. 2005. *Covariance, Subspace, and Intrinsic Cramér-Rao Bounds*, IEEE Transactions on Signal Processing **53**, no. 5, 1610–1630.

Wiesel, A. 2012. *Unified Framework to Regularized Covariance Estimation in Scaled Gaussian Models*, IEEE Transactions on Signal Processing **60**, no. 1, 29–38.



Departamento de Matemática Aplicada, Instituto de Matemática, Universidade Federal do Rio de Janeiro. Av. Athos da Silveira Ramos 149, Centro de Tecnologia - Bloco C Cidade Universitária - Ilha do Fundão. Caixa Postal 68530 21941-909 Rio de Janeiro - RJ - Brasil

*E-mail address*: `gregorio.malajovich@gmail.com`