Journal of Complexity 26 (2010) 161–171

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 28 January 2010

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
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

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

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

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

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.

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