Upper Bounds on the Distribution of the Condition Number of Singular Matrices

Bornes Supérieures pour la fonction de distribution du conditionnement des Matrices Singulières

Carlos Beltrán [a,1] Luis Miguel Pardo [b,1]

[a] *Dept. de Matemáticas, Estadística y Computación. F. de Ciencias. U. Cantabria. E-39071 SANTANDER, Spain.*
[b] *Dept. de Matemáticas, Estadística y Computación. F. de Ciencias. U. Cantabria. E-39071 SANTANDER, Spain.*

Abstract

We exhibit upper bounds for the probability distribution of the generalized condition number of singular complex matrices. To this end, we develop a new technique to study volumes of tubes about projective varieties in the complex projective space. As a main outcome, we show an upper bound estimate for the volume of the intersection of a tube with an equi–dimensional projective algebraic variety.

Résumé

Nous exhibons des bornes de la fonction de distribution du conditionnement des matrices singulières. Pour ce but nous developpons une technique nouvelle pour analyser les volumes des tubes (par rapport a la distance de Fubini-Study) autour des sous–variétés algèbriques d'un espace projectif complex. Plus spécifiquement, nous demontrons des bornes supérieueres de volumes des intersections des tubes extrinsèques (autour des sous-variétés algébriques avec une autre variété algèbrique donnée).

Version française abrégée

Le but de ces pages est celui d'exhiber des estimations fines de la fonction de distribution du conditionnement singulier dans la variété projective complexe des matrices de rang borné. Autrement dit, soit Σ^r la varété projective algèbrique complexe de tous les matrices carrées $n \times n$ et complexes de rang au plus r. Sout $\kappa_D^r(A) := ||A||_F||A^\dagger||$, le conditionnement d'un matrice $A \in \Sigma^r$, où $A^\dagger$ est la pseudo–inverse

Email addresses: `beltranc@unican.es` (Carlos Beltrán), `luis.pardo@unican.es` (Luis Miguel Pardo).
[1] Research was partially supported by MTM2004-01167.

de Moore–Penrose. Rappelons que κ_D^r mesure la perte de precission des algorithmes numériques pour le calcul du noyau de A. Nous montrons, par exemple, l'énoncé suivant :

Théorème 0.1 *Avec les notations précédentes, on a*

$$\frac{vol[A \in \Sigma^{n-1} : \kappa_D^{n-1}(A) > \varepsilon^{-1}]}{vol[\Sigma^{n-1}]} \leq 18n^{20}\varepsilon^6,$$

où $vol[\cdot]$ est la mesure de Haussdorff associée à la dimension de la variété Σ^{n-1}.

Notons que la fraction a à gauche de la formule précédente est la probabilité du fait q'une matrice complexe de rang au plus $n-1$ aie un conditionnement plus grand que ε^{-1}.

Cet énoncé est la consequence des résultats plus généraux sur les volumes des tubes (par rapport á la distance de Fubini-Study) autour des sous–variétés algèbriques d'un espace projectif complex (voir Théorème 2.1). Plus concrètement, cet Théorème est la consequence immédiate de l'énoncé suivant :

Théorème 0.2 *Soient $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ deux variétés projectives complexes à dimensions respectives $m > m' \geq 1$. Soit $0 < \varepsilon \leq 1$ un nombre réel positif. Alors, on a :*

$$\frac{\nu_m[V'_\varepsilon \cap V]}{\nu_m[V]} \leq 2\deg(V') \left(\frac{en}{n-m'}\right)^{2(n-m')} \left[e\frac{n-m'}{m-m'}\,\varepsilon\right]^{2(m-m')},$$

où ν_m est la mesure $2m-$dimensionelle naturellement associée a la variété V, et $\deg(V')$ est le degré de la variété V' dans le sense de [10].

Ces deux énoncés ont quelques variations quand on les applique a des differents sous–ensembles de l'espace de matrices. La technique de base est une combinaison de la Géometrie Intégrale et la Théorie de l'Intersections Géometrique.

1. Introduction

Condition numbers in Linear Algebra were introduced by A. Turing in [19]. They were studied by J. von Neumann and collaborators in [14]. Variations of these condition numbers may be found in the literature of Numerical Linear Algebra (cf. [2], [7], [11], [18], [21] etc.).

A relevant breakthrough was the study of the probability distribution of these condition numbers. The works by Steve Smale (cf. [16]) and mainly the works by A. Edelman (cf. [4], [5]) showed the exact values of the probability distribution of the condition number of dense complex matrices.

From a computational point of view, these statements can be translated in the following terms. Let $\mathcal{P}$ be a numerical analysis procedure whose space of input data is the space of arbitrary square complex matrices $\mathcal{M}_n(\mathbb{C})$. Then, Edelman's statements mean that *the probability that a randomly chosen dense matrix in $\mathcal{M}_n(\mathbb{C})$ is a well–conditioned input for $\mathcal{P}$ is high* (cf. also [1]).

Sometimes however we deal with procedures $\mathcal{P}$ whose input space is a proper subset $\mathcal{C} \subseteq \mathcal{M}_n(\mathbb{C})$. Additionally such procedures with particular data lead to particular condition numbers $\kappa_{\mathcal{C}}$ adapted both for the procedure $\mathcal{P}$ and the input space $\mathcal{C}$. Edelman's and Smale's results do not apply to these new conditions.

In these pages we introduce a new technique to study the probability distribution of condition numbers $\kappa_{\mathcal{C}}$. Namely, we introduce a technique to exhibit upper bound estimates of the quantity

$$\frac{vol[\{A \in \mathcal{C} : \kappa_{\mathcal{C}} > \varepsilon^{-1}\}]}{vol[\mathcal{C}]}, \tag{1}$$

where $\varepsilon > 0$ is a positive real number, and $vol[\cdot]$ is some suitable measure on the space $\mathcal{C}$ of acceptable inputs of $\mathcal{P}$.

2. The condition number for Singular matrices.

As an example of how our technique applies, let $\mathcal{C} := \Sigma^{n-1} \subseteq \mathcal{M}_n(\mathbb{C})$ be the class of all singular complex matrices. From [12] and [17], a condition number for singular matrices $A \in \mathcal{C}$ is introduced. This condition number measures the precision required to perform kernel computations . For every singular matrix $A \in \Sigma^{n-1}$ of corank 1, the condition number $\kappa_D^{n-1}(A) \in \mathbb{R}$ is defined by the following identity

$$\kappa_D^{n-1}(A) := \|A\|_F \|A^\dagger\|_2,$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix A, $A^\dagger$ is the Moore–Penrose pseudo–inverse of A and $\|A^\dagger\|_2$ is the norm of $A^\dagger$ as a linear operator.
As Σ^{n-1} is a complex homogeneous hypersurface in $\mathcal{M}_n(\mathbb{C})$(i.e. a cone of complex codimension 1), it is endowed with a natural volume *vol* induced by the $2(n^2-1)-$dimensional Hausdorff measure of its intersection with the unit disk. We state the following result.
Theorem 2.1 *With the same notations and assumptions as above, the following inequality holds:*

$$\frac{vol[A \in \Sigma^{n-1} : \kappa_D^{n-1}(A) > \varepsilon^{-1}]}{vol[\Sigma^{n-1}]} \leq 18 n^{20} \varepsilon^6.$$

Other proper subclasses of $\mathcal{M}_n(\mathbb{C})$ can also be discussed with our method. These statements are (almost) immediate consequences of a wider class of results we state below.

3. On the volume of tubes.

First of all, most condition numbers are by nature projective functions. For instance, the condition number κ_D of Numerical Linear Algebra ($\kappa_D(A) = \|A\|_F \|A^{-1}\|_2, \forall A \in \mathcal{M}_n(\mathbb{C})$) is naturally defined as a function on the complex projective space $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$ defined by the complex vector space $\mathcal{M}_n(\mathbb{C})$. Namely, we may see κ_D as a function

$$\kappa_D : \mathbb{P}(\mathcal{M}_n(\mathbb{C})) \longrightarrow \mathbb{R}_+ \cup \infty.$$

Secondly, statements like the Schmidt–Mirsky–Eckart–Young Theorem (cf. [3],[15], [13]) imply that Smale's and Edelman's estimates are, in fact, estimates of the volume of a tube about a concrete projective algebraic variety in $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$. In fact, the following two equalities hold:

$$\kappa_D(A) = \frac{1}{d_{\mathbb{P}}(A, \Sigma^{n-1})}, \qquad \forall A \in \mathbb{P}(\mathcal{M}_n(\mathbb{C})), \qquad \kappa_D^{n-1}(A) = \frac{1}{d_{\mathbb{P}}(A, \Sigma^{n-2})}, \qquad \forall A \in \mathbb{P}(\Sigma^{n-1}),$$

where $d_{\mathbb{P}}$ stands for the projective distance (namely, $d_P(x,y) = \sin d_R(x,y)$ where d_R is the Fubini-Study distance in $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$, and Σ^{n-2} stands for the set of all the matrices of rank equal to or smaller than $n-2$. We prove a general upper bound for the volume of a tube about any (possibly singular) complex projective algebraic variety (see Theorem 3.1 below).
Estimates on volumes of tubes is a classic topic that began with Weyl's Tube Formula for tubes in the affine space (cf. [20]). Formulae for the volumes of some tubes about analytic submanifolds of complex projective spaces are due to A. Gray (cf. [8], [9] and references therein). However, Gray's results do not apply to our case. First of all, Gray's statements are only valid for smooth submanifolds and not for singular varieties (as, for instance, Σ^{n-1}). Secondly, Gray's Theorems are only valid for tubes of small enough radius (depending on intrinsic features of the manifold under consideration).
These two drawbacks pushed us to look for a general statement that may be resumed as follows.

Let $d\nu_n$ be the volume form associated to the complex Riemannian structure of $\mathbb{P}_n(\mathbb{C})$. Let $V \subseteq \mathbb{P}_n(\mathbb{C})$ be any subset of the complex projective space and let $\varepsilon > 0$ be a positive real number. We define *the tube of radius ε about V in $\mathbb{P}_n(\mathbb{C})$* as the subset $V_\varepsilon \subseteq \mathbb{P}_n(\mathbb{C})$ given by the following identity.

$$V_\varepsilon := \{x \in \mathbb{P}_n(\mathbb{C}) : d_{\mathbb{P}}(x, V) < \varepsilon\}.$$

Theorem 3.1 *Let $V \subseteq \mathbb{P}_n(\mathbb{C})$ be a (possibly singular) equi–dimensional complex algebraic variety of (complex) codimension r in $\mathbb{P}_n(\mathbb{C})$. Let $0 < \varepsilon \leq 1$ be a positive real number. Then, the following inequality holds*

$$\frac{\nu_n[V_\varepsilon]}{\nu_n[\mathbb{P}_n(\mathbb{C})]} \leq 2 \deg(V) \left(\frac{e\ n\ \varepsilon}{r}\right)^{2r},$$

where $\deg(V)$ *is the degree of V (in the sense of [10]).*
This Theorem can be applied to Edelman's conditions to conclude the following estimate:

$$\frac{vol[\{A \in \mathcal{M}_n(\mathbb{C}) : \kappa_D(A) > \varepsilon^{-1}\}]}{vol[\mathcal{M}_n(\mathbb{C})]} \leq 2e^2 n^5 \varepsilon^2,$$

where $\kappa_D(A) := \|A\|_F \|A^{-1}\|_2$, and *vol* is the standard Gaussian measure in $\mathbb{C}^{n^2}$. The constants on the left–hand side of the inequality in Theorem 3.1 are essentially optimal.
The reader will observe that our bound is less sharp than Edelman's or Smale's ones although it is a particular instance of a more general statement.

4. Extrinsic tubes.

Observe that neither Smale's, Edelman's results nor Theorem 3.1 above apply to exhibit upper bounds of the probability distribution described in equation (1) above. In particular, it does not apply to prove Theorem 2.1. In order to prove this kind of statements, *we need an upper bound for the volume of the intersection of an extrinsic tube with a proper subvariety.* This is our main outcome here and can be resumed in the following statement.
Theorem 4.1 *Let $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ be two projective equi–dimensional algebraic varieties of respective dimensions $m > m' \geq 1$. Let $0 < \varepsilon \leq 1$ be a positive real number. With the same notations as in Theorem 3.1 above, the following inequality holds:*

$$\frac{\nu_m[V'_\varepsilon \cap V]}{\nu_m[V]} \leq 2 \deg(V') \left(\frac{en}{n - m'}\right)^{2(n-m')} \left[e \frac{n - m'}{m - m'}\ \varepsilon\right]^{2(m-m')},$$

where ν_m is the $2m-$dimensional natural measure in the algebraic variety V, and $\deg(V')$ *is the degree of V' in the sense of [10].*
Hence, Theorem 2.1 follows from Theorem 4.1, as the degree and codimension of Σ^{n-2} are known (cf. for example [6]):

$$codim(\Sigma^{n-2}) = 4, \qquad \deg(\Sigma^{n-2}) \leq \frac{n^4}{12}.$$

Please note that the changes suggested of the referee have been made: The word “sharp” (“fines” in the French version) has been removed, and the sentence “We state the following statement” before Th. 1 has been changed.

References

[1] D. Castro, J. L. Montaña, L. M. Pardo, and J. San Martín. The distribution of condition numbers of rational data of bounded bit length. *Found. Comput. Math.*, 2 (2002), 1–52.

[2] J. W. Demmel. The probability that a numerical analysis problem is difficult. *Math. Comp.*, 50 (1988), 449–480.

[3] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1 (1936), 211–218.

[4] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.*, 9 (1988), 543–560.

[5] A. Edelman. On the distribution of a scaled condition number. *Math. Comp.*, 58 (1992), 185–190.

[6] W. Fulton. *Intersection theory*, vol. 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1984.

[7] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[8] A. Gray. Volumes of tubes about complex submanifolds of complex projective space. *Trans. Amer. Math. Soc.*, 291 (1985), 437–449.

[9] A. Gray. *Tubes*, vol. 221 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, second edition, 2004.

[10] J. Heintz. Definability and fast quantifier elimination in algebraically closed fields. *Theoret. Comput. Sci.*, 24 (1983), 239–277.

[11] N. J. Higham. *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2002.

[12] W. Kahan. Huge Generalized Inverses of Rank-Deficient Matrices, Unpublished Manuscript.

[13] L. Mirsky. Results and problems in the theory of doubly-stochastic matrices. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 1 (1962/1963), 319–334.

[14] J. von Neumann and H. H. Goldstine. Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, 53 (1947), 1021–1099.

[15] E. Schmidt. Zur Theorie der linearen und nichtlinearen Integralgleichungen. I Tiel. Entwicklung willkurlichen Funktionen nach System vorgeschriebener. *Mahtematische Annalen*, 63 (1907), 433–476.

[16] S. Smale. On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc. (N.S.)*, 13 (1985), 87–121.

[17] G. W. Stewart and J. G. Sun. *Matrix perturbation theory*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1990.

[18] L. N. Trefethen and D. Bau, III. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[19] A. M. Turing. Rounding-off errors in matrix processes. *Quart. J. Mech. Appl. Math.*, 1 (1948), 287–308.

[20] A. Weyl. On the volume of tubes. *Amer. J. Math*, 61 (1939), 461–472.

[21] J. H. Wilkinson. *The algebraic eigenvalue problem*. Clarendon Press, Oxford, 1965.