NEARNESS OF MATRICES TO SINGULARITY

by

J. Rohn, Prague

Abstract. A measure of nearness of real matrices to singularity is introduced and described. The proof employs a characterization of singular interval matrices of a special type.

AMS Subject Classifications: 65F25, 65G10

Let $A$ be an $n \times n$ real matrix. The number

$$d(A) = \min \left\{ \|B-A\| \ ; B \text{ singular} \right\}, \quad (1)$$

where

$$\|A\| = \max_{i,j} |A_{i,j}|, \quad (2)$$

can be considered a measure of nearness of $A$ to singularity. The value of $d(A)$ was investigated by Kahan [1] for matrix norms induced by some vector norms. His result, however, cannot be applied to the norm (2) which seems to be natural in the context, since then $d(A)$ expresses the minimum deviation of coefficients which transforms $A$ to a singular matrix.

If $A$ is singular, then $d(A) = 0$; therefore we may restrict our attention only to nonsingular matrices. We shall give formulae for $d(A)$ and for the nearest singular matrix, based on a characterization of singular interval matrices of a special type. For $n \geq 1$, let $Y_n = \{y \in \mathbb{R}^n \ ; \ |y_j| = 1 \text{ for } j = 1, \ldots, n\}.$

Theorem. Let $A$ be a nonsingular $n \times n$ matrix. Then there holds

$$d(A) = \frac{1}{r(A)},$$

where
\[ r(A) = \max \{ z^T A^{-1} y \mid z, y \in Y_n \} \quad (3) \]

If \( z, y \) are vectors from \( Y_n \) for which the maximum is achieved in (3), then

\[ A_0 = A - \frac{1}{r(A)} \frac{y^T}{z^T} z \quad (4) \]

is a singular matrix nearest to \( A \) and the vector

\[ x_0 = A^{-1} y \quad (5) \]

satisfies \( A_0 x_0 = 0 \).

**Proof.** Denote \( e = (1, 1, \ldots, 1)^T \in \mathbb{R}^n \). Let \( B \) be a singular \( n \times n \) matrix. Put \( \beta = \|B - A\| \). Then \( B \) belongs to the interval matrix \([A - \beta ee^T, A + \beta ee^T]\), hence \([A - \beta ee^T, A + \beta ee^T]\) is singular and using the lemma in [2], we get that there exist \( z, y \in Y_n \) such that

\[ \beta z^T A^{-1} y \geq 1 \]

holds. Hence also \( \beta r(A) \geq 1 \), and since \( B \) was an arbitrary singular matrix, we get \( d(A) \geq \frac{1}{r(A)} \). On the other hand, for \( A_0, x_0 \) given by (4), (5), a direct computation gives \( A_0 x_0 = 0 \), hence \( A_0 \) is singular and \( \|A_0 - A\| = \frac{1}{r(A)} \); therefore \( d(A) = \frac{1}{r(A)} \).

Unfortunately, the value of \( r(A) \) is not easy to compute. However, there exists a class of matrices for which \( r(A) \) can be expressed explicitly:

**Corollary.** Let \( A \) be a nonsingular \( n \times n \) matrix for which there exist \( z, y \in Y_n \) such that

\[ z_i A_{1}^{-1} y_j \geq 0 \quad (i, j = 1, \ldots, n) \quad (6) \]

holds. Then \( d(A) = \frac{1}{r(A)} \), where

\[ r(A) = \sum_{i, j} |A_{1}^{-1}_{i j}| \]
Proof. Under the assumption, we have $Z^T\Lambda^{-1}y \leq r(\Lambda) \leq \sum_{i,j} |A_{i,j}| = \sum_{i,j} \tilde{z}_i \tilde{A}_{i,j} \tilde{y}_j = z^T\Lambda^{-1}y$, hence $r(\Lambda) = \sum_{i,j} |A_{i,j}^{-1}|$. \hfill \blacksquare

Especially, for inverse nonnegative matrices (where $\Lambda^{-1} \geq 0$, so that (6) is satisfied with $Z = \tilde{y} = e$) we get that $r(\Lambda) = \sum_{i,j} \Lambda_{i,j}^{-1}$ and the nearest singular matrix can be obtained by subtracting the value of $r(\Lambda)^{-1}$ from all coefficients of $\Lambda$.

Let us now return to the general case. If the maximum in (3) is achieved at some $z, y \in Y_n$, then, since $r(\Lambda) = z^T\Lambda^{-1}y = \sum_i z_i (\Lambda^{-1}y)_i = \sum_j (z^T\Lambda^{-1})_j y_j$, there must hold

\begin{equation}
    z_i (\Lambda^{-1}y)_i \geq 0 \quad \text{for } i = 1, \ldots, n
\end{equation}

and

\begin{equation}
    (z^T\Lambda^{-1})_j y_j \geq 0 \quad \text{for } j = 1, \ldots, n
\end{equation}

for otherwise the value of $z^T\Lambda^{-1}y$ could be increased. Thus, using the vector norm $\|x\|_1 = \sum_i |x_i|$, we may also write

\begin{equation*}
    r(\Lambda) = \max \left\{ \|A^{-1}y\|_1 : y \in Y_n \right\}
\end{equation*}

If $n$ is large, then $r(\Lambda)$ cannot be computed in this way since $Y_n$ has $2^N$ elements. In this case, we propose the following algorithm which stops after reaching a pseudo-optimal solution satisfying the necessary optimality conditions (7) and (8):

\textbf{Algorithm.}

0. Select $z, y \in Y_n$.
1. Set $z_i := -z_i$ for each $i$ with $z_i (\Lambda^{-1}y)_i < 0$.
2. Set $y_j := -y_j$ for each $j$ with $(z^T\Lambda^{-1})_j y_j < 0$.
3. If (7) holds, terminate. Otherwise go to step 1.

The algorithm is finite since $Y_n$ is finite and the value of $z^T\Lambda^{-1}y$ is always increased during step 1 or 2, so that cycling cannot occur. The condition (8) is always satisfied after
step 2, hence it need not be tested in step 3. If the algorithm terminates with some \( z, y \in Y_n \) in step 3, then

\[
d(A) \leq \frac{1}{z A^{-1} y}
\]

and the matrix

\[ A_0^* = A - \frac{yz^T}{z A^{-1} y} \]

is a singular matrix with \( \| A_0^* - A \| = \frac{1}{z A^{-1} y} \) and

\[ A_0^* x_0^* = 0 \text{ for } x_0^* = A^{-1} y . \]

It is perhaps worth mentioning that according to (4), each square nonsingular matrix \( A \) can be decomposed as \( A = A_0 + A_1 \), where \( A_0 \) is singular and \( A_1 \) is of rank one. Also,

\[ \| A^{-1} \| \geq \frac{1}{n^2 d(A)} \text{ for each nonsingular } n \times n \text{ matrix } A . \]

References


Author's address: J. Rohn, Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 118 00 Prague, Czechoslovakia