J. M. Martínez

A bound for the Moore-Penrose pseudoinverse of a matrix

Commentationes Mathematicae Universitatis Carolinae, Vol. 20 (1979), No. 2, 357--360

Persistent URL: `http://dml.cz/dmlcz/105934`

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE
20,2 (1979)

A BOUND FOR THE MOORE-PENROSE PSEUDOINVERSE OF A MATRIX

J. M. MARTÍNEZ

Abstract: A geometric bound is obtained for the norm of $(A^tA)^{-1}A^t$, when A is an m x n matrix of full rank with $m \geq n$. Hence, a similar bound holds for the Moore-Penrose pseudoinverse of any m x n matrix, with $m \geq n$. The new bound gives a geometrical meaning to the well-known relation between condition number, scaling and angle between columns.

Key words: Norm of a matrix, Moore-Penrose pseudoinverse, Condition number.

AMS: 65F20, 65F35, 15A09, 15A12

Notation. $[v_1,\dots,v_p]$ will denote the subspace spanned by the vectors $v_1,\dots,v_p$, and $[v_1,\dots v_p]^{\perp}$ its orthogonal complement. $\|\cdot\|$ will always be any norm, unless specified.

Lemma 1. Let A be a real n x n matrix, $A = (a_1,\dots,a_n)$ and let α_1 be equal to $\pi/2$ and α_j, $j = 2,\dots,n$ the angle between a_j and $[a_1,\dots,a_{j-1}]$. Then,

$$|\det A| = \prod_{i=1}^{n} \|a_i\|_2 \, |\sin \alpha_i|.$$

Proof. See [2].

Lemma 2. Let A be a real m x n matrix of full rank with $m \geq n$; $A = (a_1,\dots,a_n)$; and define $\alpha_j = \alpha_j(A)$ as in Lemma 1 for $j = 1,\dots,n$. Define $P(A) = \prod_{i=1}^{n} |\sin \alpha_i|$. Then

$P(A)$ is invariant under permutations of the columns of A.

Proof. If $m = n$ the thesis is true because of Lemma 1. Suppose $m > n$ and define $A' = (a_1,\dots,a_n,a_{n+1},\dots,a_m)$, where $\|a_i\|_2 = 1$, $\langle a_i,a_j\rangle = 0$ if $i \neq j$, $i, j = n+1,\dots,m$, and $[a_{n+1},\dots,a_m] = [a_1,\dots,a_n]^{\perp}$. Then $P(A') = P(A)$. But $P(A')$ is invariant under permutations of the columns of A'; so the same holds for A.

Lemma 3. Let A be as in Lemma 2, and let $\beta_i = \beta_i(A)$, $i = 1,\dots,n$, be the angle between a_i and $[a_1,\dots,a_{i-1},a_{i+1},\dots$ $\dots,a_n]$. Then $|\sin \beta_i| \geq P(A)$.

Proof. Define $A' = (a_1,\dots,a_{i-1},a_{i+1},\dots,a_n,a_i)$. Then $\beta_i(A) = \alpha_n(A')$ and so, $|\sin \beta_i(A)| = |\sin \alpha_n(A')| \geq$ $\geq P(A') = P(A)$.

Lemma 4. Let A be as in Lemma 2, and define $A^+ =$ $= (A^tA)^{-1}A^t = (b_1,\dots,b_n)^t$. Then $\|b_i\|_2 \leq 1/(P(A)\,\|a_i\|_2)$ for all $i = 1,\dots,n$.

Proof. $A^+A = I$ implies that $b_i \in [a_1,\dots,a_{i-1},a_{i+1},\dots$ $\dots,a_n]^{\perp}$ and $\langle a_i,b_i\rangle = 1$. Then, $\|a_i\|_2\,\|b_i\|_2 \cos \gamma_i = 1$, where γ_i is the angle between a_i and b_i. But $A^+ = (A^tA)^{-1}A^t$ implies that $b_i \in [a_1,\dots,a_n]$. Then $\gamma_i = \pi/2 - \beta_i$, with β_i defined as in Lemma 3; and so, $\|b_i\|_2 = 1/(\|a_i\|_2\,|\sin \beta_i|) \leq$ $\leq 1/(\|a_i\|_2\, P(A))$.

Theorem 1. Let $\|\cdot\|$ be a norm in $R^{m \times n}$. Then there exists $K > 0$, $K = K(m,n)$ such that for all A with the hypotheses of Lemma 4,

$$\|A^+\| \leq K \max\{1/\|a_i\|_2,\ i = 1,\dots,n\}/P(A).$$

Proof. It follows immediately from Lemma 4.

Theorem 2. Let A be a real m x n matrix of rank p with $m \geq n$. Suppose A = (B,C), where rank B = p; and let A be the Moore-Penrose pseudoinverse of A (see [3]). Then there exists K = K(m,p) such that

$$\|A^+\| \leq K \max\{1/\|a_i\|_2,\ i = 1,\dots,p\} /P(B).$$

Proof. Define $A' = \binom{B^+}{0}$. Then, $A'b$ is a solution of the least - squares problem $Ax \cong b$ for all $b \in R^m$. Then $\|A^+b\|_2 \leq \|A'b\|_2$ for all $b \in R^m$. Thus $\|A^+\|_2 \leq \|A'\|_2$, and the thesis follows easily from this inequality.

Final remarks.

a) If k(A) is the condition number of an n x n nonsingular matrix (see [1]), then it follows from Theorem 1 that

$$k(A) \leq K \max\{\|a_i\|_2,\ i = 1,\dots,n\} \max\{1/\|a_i\|_2,\ i = 1,\dots \\ \dots,n\} /P(A).$$

This is an interesting inequality which shows that when the condition number grows, then either the matrix is not "well scaled" or the columns of A are nearly dependent.

b) The sharpness of the bounds on Theorems 1 and 2 depends on the sharpness of the inequalities $|\sin \beta_i| \geq P(A)$ in Lemma 3. If more than one column is nearly dependent from the other columns, it may happen that $|\sin \beta_i| >> P(A)$.

c) We may, mutatis mutandi, reformulate the results of this section for full rank matrices $A \in R^{m\times n}$, with $m \leq n$ and A^d (right inverse) $= A^t(AA^t)^{-1}$.

References

[1] G. FORSYTHE and V.B. MOLER (1967): Computer solution of

linear algebraic systems, Prentice Hall, Englewood Cliffs, N.J.

[2] J.M. MARTINEZ (1978): On the order of convergence of Broyden-Gay-Schnabel's method, Comment. Math. Univ. Carolinae 19, 107-118.

[3] C.R. RAO and S.K. MITRA (1971): Generalized Inverse Matrices and its Applications, Wiley, New York.

IMECC - UNICAMP
CP 1170
13100 Campinas, SP
Brazil

(Oblatum 23.1. 1979)