Universidad Central de Venezuela
Facultad de Ciencias
Escuela de Computación

Lecturas en Ciencias de la Computación
ISSN 1316-6239

Geometrical properties of the Frobenius condition number for positive definite matrices

J.P. Chehab y M. Raydan.

RT 2007-05

Centro de Cálculo Científico y Tecnológico de la UCV
CCCT-UCV
Caracas, Mayo, 2007.

Geometrical properties of the Frobenius condition number for positive definite matrices

Jean-Paul Chehab[*] Marcos Raydan [†]

May 24, 2007

Abstract

We study the geometrical properties of the Frobenius condition number on the cone of symmetric and positive definite matrices. This number, related to the cosine of the angle between a given matrix and its inverse, is equivalent to the classical 2-norm condition number, but has a direct and natural geometrical interpretation. In particular we establish sharp bounds for the ratio between the angle that a matrix form with the identity ray and the angle that the inverse of that matrix form with the identity ray. These bounds allow us to show new lower bounds for the condition number, that only require the trace and the Frobenius norm of the matrix.

Key words: Condition number, cones of matrices, Frobenius norm.

1 Introduction

The space of square real $n \times n$ matrices can be equipped with the Frobenius inner product defined by

$$< A, B >_F = tr(A^T B),$$

for which we have the associated norm that satisfies $\|A\|_F^2 =< A, A >_F$. In here, $tr(A) = \sum_i a_{ii}$ is the trace of the matrix A. In this inner product space, the Frobenius condition number of positive definite matrices has a geometrical interpretation that cannot be shared by the extensively used 2-norm condition number. From that geometry several results will follow that are useful for practical aspects like estimating the condition number of a given matrix. In particular we explore in this work the relationship between the angle that a matrix forms with its inverse and the Frobenius condition number of the matrix. In this context, as we will see later, the angle that any matrix form with the identity ray, αI for $\alpha > 0$, plays a very important role.

[*]Laboratoire de Mathématiques Paul Painlevé, UMR 8524, Université de Lille 1, France,(chehab@math.univ-lille1.fr), and Laboratoire de Mathématiques, CNRS, UMR 8628, Equipe ANEDP, Université Paris Sud, Orsay, France.

[†]Departamento de Computación, Facultad de Ciencias, Universidad Central de Venezuela, Ap. 47002, Caracas 1041-A, Venezuela (mraydan@kuaimare.ciens.ucv.ve). Sponsored by the Scientific Computing Center at UCV.

This work has been motivated by the rich geometrical structure of the positive semidefinite (PSD) cone of matrices and specially by the discussion presented by Tarazaga [9, 10], in which, among other results, we learned that the identity ray is identified as the "center ray" of the cone, therefore representing the natural reference in the cone, for studying the location of a given matrix. For example, all the rank one symmetric matrices (i.e. matrices of the form xx^t, for $x \neq 0$) form the same angle with I, and moreover they are the "farthest away" from it, among all the matrices in the cone. These rank one matrices have only one positive eigenvalue. Any other singular matrix with at least two nonzero eigenvalues is "closer" to the identity matrix than the rank one matrices, but still they are at the boundary of the cone. An additional interesting discussion on the structure of the PSD cone can be found in [4], and more recently for a general setting in [5, 6].

On the other hand, the identity ray represents the best conditioned matrices in the cone, and any well-conditioned matrix should form a "small" angle with it. Moreover, a very ill-conditioned matrix is "almost" singular (near the boundary), and so, it should form a "large" angle with I. Similarly, in that case the inverse matrix is also very ill-conditioned and as a consequence it should also form a "large" angle with I. Is it possible to conclude that an ill-conditioned matrix and its inverse form a "large" angle among them? We believe that the results in this work add understanding for answering this question.

Clearly, all the comments in the last two paragraphs carry a significant geometrical intuition. However, some of them became formal facts in [9] and [10]. The main purpose of this work is to continue the formal study of these geometrical ideas, and specially to get practical bounds, as sharp as possible, for the angle between the inverse of a given matrix and I, and also for the Frobenius condition number.

2 Notation and basic results

Let us denote by S_n the set of symmetric real matrices of order n, and by PD_n the matrices in S_n that are positive definite. Notice that PD_n is the interior of PSD, which is a closed convex cone. We are interested in PD_n, and not in PSD, because our main concern is to study the properties of the Frobenius condition number, which is not defined for singular matrices.

The Frobenius inner product allow us to define the cosine of the angle between two given real $n \times n$ matrices as

$$\cos(A, B) = \frac{< A, B >_F}{\|A\|_F \|B\|_F}.$$

In particular, for a given matrix A in PD_n,

$$\cos(A, I) = \frac{tr(A)}{\|A\|_F \sqrt{n}}, \tag{1}$$

and hence

$$\cos(A, I) \cos(A^{-1}, I) = \frac{tr(A) tr(A^{-1})}{n\, \kappa_F(A)}, \tag{2}$$

where $\kappa_F(A) = \|A\|_F \|A^{-1}\|_F$. Note that for any nonsingular symmetric matrix A, using Cauchy-Schwarz inequality, we have

$$n = tr(I) = < A, A^{-1} >_F \leq \; \|A\|_F \|A^{-1}\|_F = \kappa_F(A), \tag{3}$$

and so, n is a lower bound for $\kappa_F(A)$. Taking advantage of the basic properties of any inner product space, we can characterize the distance between n and $\kappa_F(A)$.

Lemma 2.1 *If A is a nonsingular symmetric matrix, then*

$$\kappa_F(A) = n + \frac{1}{2}\left[\|A - A^{-1}\|_F^2 - (\|A\|_F - \|A^{-1}\|_F)^2\right]. \tag{4}$$

Proof. Since

$$\|A - A^{-1}\|_F^2 = < A, A >_F + < A^{-1}, A^{-1} >_F - 2n,$$

and

$$(\|A\|_F - \|A^{-1}\|_F)^2 = < A, A >_F + < A^{-1}, A^{-1} >_F - 2\kappa_F(A),$$

then

$$\left[\|A - A^{-1}\|_F^2 - (\|A\|_F - \|A^{-1}\|_F)^2\right] = -2n + 2\kappa_F(A),$$

and the result follows. ∎

Notice that for any symmetric matrix such that $A^{-1} = A$ (e.g. Householder orthogonal transformations) we have that $\kappa_F(A) = n$. Moreover, from the triangle inequality it follows that $(\|A\|_F - \|A^{-1}\|_F)^2 \leq \|A - A^{-1}\|_F^2$, and hence (4) is coherent with (3). It is also worth noticing that an equality similar to (4) can hardly be obtained for the classical 2-norm condition number. Nevertheless, recalling the usual notation $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$, it is well-known that

$$\kappa_F(A)/n \leq \kappa_2(A) \leq \kappa_F(A), \tag{5}$$

which combined with Lemma 2.1 yields the following inequality for the 2-norm condition number

$$1 + \frac{1}{2n} D(A, A^{-1}) \leq \kappa_2(A) \leq n + \frac{1}{2} D(A, A^{-1}),$$

where $D(A, A^{-1}) = \left[\|A - A^{-1}\|_F^2 - (\|A\|_F - \|A^{-1}\|_F)^2\right]$.

Using Cauchy-Schwarz inequality again, we can now establish a simple but useful lemma.

Lemma 2.2 *If A is a square matrix in PD_n, then*

$$tr(A) * tr(A^{-1}) \geq n^2. \tag{6}$$

Proof. Since $A \in PD_n$ then the square roots $A^{1/2}$ and $A^{-1/2}$ are well defined, and

$$n^2 = tr^2(I) = < A^{1/2}, A^{-1/2} >_F^2 \leq \|A^{1/2}\|_F^2 \|A^{-1/2}\|_F^2 = tr(A) * tr(A^{-1}). \quad ∎$$

The next result will play an important role in the rest of this note.

Lemma 2.3 *If A is a square matrix in PD_n, then*

$$\frac{1}{\kappa_2(A)} \le \cos(A, A^{-1}) \le \cos(A, I)\ \cos(A^{-1}, I). \tag{7}$$

Proof. Combining (2) and (6) we obtain

$$\cos(A, I)\cos(A^{-1}, I) \ge \frac{n}{\kappa_F(A)}, \tag{8}$$

or equivalently

$$\frac{n}{\kappa_F(A)\cos(A, I)} \le \cos(A^{-1}, I) \le 1. \tag{9}$$

In (9) equality is attained everywhere when $A = I$.

Moreover,

$$\cos(A, A^{-1}) = \frac{tr(I)}{\kappa_F(A)} = \frac{n}{\kappa_F(A)}, \tag{10}$$

that combined with (9), yields

$$\cos(A, A^{-1}) \le \cos(A, A^{-1})/\cos(A, I) \le \cos(A^{-1}, I) \le 1. \tag{11}$$

In here, we use the fact that for any given matrices A and B in PD_n, the maximum possible angle between them is $\pi/2$, i.e., $0 \le \cos(A, B) \le 1$, and so $tr(AB) \ge 0$ (see e.g. Iusem and Seeger [5, 6]). Recalling now that for any A in PD_n

$$\|A\|_2 \le \|A\|_F \le \sqrt{n}\, \|A\|_2,$$

we obtain

$$\kappa_2(A) \le \kappa_F(A) \le n\, \kappa_2(A),$$

and hence, (9) can be written as

$$\frac{1}{\kappa_2(A)} \le \frac{1}{\kappa_2(A)\cos(A, I)} \le \cos(A^{-1}, I) \le 1. \tag{12}$$

Finally, we can conclude from (10), (11) and (12) that

$$\frac{1}{\kappa_2(A)} \le \cos(A, A^{-1}) \le \cos(A, I)\ \cos(A^{-1}, I),$$

and the proof is complete. ∎

Notice that the inequality in Lemma 2.3 is sharp when $A = I$. Notice also that using similar arguments to the ones used to obtain (11) we have

$$\cos(A, A^{-1}) \le \cos(A, I) \le 1,$$

which means, together with (11) , that the angle between A and A^{-1} is larger than the angle between A and I or between A^{-1} and I.

3 Main result

Lemma 2.3 together with the last observations accounts for some of the geometrical properties of the Frobenius condition number. From (5) we know that $\kappa_F(A)$ is equivalent to $\kappa_2(A)$. Nevertheless, $\kappa_F(A)$ can be associated with the geometrical intuition described in Section 1.

In order for this intuitive line of arguments to be formal and complete, we need to study the relationship between $\cos(A^{-1}, I)$ and $\cos(A, I)$. Our next theorem establishes a bound that adds understanding to this relationship.

Theorem 3.1 *If A is a square matrix in PD_n, then*

$$1/\sqrt{n} \leq \frac{\cos(A, I)}{\cos(A^{-1}, I)} \leq \sqrt{n}. \tag{13}$$

Proof. Using (1) for A and A^{-1} and recalling that $tr(A) = \sum_{i=1}^{n} \lambda_i$ and $\|A\|_F^2 = \sum_{i=1}^{n} \lambda_i^2$, we have

$$\cos^2(A, I)/\cos^2(A^{-1}, I) = \frac{(\sum_{i=1}^{n} \lambda_i)^2 (\sum_{i=1}^{n} 1/\lambda_i^2)}{(\sum_{i=1}^{n} 1/\lambda_i)^2 (\sum_{i=1}^{n} \lambda_i^2)},$$

where

$$0 < \lambda_1 \leq \ldots \leq \lambda_n < \infty$$

are the eigenvalues of A.

Consider now the vector $\lambda \in \Re^n$ with entries λ_i, and consider also the vector $y \in \Re^n$ with entries $1/\lambda_i$, for $i = 1, \ldots, n$. Using these two vectors, we can write

$$\cos^2(A, I)/\cos^2(A^{-1}, I) = (\lambda^T e)^2 \, (y^T y)/(\lambda^T \lambda) \, (y^T e)^2, \tag{14}$$

where e is the vector of all ones. Using Cauchy-Schwarz inequality and the fact that $\|x\|_2 \leq \|x\|_1$ for any vector $x \in \Re^n$, it follows that

$$1 \leq \|\lambda\|_1^2/\|\lambda\|_2^2 = (\lambda^T e)^2/\|\lambda\|_2^2 \leq ((\lambda^T \lambda) \, (e^T e))/\|\lambda\|_2^2 = e^T e = n,$$

and also that

$$1 \geq \|y\|_2^2/\|y\|_1^2 = (y^T y)/(y^T e)^2 \geq (y^T y)/(\|y\|_2^2 \, \|e\|_2^2) = 1/(e^T e) = 1/n.$$

Therefore, substituting the last two inequalities in (14) we obtain that

$$1/n \leq \cos^2(A, I)/\cos^2(A^{-1}, I) \leq n,$$

and the result is established. ∎

Combining (13) and (12) we obtain that, for any matrix in PD_n,

$$\frac{1}{\cos(A, I)} \leq \sqrt{n} \cos(A^{-1}, I) \leq \sqrt{n},$$

In here, equality is attained at any extreme ray of the PSD cone, as established by Tarazaga [10]. Moreover, combining Lemma 2.3 and Theorem 3.1 we obtain the following inequality

$$\cos(A, A^{-1}) \leq \sqrt{n}\, \cos^2(A, I),$$

which together with (3) yields the following new practical lower bound for the Frobenius condition number $\kappa_F(A)$

$$\kappa_F(A) \geq \max\left(n, \frac{\sqrt{n}}{\cos^2(A, I)}\right). \tag{15}$$

4 Combining and comparing with other bounds

From Wolkowicz and Styan [11] (Theorem 2.1), we have the following inequalities related to a given matrix A with extreme real eigenvalues $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$

$$m - sp \leq \lambda_{\min}(A) \leq m - \frac{s}{p},$$

and

$$m + \frac{s}{p} \leq \lambda_{\max}(A) \leq m + sp,$$

where $p = \sqrt{(n-1)}$, $m = tr(A)/n$ and $s^2 = (\|A\|_F^2/n) - m^2$. If $m - sp > 0$, and A is positive definite, then the following upper bound is also obtained in [11] (Corollary 2.1)

$$\kappa_2(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \leq \frac{m + sp}{m - sp}.$$

However, even if A is positive definite, $m - sp$ might not be positive. In that case, it is clear that

$$\kappa_2(A) \leq \frac{m + sp}{\lambda_{\min}(A)},$$

which of course requires the knowledge (or an approximation) of $\lambda_{\min}(A)$. Under the additional assumption that $tr(A)^2 > (n-1)tr(A^2)$, some other bounds are obtained in [11]; and assuming that the determinant of A is also available additional upper bounds are established in [7].

On the other direction, for symmetric and positive definite matrices, and without any additional assumption, it is also established in [11] (Corollary 2.3) that

$$\kappa_2(A) \geq 1 + \frac{2s}{m - \frac{s}{p}}. \tag{16}$$

Moreover, If n is odd they proved a slightly sharper result

$$\kappa_2(A) \geq 1 + \frac{2sr}{m - \frac{s}{p}},$$

where $r = n/\sqrt{(n^2-1)}$. Clearly, r tends to 1 when n increases.

Using (5), all these bounds can be combined with the results of the previous sections to produce practical bounds for $\kappa_F(A)$. In particular, combining (5), (15), and the Wolkowicz-Styan (WS) bound (16) we present the following practical new bound

$$\kappa_F(A) \geq \max\left(n, \frac{\sqrt{n}}{\cos^2(A, I)}, (1 + \frac{2s}{m - \frac{s}{p}})\right). \tag{17}$$

Notice, that the computational cost for obtaining (15), (16), or (17) is the same, since only $tr(A)$ and $\|A\|_F$ are required.

We now present some numerical experiments to evaluate and compare the accuracy of the new bounds. In Figure 1 we compare the value of $\kappa_F(A)$ with (15), with the WS bound given by (16), and with (17) for two different matrices B: $B = ee^T + \lambda I$ where e is the vector of all ones in the first case, and $B = RR^T + \lambda I$ where R is a random square matrix in the second case (R is built with the `rand` function in MATLAB). In both cases the matrices are of dimension 500×500.

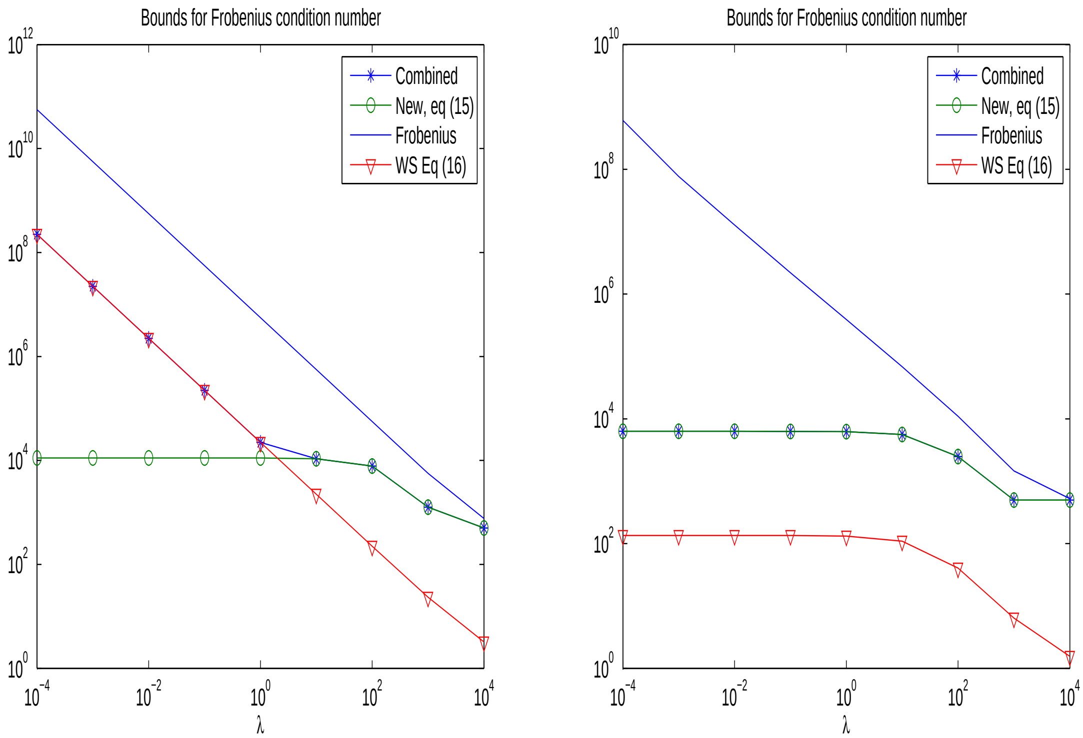

Figure 1: Comparison of different bounds for the Frobeniuis condition number of the 500×500 matrices $B = ee^T + \lambda I$ (left) and of $B = RR^T + \lambda I$ (right) vs lambda in a loglog scale.

In Figure 1 we observe that for matrices that form a small angle with an extreme ray of the PSD cone, the WS bound is sharper than the new bound (15). However, away from an extreme ray of the PSD cone, (15) is sharper than WS. For instances, in the second (right) experiment (15) is sharper than WS for all λ, including the ill-conditioned and also the well-conditioned matrices that are obtained when $\lambda > 0$ increases. Therefore, the best option out of the ones considered here is the combined bound (17) which is obviously a winner at the same computational cost.

5 Final remarks and perspectives

In recent years the art of building inverse preconditioners has received much attention; see [3] for a survey and [8] for a full description of the topic. The geometrical understanding, obtained from the preceding results, can possibly lead to new approaches for building inverse preconditioners.

As a first attempt consider the following equality

$$\begin{aligned} \|I - AQ\|_F^2 &= \|I\|_F^2 + \|AQ\|_F^2 - 2tr(AQ) \\ &= \|I\|_F^2 + \|AQ\|_F^2 - 2\cos(AQ, I)\|AQ\|_F\|I\|_F \\ &= \|I\|_F^2 + \|AQ\|_F^2 - 2\|AQ\|_F\|I\|_F + 2(1 - \cos(AQ, I))\|AQ\|_F\|I\|_F \\ &= (\|I\|_F - \|AQ\|_F)^2 + 2(1 - \cos(AQ, I))\|AQ\|_F\|I\|_F. \end{aligned}$$

If we now replace Q by Q_k, a sequence of matrices converging to A^{-1}, the term $1 - \cos(AQ_k, I)$ measures the weak convergence of AQ_k to I. Indeed the (strong) convergence in any inner product space is typically decomposed into the convergence of the norms $(\|I\|_F - \|AQ_k\|_F)^2$ and the weak convergence $(1 - \cos(AQ_k, I))\|AQ_k\|_F\|I\|_F$, that can be estimated by $1 - \cos(AQ_k, I)$ when $\|Q_k\|_F$ is bounded. In that case we would consider the objective function

$$1 - \cos(AQ, I) = (\|AQ\|_F\sqrt{n} - tr(AQ))/\|AQ\|_F\sqrt{n}$$

and then it would be convenient to study the expression $\|AQ\|_F\sqrt{n} - tr(AQ)$. Of course, to guarantee the convergence of the sequence Q_k we must impose some scaling condition at each step; otherwise we could observe convergence of $\cos(AQ_k, I)$ to 1 while the norm of Q_k grows consistently. Summing up, we could obtain an approximation to the inverse of the matrix A as the (*inexact*) solution of any of the two following problems

- The *weak convergence* problem

$$\min_{Q} \left(\|AQ\|_F\sqrt{n} - tr(AQ)\right)$$

- The *strong convergence* problem

$$\min_{\|AQ\|_F=\sqrt{n}} \left(\|AQ\|_F\sqrt{n} - tr(AQ)\right)$$

The first formulation is an optimization problem without constraint whereas the second one is a constrained problem. The scaling constraint can be relaxed classically by adding a penalization term or, in an iterative process, by scaling Q_k at each iteration.

Finally, it is also worth mentioning that if the matrix A is not available but matrix-vector products with it are easy to evaluate, we can still recover all the practical bounds previously described. For example, if the products Ae_i are computed for all canonical vectors e_i, then $tr(A)$ can be recovered as follows

$$tr(A) = \sum_{i=1}^{n} e_i^T A e_i,$$

and

$$\|A\|_F^2 = tr(A^2) = \sum_{i=1}^{n} (Ae_i)^T A e_i.$$

Unfortunately, if n is large, these formulas are of limited practical use. Additional ideas for computing traces for large scale problems, without the explicit knowledge of A, can be found in [1, 2].

References

[1] Z. Bai, M. Fahey and G. Golub [1996], Some large-scale matrix computation problems, *Journal of Computational and Applied Mathematics*, 74, 71–89.

[2] C. Bekas, E. Kokiopoulou and Y. Saad [2006], An estimator for the diagonal of a matrix, *Applied Numerical Mathematics*, to appear.

[3] M. Benzi [2002], Preconditioning techniques for large linear systems: A survey, *Journal of Computational Physics*, 182, 418–477.

[4] R. D. Hill and S. R. Waters [1987], On the cone of positive semidefinite matrices, *Linear Algebra and its Applications*, 90, 81–88.

[5] A. N. Iusem and A. Seeger [2006], Measuring the degree of pointedness of a closed convex cone: A metric approach, *Math. Nachr.*, 279, 599–618.

[6] A. N. Iusem and A. Seeger [2005], On pairs of vectors achieving the maximal angle of a convex cone, *Math. Program.*, Ser. B, 104, 501–523.

[7] J. K. Merikoski, U. Urpala, A. Virtanen, T.-Y. Tam and F. Uhlig [1997], A best upper bound for the 2-norm condition number of a matrix, *Linear Algebra and its Applications*, 254, 355–365.

[8] Y. Saad [2003], *Iterative Methods for Sparse Linear Systems (2nd. Edition)*, SIAM, Philadelphia.

[9] P. Tarazaga [1990], Eigenvalue estimates for symmetric matrices, *Linear Algebra and its Applications*, 135, 171–179.

[10] P. Tarazaga [1991], More estimates for eigenvalues and singular values, *Linear Algebra and its Applications*, 149, 97–110.

[11] H. Wolkowicz and G. P. H. Styan [1980], Bounds for eigenvalues using traces, *Linear Algebra and its Applications*, 29, 471–506.