ARTICLE IN PRESS

Journal of Complexity ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Journal of Complexity

journal homepage: www.elsevier.com/locate/jco

Stochastic perturbations and smooth condition numbers

Diego Armentano

Centro de Matemática, Facultad de Ciencias, Universidad de la República, Calle Iguá 4225, 11400 Montevideo, Uruguay

ARTICLE INFO

Article history:
Received 18 December 2008
Accepted 15 December 2009
Available online xxxx

This paper is dedicated to Mario Wschebor, on his 70th birthday

Keywords:
Condition numbers
Random matrices
Systems of linear equations
Finding kernels
Eigenvalue and eigenvector problems
Solving polynomial systems of equations

ABSTRACT

In this paper we define a new condition number adapted to directionally uniform perturbations in a general framework of maps between Riemannian manifolds. The definitions and theorems can be applied to a large class of problems. We show the relation with the classical condition number and study some interesting examples.

1. Introduction and main result

Let X and Y be two real (or complex) Riemannian manifolds of real dimensions m and n ($m \geq n$) associated respectively to some computational problem, where X is the space of *inputs* and Y is the space of *outputs*. Let $V \subset X \times Y$ be the *solution variety*, i.e. the subset of pairs (x, y) such that y is an output corresponding to the input x. Let $\pi_1 : V \to X$ and $\pi_2 : V \to Y$ be the canonical projections. The set of critical points of the projection π_1 is denoted by Σ', and let $\Sigma := \pi_1(\Sigma')$.

When $\dim V = \dim X$, for each $(x, y) \in V \setminus \Sigma'$, there is a differentiable function locally defined between some neighborhoods U_x and U_y of $x \in X$ and $y \in Y$ respectively, namely

$$G := \pi_2 \circ \pi_1^{-1}|_{U_x} : U_x \to U_y.$$

Let us denote by $\langle \cdot, \cdot \rangle_x$ and $\langle \cdot, \cdot \rangle_y$ the Riemannian (or Hermitian) inner product in the tangent spaces T_xX and T_yY at x and y respectively. The derivative $DG(x) : T_xX \to T_yY$ is called the *condition linear*

E-mail addresses: armentano.diego@gmail.com, diego@cmat.edu.uy.

0885-064X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jco.2010.01.003

operator at (x, y). The *condition number* at $(x, y) \in V \setminus \Sigma'$ is defined as

$$\kappa(x, y) := \max_{\substack{\dot{x} \in T_x X \\ \|\dot{x}\|_x^2 = 1}} \|DG(x)\dot{x}\|_y. \quad (1)$$

This number is an upper-bound—to first-order approximation—of the *worst-case* sensitivity of the output error with respect to small perturbations of the input. There is an extensive literature about the role of the condition number in the accuracy of algorithms, see for example Higham [12] and references therein.

Remark 1.1. Our general framework of maps between Riemannian manifolds was motivated by Shub–Smale [14] and Dedieu [8]. This general framework for a *computational problem* differs from the usual one, where the problem being solved can be described by a univalent function G. In the given context, we allow multi-valued functions, that is, we allow inputs with different outputs. In this way, one can define the condition number for the input $x \in X$ as a certain functional defined over $(\kappa(x, y))_{\{y \in \pi_2(\pi_1^{-1}(x))\}}$. When the function G is univalent the condition number $\kappa(x) := \kappa(x, y)$ coincides with the classical condition number (see Higham [12], p. 8). In what follows, we will restrict ourselves to study the condition number given by (1), but it is worth pointing out that all the analysis we will pursue here can be carried out to this kind of condition numbers without modifications.

In many practical situations, however, there exists a discrepancy between worst case theoretical analysis and observed accuracy of an algorithm. There exist several approaches that attempt to rectify this discrepancy. Among them we find *average-case analysis* (see Edelman [10], Smale [15]) and *smooth analysis* (see Spielman–Teng [16], Bürgisser–Cucker–Lotz [7], Wschebor [21]). For a comprehensive review on this subject with historical notes see Bürgisser [5].

In many problems, the space of inputs has a much larger dimension than the one of the space of outputs ($m \gg n$). Then, it is natural to assume that infinitesimal perturbations of the input will produce drastic changes in the output only when they are performed in a few directions. Then, a possibly different approach to analyze accuracy of algorithms is to replace "worst direction" by a certain mean over all possible directions. This alternative was already suggested and studied in Weiss et al. [19] in the case of the linear system solving $Ax = b$, and more generally, in Stewart [17] in the case of matrix perturbation theory, where the first-order perturbation expansion is assumed to be random.

In this paper we extend this approach to a large class of computational problems, restricting ourselves to the case of directionally uniform perturbations.

Generalizing the concept introduced in Weiss et al. [19] and Stewart [17], we define the *pth-stochastic condition number* at (x, y) as

$$\kappa_{st}{}^{[p]}(x, y) := \left[\frac{1}{\mathrm{vol}(S_x^{m-1})} \int_{\dot{x} \in S_x^{m-1}} \|DG(x)\dot{x}\|_y^p \, \mathrm{d}S_x^{m-1}(\dot{x}) \right]^{1/p}, \quad (p = 1, 2, \ldots), \quad (2)$$

where $\mathrm{vol}(S_x^{m-1}) = \frac{2\pi^{m/2}}{\Gamma(m/2)}$ is the measure of the unit sphere S_x^{m-1} in T_xX and $\mathrm{d}S_x^{m-1}$ is the induced volume element. We will be mostly interested in the case $p = 2$, which we simply write κ_{st} and call it the *stochastic condition number*.

Before stating the main theorem, we define the *Frobenius condition number* as

$$\kappa_F(x, y) := \|DG(x)\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2},$$

where $\|\cdot\|_F$ is the Frobenius norm and $\sigma_1, \ldots, \sigma_n$ are the singular values of the condition operator. Note that $\kappa_F(x, y)$ is a smooth function in $V \setminus \Sigma'$, where its differentiability class depends on the differentiability class of G.

Theorem 1.

$$\kappa_{st}{}^{[p]}(x, y) = \frac{1}{\sqrt{2}} \left[\frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+p}{2}\right)} \right]^{1/p} \cdot \mathbb{E}(\|\eta_{\sigma_1, \ldots, \sigma_n}\|^p)^{1/p},$$

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

where $\|\cdot\|$ *is the Euclidean norm in* $\mathbb{R}^n$ *and* $\eta_{\sigma_1,\ldots,\sigma_n}$ *is a centered Gaussian vector in* $\mathbb{R}^n$ *with the diagonal covariance matrix* $\mathrm{Diag}(\sigma_1^2, \ldots, \sigma_n^2)$.

In particular, for $p = 2$

$$\kappa_{st}(x, y) = \frac{\kappa_F(x, y)}{\sqrt{m}}. \tag{3}$$

Remark 1.2. Since $\kappa(x, y) \le \kappa_F(x, y) \le \sqrt{n} \cdot \kappa(x, y)$, we have from (3) that

$$\frac{1}{\sqrt{m}} \cdot \kappa(x, y) \le \kappa_{st}(x, y) \le \sqrt{\frac{n}{m}} \cdot \kappa(x, y).$$

This result is most interesting when $m \gg n$, for in that case $\kappa_{st}(x, y) \ll \kappa(x, y)$. Thus, in these cases one may expect much better stability properties than those predicted by classical condition numbers.

Remark 1.3. In many situations, one needs to analyze how the condition number varies in order to study (or to improve) the accuracy of an algorithm. In this way, the replacement of the usual non-smooth condition number κ given in (1) by a smooth one has an important theoretical and practical application.

In numerical analysis, many authors are interested in relative errors. Thus, when $(X, \langle\cdot,\cdot\rangle_X)$ and $(Y, \langle\cdot,\cdot\rangle_Y)$ are real (or complex) finite dimensional vector spaces with an inner (or Hermitian) product, instead of considering the (absolute) condition number (1), one can take the *relative condition number* defined as

$$\kappa_{rel}(x, y) := \frac{\|x\|_X}{\|y\|_Y} \cdot \kappa(x, y), \quad x \ne 0,\ y \ne 0;$$

and the *relative Frobenius condition number* as

$$\kappa_{relF}(x, y) := \frac{\|x\|_X}{\|y\|_Y} \cdot \kappa_F(x, y), \quad x \ne 0,\ y \ne 0,$$

where $\|\cdot\|_X$ and $\|\cdot\|_Y$ are the respective induced norms. In the same way, we define the *relative pth-stochastic condition number* as

$$\kappa_{relst}^{[p]}(x, y) := \frac{\|x\|_X}{\|y\|_Y} \cdot \kappa_{st}^{[p]}(x, y), \quad (p = 1, 2, \ldots). \tag{4}$$

For the case $p = 2$ we simply write κ_{relst} and call it the *relative stochastic condition number*.

In this case, we can define Riemannian structures on $X \setminus \{0\}$ and $Y \setminus \{0\}$ in the following way: for each $x \in X,\ x \ne 0$, and $y \in Y,\ y \ne 0$, we define

$$\langle\cdot,\cdot\rangle_x := \frac{\langle\cdot,\cdot\rangle_X}{\|x\|_X^2}, \quad \text{and} \quad \langle\cdot,\cdot\rangle_y := \frac{\langle\cdot,\cdot\rangle_Y}{\|y\|_Y^2}.$$

Notice that, in these Riemannian structures the usual condition number defined in (1) turns to be the relative condition number defined before. Then, Theorem 1 remains true if one exchanges the (absolute) condition number by the relative condition number. In particular,

$$\kappa_{relst}(x, y) := \frac{\kappa_{relF}(x, y)}{\sqrt{m}}.$$

2. Componentwise analysis

In the case $Y = \mathbb{R}^n$, we define the *kth-componentwise condition number* at $(x, y) \in V \setminus \Sigma'$ as

$$\kappa(x, y; k) := \max_{\substack{\dot{x} \in T_x X \\ \|\dot{x}\|_X^2 = 1}} |(DG(x)\dot{x})_k|, \quad (k = 1, \ldots, n), \tag{5}$$

where $|\cdot|$ is the absolute value and w_k indicates the kth-component of the vector $w \in \mathbb{R}^n$.

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

Following Weiss et al. [19] for the linear case, we define the *pth-stochastic kth-componentwise condition number* as

$$\kappa_{st}{}^{[p]}(x, y; k) := \left[\frac{1}{\mathrm{vol}(S_x^{m-1})} \int_{\dot{x}\in S_x^{m-1}} |(DG(x)\dot{x})_k|^p \, \mathrm{d}S_x^{m-1}(\dot{x})\right]^{1/p}, \quad (p = 1, 2, \ldots). \tag{6}$$

Then we have the following proposition.

Proposition 1.

$$\kappa_{st}{}^{[p]}(x, y; k) = \left[\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+p}{2}\right)} \cdot \Gamma\left(\frac{p+1}{2}\right)\right]^{1/p} \cdot \kappa(x, y; k).$$

In particular,

$$\kappa_{st}(x, y; k) = \frac{\kappa(x, y; k)}{\sqrt{m}}.$$

Proof. Observe that $\kappa_{st}{}^{[p]}(x, y; k)$ is the pth-stochastic condition number for the problem of finding the kth-component of $G = (G_1, \ldots, G_n) : X \to \mathbb{R}^n$. Theorem 1 applied to G_k yields

$$\kappa_{st}{}^{[p]}(x, y; k) = \frac{1}{\sqrt{2}} \left[\frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+p}{2}\right)}\right]^{\frac{1}{p}} \cdot \mathbb{E}(|\eta_{\sigma_1}|^p)^{1/p}$$

where $\sigma_1 = \|DG_k(x)\| = \kappa(x, y; k)$. Then,

$$\mathbb{E}(|\eta_{\sigma_1}|^p)^{1/p} = \kappa(x, y; k) \cdot \mathbb{E}(|\eta_1|^p)^{1/p},$$

where η_1 is a standard normal in $\mathbb{R}$. Finally,

$$\mathbb{E}(|\eta_1|^p) = \frac{2}{\sqrt{2\pi}} \int_0^\infty \rho^p \mathrm{e}^{-\rho^2/2} \mathrm{d}\rho = \frac{2}{\sqrt{2\pi}} 2^{\frac{p-1}{2}} \Gamma\left(\frac{p+1}{2}\right),$$

and the proposition follows. □

3. Examples

In this section we will compute the stochastic condition number for different problems: systems of linear equations, eigenvalue and eigenvector problems, finding kernels of linear transformations and solving polynomial systems of equations. The first two have been computed in Stewart [17] and are an easy consequence of Theorem 1 and the usual condition number κ.

The computations of κ for the case of systems of linear equations, eigenvalue and eigenvector problems, and solving polynomial systems of equations are fairly well-known. However, as far as we know, previous results of κ for the problem of finding kernels of linear transformations only offers bounds (see Kahan [13], Stewart–Sun [18], Beltran–Pardo [2]). In Section 3.3 we give an explicit computation of κ for this problem.

In what follows, we will drop the output in the notation of condition number when the input–output map is univalued.

3.1. Systems of linear equations

We consider the problem of solving for $y \in \mathbb{R}^n$ the system of linear equations $Ay = b$, $y \neq 0$, where $A \in \mathscr{M}_n(\mathbb{R})$ (the space of $n \times n$ real matrices), and $b \in \mathbb{R}^n$. If we assume that b is fixed, then we can consider the input space $X = \mathscr{M}_n(\mathbb{R})$ equipped with the Frobenius inner product

$$\langle A, B\rangle_F = \mathrm{trace}(AB^t), \tag{7}$$

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

where B^t is the transpose of B, and the output space $Y = \mathbb{R}^n$ equipped with the Euclidean inner product. It is easy to see that Σ equals the subset of non-invertible matrices. Then, the map $G : \mathscr{M}_n(\mathbb{R}) \setminus \Sigma \to \mathbb{R}^n$ is globally defined and differentiable, namely

$$G(A) = A^{-1}b(= y).$$

By implicit differentiation,

$$DG(A)\dot{A} = -A^{-1}\dot{A}y. \tag{8}$$

It is easy to see from (8) that

$$\kappa(A) = \|A^{-1}\| \cdot \|y\|.$$

Let H be the orthogonal complement of $\ker DG(A)$, i.e. H is the set of rank one matrices of the form uy^t, $u \in \mathbb{R}^n$, where y^t denotes the transpose of $y \in \mathbb{R}^n$. Then, the map $u \mapsto uy^t/\|y\|$ is a linear isometry between $\mathbb{R}^n$ and H. Under this identification, it is easy to see from (8) that $DG(A)|_H$ coincides with the map $-\|y\| \cdot A^{-1}$, from where we conclude

$$\kappa_F(A) = \|A^{-1}\|_F \cdot \|y\|.$$

Then, from Theorem 1 we get

$$\kappa_{st}(A) = \frac{\|A^{-1}\|_F \cdot \|y\|}{n},$$

and therefore

$$\kappa_{st}(A) \le \frac{\kappa(A)}{\sqrt{n}}. \tag{9}$$

A similar result was proved in Stewart [17].

For the general case, we consider $X = \mathscr{M}_n(\mathbb{R}) \times \mathbb{R}^n$ equipped with the product metric structure of the Frobenius inner product in $\mathscr{M}_n(\mathbb{R})$ and the Euclidean inner product in $\mathbb{R}^n$. Then, $G : \mathscr{M}_n(\mathbb{R}) \setminus \Sigma \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies $G(A, b) = A^{-1}b$.

Similar to the preceding case, we have $\kappa(A, b) = \|A^{-1}\| \cdot \sqrt{1 + \|y\|^2}$ and $\kappa_F(A, b) = \|A^{-1}\|_F \cdot \sqrt{1 + \|y\|^2}$. Again from Theorem 1 we get

$$\kappa_{st}(A, b) = \frac{\|A^{-1}\|_F \cdot \sqrt{1 + \|y\|^2}}{\sqrt{n^2 + n}},$$

and therefore

$$\kappa_{st}(A, b) \le \frac{\kappa(A, b)}{\sqrt{n + 1}}.$$

For the kth-componentwise condition number, we have that

$$\kappa_{st}{}^{[p]}((A, b); k) = \left[\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma\left(\frac{n^2+n}{2}\right)}{\Gamma\left(\frac{n^2+n+p}{2}\right)} \cdot \Gamma\left(\frac{p+1}{2}\right) \right]^{1/p} \cdot \kappa((A, b); k),$$

and

$$\kappa_{st}((A, b); k) = \frac{\kappa((A, b); k)}{\sqrt{n^2 + n}}.$$

A similar result was proved in Weiss et al. [19], where the average in (6) is performed over the unit ball instead of the unit sphere.

In Edelman [10], it is proved that the expected value of the relative condition number $\kappa_{rel}(A) = \|A\| \cdot \|A^{-1}\|$ of a random matrix A, whose elements are i.i.d standard normal, satisfies

$$\mathbb{E}(\log \kappa_{rel}(A)) = \log n + c + o(1),$$

as $n \to \infty$, where $c \approx 1.537$. If we consider the relative stochastic condition number defined in (4), we get from (9)

$$\mathbb{E}(\log \kappa_{relst}(A)) \leq \frac{1}{2}\log n + c + o(1),$$

as $n \to \infty$.

3.2. Eigenvalue and eigenvector problem

We focus on the complex case. The real case is analogue. We consider the problem of solving for $(\lambda, v) \in \mathbb{C} \times \mathbb{C}^n$ the system of equations $(\lambda I_n - A)v = 0,\ v \neq 0$, where $A \in \mathscr{M}_n(\mathbb{C})$ (the space of $n \times n$ complex matrices). Since this system of equations is homogenous in v, we define the solution variety associated with this problem as

$$V = \{(A, v, \lambda) \in \mathscr{M}_n(\mathbb{C}) \times \mathbb{P}(\mathbb{C}^n) \times \mathbb{C} : (\lambda I_n - A)v = 0\},$$

where $\mathbb{P}(\mathbb{C}^n)$ denotes the projective space associated with $\mathbb{C}^n$.

Following Shub–Smale [14], let $X = \mathscr{M}_n(\mathbb{C})$ be equipped with the Frobenius Hermitian inner product, i.e. the complex analogue of (7), and $Y = \mathbb{P}(\mathbb{C}^n) \times \mathbb{C}$ be equipped with the canonical product metric structure.

Then, for $(A, v, \lambda) \in V \setminus \Sigma'$, i.e. when λ is a simple eigenvalue (cf. Wilkinson [20]), the condition linear operators DG_1 and DG_2 associated with the eigenvector and eigenvalue problem are

$$DG_1(A)\dot{A} = \left(\pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\right)^{-1}\left(\pi_{v^\perp}\dot{A}v\right) \quad \text{and} \quad DG_2(A)\dot{A} = \frac{\langle \dot{A}v, u\rangle}{\langle v, u\rangle},$$

where $\pi_{v^\perp}$ denotes the orthogonal projection onto $v^\perp$ and u is some left eigenvector associated with λ, that is, $u^*A = \overline{\lambda}u^*$.

The associated condition numbers are

$$\kappa_1(A, v) = \left\|\left(\pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\right)^{-1}\right\| \quad \text{and} \quad \kappa_2(A, \lambda) = \frac{\|v\| \cdot \|u\|}{|\langle v, u\rangle|}. \tag{10}$$

From our Theorem 1, we get the respective stochastic condition numbers:

$$\kappa_{1st}(A, v) = \frac{1}{n}\left\|\left(\pi_{v^\perp}(\lambda I_n - A)|_{v^\perp}\right)^{-1}\right\|_F \leq \frac{1}{\sqrt{n}}\kappa_1(A, v),$$

$$\kappa_{2st}(A, \lambda) = \frac{1}{n}\kappa_2(A, \lambda).$$

A similar result for $\kappa_{2st}(A, \lambda)$ was proved in Stewart [17].

3.3. Finding kernels of linear transformations

For the sake of completeness of the exposition, we focus on the complex case. All ideas carry over naturally on the real case. Let $\mathscr{M}_{k,p}(\mathbb{C})$ be the linear space of $k \times p$ complex matrices with the Frobenius Hermitian inner product, and $\mathscr{R}_r \subset \mathscr{M}_{k,p}(\mathbb{C})$ be the subset of matrices of rank r. Given $A \in \mathscr{R}_r$ we consider the problem of finding the subspace F of $\mathbb{C}^p$ such that $Ax = 0$ for all $x \in F$, i.e. finding the kernel subspace $\ker(A)$ of A. For this purpose, we introduce the *Grassmannian* manifold $\mathbb{G}_{p,\ell}$ of complex subspaces of dimension ℓ in $\mathbb{C}^p$, where $\ell = p - r$ is the dimension of $\ker(A)$.

The input space $X = \mathscr{R}_r$ is a smooth submanifold of $\mathscr{M}_{k,p}(\mathbb{C})$ of complex dimension $(k + p)r - r^2$ (see Dedieu [9]). Thus, it has a natural Hermitian structure induced by the Frobenius Hermitian inner product on $\mathscr{M}_{k,p}(\mathbb{C})$.

In what follows, we identify $\mathbb{G}_{p,\ell}$ with the quotient $\mathbb{S}_{p,\ell}/\mathcal{U}_\ell$ of the *Stiefel* manifold

$$\mathbb{S}_{p,\ell} := \{M \in \mathscr{M}_{p,\ell}(\mathbb{C}) : M^*M = I\}$$

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

by the unitary group $\mathcal{U}_\ell \subset \mathcal{M}_\ell(\mathbb{C})$, which acts on the right of $\mathbb{S}_{p,\ell}$ in the natural way (see Dedieu [9]). Then, the complex dimension of the output space $Y = \mathbb{G}_{p,\ell}$ is $(p-r)r$. (We will use the same letter to represent an element of $\mathbb{S}_{p,\ell}$ and its class in $\mathbb{G}_{p,\ell}$).

The manifold $\mathbb{S}_{p,\ell}$ has a canonical Riemannian structure induced by the real part of the Frobenius Hermitian structure in $\mathcal{M}_{p,\ell}(\mathbb{C})$. On the other hand, $\mathcal{U}_\ell$ is a Lie group of isometries acting on $\mathbb{S}_{p,\ell}$. Therefore, $\mathbb{G}_{p,\ell}$ is a homogeneous space (see Gallot–Hulin–Lafontaine [11]), with a natural Riemannian structure that makes the quotient projection $\pi : \mathbb{S}_{p,\ell} \to \mathbb{G}_{p,\ell}$ a Riemannian submersion. More precisely, the orbit of $M \in \mathbb{S}_{p,\ell}$ under the action of the unitary group $\mathcal{U}_\ell$, namely, $\pi^{-1}(M) = \{MU : U \in \mathcal{U}_\ell\}$, defines a smooth submanifold of $\mathbb{S}_{p,\ell}$. In this way, the tangent space $T_M\mathbb{S}_{p,\ell}$ splits into two orthogonally complementary subspaces, namely,

$$T_M\mathbb{S}_{p,\ell} = T_M\pi^{-1}(M) \oplus \left(T_M\pi^{-1}(M)\right)^{\perp},$$

where $T_M\pi^{-1}(M)$ is the tangent space of $\pi^{-1}(M)$ at M. Then, we can naturally identify the tangent space $T_M\mathbb{G}_{p,\ell}$ with $\left(T_M\pi^{-1}(M)\right)^{\perp}$ with the inherited Riemannian structure induced by $\mathbb{S}_{p,\ell}$. Moreover, in this fashion, we can carry out all computations over the quotient manifold $\mathbb{G}_{p,\ell}$ onto $\mathbb{S}_{p,\ell}$.

To compute the derivative of the input–output map $G : \mathcal{R}_r \to \mathbb{G}_{p,\ell}$ which maps A onto $\ker(A)$, notice that if $M \in \mathbb{S}_{p,\ell}$ is any representative in $\pi^{-1}(\ker(A))$, then $AM = 0$. Then, implicit differentiation in the lift $\mathbb{S}_{p,\ell}$ yields

$$\dot{A}M + A(DG(A)\dot{A}) = 0,$$

where $\dot{A} \in T_A\mathcal{R}_r$, and $DG(A)\dot{A} \in T_M\mathbb{G}_{p,\ell}$. Then,

$$DG(A)\dot{A} = -A^{\dagger}\dot{A}M, \tag{11}$$

where $A^{\dagger}$ is the Moore–Penrose inverse of A.

We have concluded that the condition operator $DG(A)$ is a linear map from $T_A\mathcal{R}_r$ (with the Hermitian structure induced by $\mathcal{M}_{k,p}(\mathbb{C})$) onto $\left(T_M\pi^{-1}(M)\right)^{\perp}$ (with the inherited Riemannian structure of $\mathbb{S}_{p,\ell}$), and given by Eq. (11).

One way to compute the singular values of the condition operator described in (11) is to take an orthonormal basis in $\mathcal{M}_{k,p}(\mathbb{C})$ which diagonalizes A. From the singular value decomposition, there exists positive numbers $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and orthonormal basis $\{u_1, \ldots, u_k\}$ of $\mathbb{C}^k$ and $\{v_1, \ldots, v_p\}$ of $\mathbb{C}^p$, such that $A = \sum_{i=1}^{r} \sigma_i u_i v_i^*$ and $A^{\dagger} = \sum_{i=1}^{r} \sigma_i^{-1} v_i u_i^*$. Here w^* denotes the conjugate transpose of the vector w. Thus, $\{u_i v_j^* : i = 1, \ldots, k;\ j = 1, \ldots, p\}$ is an orthonormal basis of $\mathcal{M}_{k,p}(\mathbb{C})$ which diagonalizes A. On this basis the tangent space $T_A\mathcal{R}_r$ is the orthogonal complement of the subspace generated by $\{u_i v_j^* : i = r+1, \ldots, k;\ j = r+1, \ldots, p\}$.

Acting by an element $U \in \mathcal{U}_\ell$, if necessary, one can assume $M = \sum_{h=1}^{\ell} v_{h+r} e_h^*$, where $\{e_1, \ldots, e_\ell\}$ is the canonical basis of $\mathbb{C}^\ell$. Observe that $\|A^{\dagger}\dot{A}M\|_F \leq \|A^{\dagger}\| \cdot \|\dot{A}M\|_F$. Then,

$$\kappa(A) = \|A^{\dagger}\|,$$

where the maximum is attained, for example, at $\dot{A} = u_r v_{r+1}^* \in T_A\mathcal{R}_r$.

Observe that $\kappa_F(A)^2 = \sum_{i,j} \|DG(A)u_i v_j^*\|_F^2$, where the sum runs over all elements $u_i v_j^* \in T_A\mathcal{R}_r$. As $u_i v_j^* \in \ker DG(A)$, for $i = r+1, \ldots, p$ and $j = 1, \ldots, k$, then,

$$\kappa_F(A)^2 = \sum_{i=1}^{r}\sum_{j=1}^{p} \|A^{\dagger}u_i v_j^* M\|_F^2 = \sum_{i=1}^{r}\sum_{j=r+1}^{p} \|\sigma_i^{-1} v_i e_{j-r}^*\|_F^2 = (p-r) \cdot \sum_{i=1}^{r} \sigma_i^{-2}.$$

That is,

$$\kappa_F(A) = \sqrt{p-r} \cdot \|A^{\dagger}\|_F.$$

From our Theorem 1,

$$\kappa_{st}(A) = \frac{\sqrt{p-r}}{\sqrt{(k+p-r)r}} \cdot \|A^{\dagger}\|_F \leq \sqrt{\frac{p(p-r)}{(k+p-r)r}} \cdot \kappa(A).$$

In Beltŕan [1], it is proved that

$$\mathbb{E}(\log \kappa_{rel}(A) : A \in \mathcal{R}_r) \le \log\left[\frac{k+p-r}{k+p-2r+1}\right] + 2.6,$$

where the expected value is computed with respect to the normalized naturally induced measure in $\mathcal{R}_r$. Our Theorem 1 immediately yields a bound for the stochastic relative condition number, namely,

$$\mathbb{E}(\log \kappa_{relst}(A) : A \in \mathcal{R}_r) \le \frac{1}{2}\log\left[\frac{(k+p-r)r}{(k+p-2r+1)^2 p(p-r)}\right] + 2.6.$$

3.4. Finding roots problem I: Univariate polynomials

We start with the case of one polynomial in one complex variable. Let $X = \mathcal{P}_d = \{f : f(z) = \sum_{i=0}^{d} f_i z^i,\ f_i \in \mathbb{C}\}$. Identifying $\mathcal{P}_d$ with $\mathbb{C}^{d+1}$, we can define two standard Hermitian inner products in the space $\mathcal{P}_d$:

- Weyl inner product:

$$\langle f, g\rangle_W := \sum_{i=0}^{d} f_i \overline{g_i} \binom{d}{i}^{-1}; \tag{12}$$

- Canonical Hermitian inner product:

$$\langle f, g\rangle_{\mathbb{C}^{d+1}} := \sum_{i=0}^{d} f_i \overline{g_i}. \tag{13}$$

The solution variety is given by $V = \{(f, z) \in \mathcal{P}_d \times \mathbb{C} : f(z) = 0\}$ and $\Sigma' = \{(f, z) \in V : f'(z) = 0\}$. Thus, by implicit differentiation,

$$DG(f)(\dot{f}) = -\left(f'(\zeta)\right)^{-1} \dot{f}(\zeta).$$

We denote by κ_W and $\kappa_{\mathbb{C}^{d+1}}$ the condition numbers with respect to the Weyl and Hermitian inner product. The reader may check that

$$\kappa_W(f, \zeta) = \frac{(1+|\zeta|^2)^{d/2}}{|f'(\zeta)|} \quad \text{and} \quad \kappa_{\mathbb{C}^{d+1}}(f, \zeta) = \frac{\sqrt{\sum_{i=0}^{d} |\zeta|^{2i}}}{|f'(\zeta)|},$$

(for a proof see Blum et al. [4], p. 228). From Theorem 1, we get

$$\kappa_{W\,st}(f, \zeta) = \frac{1}{\sqrt{2(d+1)}}\kappa_W(f, \zeta), \qquad \kappa_{\mathbb{C}^{d+1}\,st}(f, \zeta) = \frac{1}{\sqrt{2(d+1)}}\kappa_{\mathbb{C}^{d+1}}(f, \zeta).$$

3.5. Finding roots problem II: systems of polynomial equations

We now study the case of complex homogeneous polynomial systems. Let $\mathcal{H}_d$ be the space of homogeneous polynomials in $n+1$ complex variables of degree $d \in \mathbb{N}\setminus\{0\}$. We consider $\mathcal{H}_d$ with the Hermitian inner product $\langle\cdot,\cdot\rangle_d$, namely, the homogeneous analogous of the Weyl structure defined above (see Chapter 12 of Blum et al. [4] for details).

Fix $d_1, \ldots, d_n \in \mathbb{N}\setminus\{0\}$ and let $\mathcal{H}_{(d)} = \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$ be the vector space of polynomial systems $f : \mathbb{C}^{n+1} \to \mathbb{C}^n$, $f = (f_1, \ldots, f_n)$, where $f_i \in \mathcal{H}_{d_i}$. The space $\mathcal{H}_{(d)}$ is naturally endowed with the Hermitian inner product $\langle f, g\rangle_W = \sum_{i=1}^{n} \langle f_i, g_i\rangle_{d_i}$.

Let $X = \mathbb{P}(\mathcal{H}_{(d)})$ and $Y = \mathbb{P}(\mathbb{C}^{n+1})$, then the solution variety is given by $V = \{(f, \zeta) \in \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1}) : f(\zeta) = 0\}$ and $\Sigma' = \{(f, \zeta) \in V : Df(\zeta)|_{\zeta^\perp} \text{ is singular}\}$.

We denote by $N = \sum_{i=1}^{n} \binom{d_i+n}{n} - 1$ the complex dimension of X. We may think of $2N$ as the size of the input.

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

Then, for $(f, \zeta) \in V \setminus \Sigma'$, we have

$$DG(f)\dot{f} = -\left(Df(\zeta)|_{\zeta^\perp}\right)^{-1}\dot{f}(\zeta),$$

and the condition number is

$$\kappa_W(f, \zeta) = \left\|\left(Df(\zeta)|_{\zeta^\perp}\right)^{-1}\right\|,$$

where some norm 1 affine representatives of f and ζ have been chosen (cf. Blum et al. [4]).

For the complexity analysis of path-following methods, it is convenient to consider the *normalized condition number* defined by

$$\kappa_{norm}(f, \zeta) = \left\|\left(Df(\zeta)|_{\zeta^\perp}\right)^{-1} \cdot \mathrm{Diag}(d_1^{1/2}, \ldots, d_n^{1/2})\right\|,$$

where $\mathrm{Diag}(d_1^{1/2}, \ldots, d_n^{1/2})$ denotes the diagonal matrix with entries $d_1^{1/2}, \ldots, d_n^{1/2}$. (Notice that κ_{norm} is the usual condition number for the slightly modified Hermitian inner product in $\mathcal{H}_{(d)}$ given by $\langle f, g\rangle_{norm} = \sum_{i=1}^n \frac{1}{d_i}\langle f_i, g_i\rangle_{d_i}$.)

Associated with κ_{norm}, we consider

$$\kappa_{norm}(f)^2 := \frac{1}{\mathcal{D}} \sum_{\{\zeta : f(\zeta)=0\}} \kappa_{norm}(f, \zeta)^2, \tag{14}$$

where $\mathcal{D} = d_1 \cdots d_n$ is the number of projective solutions of a generic system.

The expected value of $\kappa^2_{norm}(f)$ is an essential ingredient in the complexity analysis of path-following methods (cf. Shub–Smale [14], Beltran–Pardo [3], and recently Bürgisser–Cucker [6]). In Beltran–Pardo [3], the authors proved that

$$\mathbb{E}_f\left[\kappa_{norm}(f)^2\right] \le 8nN, \tag{15}$$

where f is chosen at random with the Weyl distribution.

The relation between complexity theory and the stochastic condition number is not clear yet. However, it is interesting to study the expected value of the κ_{st}-analogue of Eq. (14), namely

$$\kappa_{norm\,st}(f)^2 := \frac{1}{\mathcal{D}} \sum_{\{\zeta : f(\zeta)=0\}} \kappa_{norm\,st}(f, \zeta)^2.$$

Here $\kappa_{norm\,st}(f, \zeta)$ is the stochastic condition number for the modified condition operator, given by

$$\dot{f} \mapsto \left(Df(\zeta)|_{\zeta^\perp}\right)^{-1} \cdot \mathrm{Diag}(d_1^{1/2}, \ldots, d_n^{1/2}) \cdot \dot{f}(\zeta).$$

(Notice that, $\kappa_{norm\,st}(f, \zeta)$ is the stochastic condition number for the modified Hermitian inner product in $\mathcal{H}_{(d)}$ given by $\langle\cdot,\cdot\rangle_{norm}$).

From our Theorem 1 we get,

$$\kappa_{norm\,st}(f, \zeta) \le \frac{\kappa_{norm}(f, \zeta)}{\sqrt{N/n}}, \qquad \mathbb{E}_f\left[\kappa_{norm\,st}(f)^2\right] \le 8n^2.$$

Note that the last bound depends on the number of unknowns n, and not on the size of the input $N \gg n$.

4. Proof of the main theorem

In the case of complex manifolds, the condition matrix turns to be an $n \times n$ complex matrix. In what follows, we identify it with the associated $2n \times 2n$ real matrix. We focus on the real case.

The main theorem follows immediately from Lemma 1 and Proposition 2 below.

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

Lemma 1. *Let η be a Gaussian standard random vector in $\mathbb{R}^m$. Then*

$$\kappa_{st}{}^{[p]}(x,y) = \frac{1}{\sqrt{2}}\left[\frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m+p}{2}\right)}\right]^{1/p} \cdot \left[\mathbb{E}(\|DG(x)\eta\|^p)\right]^{1/p},$$

where $\mathbb{E}$ is the expectation operator and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^n$.

Proof. Let $f : \mathbb{R}^m \to \mathbb{R}$ be the continuous function given by

$$f(v) = \|DG(x)v\|.$$

Then,

$$\left[\mathbb{E}(\|DG(x)\eta\|^p)\right]^{1/p} = \left[\frac{1}{(2\pi)^{m/2}}\int_{\mathbb{R}^m} f(v)^p \cdot \mathrm{e}^{-\|v\|^2/2}\mathrm{d}v\right]^{1/p}.$$

Integrating in polar coordinates, we get

$$\mathbb{E}(\|DG(x)\eta\|^p) = \frac{I_{m+p-1}}{(2\pi)^{m/2}} \cdot \int_{S^{m-1}} f^p \mathrm{d}S^{m-1}, \tag{16}$$

where

$$I_j = \int_0^{+\infty} \rho^j \mathrm{e}^{-\rho^2/2}\mathrm{d}\rho, \quad j \in \mathbb{N}.$$

Making the change of variable $u = \rho^2/2$, we obtain

$$I_j = 2^{\frac{j-1}{2}}\Gamma\left(\frac{j+1}{2}\right);$$

therefore

$$I_{m+p-1} = 2^{\frac{m+p-2}{2}} \cdot \Gamma\left(\frac{m+p}{2}\right). \tag{17}$$

Then, joining together (16) and (17), we obtain the result. □

Proposition 2. *If η is a Gaussian standard random vector in $\mathbb{R}^m$, then*

$$\mathbb{E}(\|DG(x)\eta\|^p) = \mathbb{E}(\|\eta_{\sigma_1,\ldots,\sigma_n}\|^p),$$

where $\eta_{\sigma_1,\ldots,\sigma_n}$ is a centered Gaussian vector in $\mathbb{R}^n$ with the diagonal covariance matrix $\mathrm{Diag}(\sigma_1^2,\ldots,\sigma_n^2)$, *and $\sigma_1,\ldots,\sigma_n$ are the singular values of $DG(x)$.*

Proof. Let $DG(x) = UDV$ be a singular value decomposition of $DG(x)$, where V and U are orthogonal transformations of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, and $D := \mathrm{Diag}(\sigma_1,\ldots,\sigma_n)$. By the invariance of the Gaussian distribution under the action of the orthogonal group in $\mathbb{R}^m$, $V\eta$ is again a Gaussian standard random vector in $\mathbb{R}^m$. Then,

$$\mathbb{E}(\|DG(x)\eta\|^p) = \mathbb{E}(\|UD\eta\|^p),$$

and by the invariance under the action of the orthogonal group of the Euclidean norm, we get

$$\mathbb{E}(\|DG(x)\eta\|^p) = \mathbb{E}(\|D\eta\|^p).$$

Finally $D\eta$ is a centered Gaussian vector in $\mathbb{R}^n$ with the covariance matrix $\mathrm{Diag}(\sigma_1^2,\ldots,\sigma_n^2)$, and the proposition follows. For the case $p = 2$,

$$\kappa_{st}(x,y) = \left[\mathbb{E}\left(\sigma_1^2\eta_1^2 + \cdots + \sigma_n^2\eta_n^2\right)\right]^{1/2},$$

Please cite this article in press as: D. Armentano, Stochastic perturbations and smooth condition numbers, Journal of Complexity (2010), doi:10.1016/j.jco.2010.01.003

where $\eta_1, \ldots, \eta_n$ are i.i.d. standard normal in $\mathbb{R}$. Then,

$$\kappa_{st}(x, y) = \left(\sum_{i=1}^{n} \sigma_i^2\right)^{1/2} = \kappa_F(x, y). \quad \square$$

Acknowledgments

The author thanks Carlos Beltrán, Jean–Pierre Dedieu, Michael Shub and Mario Wschebor for useful discussions. Also, the author thanks the anonymous referee whose comments helped to improve the paper. The author was partially supported by CSIC, Uruguay.

References

[1] C. Beltrán, Estimates on the condition number of random, rank-deficient matrices, IMA J. Numer. Anal. (2009), doi:10.1093/imanum/drp035.
[2] C. Beltrán, L.M. Pardo, Estimates on the distribution of the condition number of singular matrices, Found. Comput. Math. 7 (1) (2007) 87–134.
[3] C. Beltrán, L.M. Pardo, Fast linear homotopy to find approximate zeros of polynomial systems (in press).
[4] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer-Verlag, New York, 1998.
[5] P. Bürgisser, Smoothed analysis of condition numbers, in: Foundations of Computational Mathematics, Hong Kong, in: London Mathematical Society Lecture Note Series, vol. 363, Cambridge University Press, 2009, pp. 1–41.
[6] P. Bürgisser, F. Cucker, On a problem posed by steve smale. arxiv:0909.2114v3 [math.NA].
[7] P. Bürgisser, F. Cucker, M. Lotz, Smoothed analysis of complex conic condition numbers, J. Math. Pures Appl. 86 (2006) 293–309.
[8] J.P. Dedieu, Approximate solutions of numerical problems, condition number analysis and condition number theorem, in: The Mathematics of Numerical Analysis (Park City, UT, 1995), in: Lectures in Appl. Math., vol. 32, Amer. Math. Soc., Providence, RI, 1996, pp. 263–283.
[9] J.P. Dedieu, Points Fixes, Zéros et la Méthode de Newton. Mathématiques & Applications, SMAI-Springer Verlag, 2006, p. 54.
[10] A. Edelman, Eigenvalues and condition numbers of random matrices, SIAM J. Matrix Anal. Appl. 9 (1988) 543–560.
[11] S. Gallot, D. Hulin, J. Lafontaine, Riemannian Geometry (Universitext), Third ed., Springer-Verlag, Berlin, 2004.
[12] N. Higham, Accuracy and Stability of Numerical Algorithms, second ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.
[13] W. Kahan, Huge generalized inverses of rank-deficient matrices, Pre-print.
[14] M. Shub, S. Smale, Complexity of Bezoutś theorem. IV. Probability of success; extensions, SIAM J. Numer. Anal. 33 (1) (1996) 128–148.
[15] S. Smale, On the eficiency of algorithms of analysis, Bull. Amer. Math. Soc. 13 (1985) 87–121.
[16] D.A. Spielman, S.H. Teng, Smoothed analysis of algorithms, ICM 2002, Beijing, vol. I, 2002, pp. 597–606.
[17] G.W. Stewart, Stochastic perturbation theory, SIAM Rev. 32 (4) (1990) 579–610.
[18] G.W. Stewart, G.J. Sun, Matrix perturbation theory, in: Computer Science and Scientific Computing, Academic Press, Inc., Boston, MA, 1990.
[19] N. Weiss, G.W. Wasilkowski, H. Wozniakowski, M. Shub, Average condition number for solving linear equations, Linear Algebra Appl. 83 (1986) 79–102.
[20] J.H. Wilkinson, Note on matrices with a very ill-conditioned eigenproblem, Numer. Math. 19 (1972) 176–178.
[21] M. Wschebor, Smoothed analysis of $\kappa(a)$, J. Complexity 20 (2004) 97–107.