ON BEST CONDITIONED MATRICES

G. E. FORSYTHE AND E. G. STRAUS

1. Main theorems. Let $A$ be a positive definite Hermitian matrix of finite order, and let $\Lambda$ and $\lambda$ be its maximal and minimal eigenvalue respectively. The condition number of $A$ is the ratio $P(A) = \frac{\Lambda}{\lambda}$ introduced by Todd [1]. Let $\mathcal{C}$ be a class of regular linear transformations. Define $A^T = T^*AT$. We say that $A$ is best conditioned with respect to $\mathcal{C}$ if $P(A^T) \geq P(A)$ for all $T \in \mathcal{C}$.

In order to investigate whether $A$ is best conditioned we remember that

\[
\Lambda = \max_{x} \frac{x^*Ax}{x^*x}, \quad \lambda = \min_{x} \frac{x^*Ax}{x^*x}
\]

and hence

\[
P(A) = \max \frac{x^*Ax}{|x|^2}.
\]

We introduce the abbreviation $R = R(T) = (T^*)^{-1}T^{-1}$. Now let $\Lambda^T, \lambda^T$ be the extremal eigenvalues of $A^T$. Setting $u = Tx$, we obtain from (1) and (2):

\[
\Lambda^T = \max_{x} \frac{x^*T^*ATx}{x^*x} = \max_{u} \frac{u^*Au}{u^*Ru},
\]

\[
\lambda^T = \min_{x} \frac{x^*T^*ATx}{x^*x} = \min_{u} \frac{u^*Au}{u^*Ru},
\]

\[
P(A^T) = \max \frac{u^*Au}{v^*Rv}.
\]

Thus, if we let $S_\Lambda, S_\lambda$ be the sets of unit eigenvectors of $A$ belonging to $\Lambda$ and $\lambda$ respectively, then we obtain from (2)

\[
P(A) = \max \frac{x^*Ax}{y^*Ay}, \quad x \in S_\Lambda, \ y \in S_\lambda.
\]

Presented to the International Congress of Mathematicians, September 8, 1954; received by the editors August 13, 1954.

1 The first author wishes to acknowledge the sponsorship of the Office of Naval Research—at the beginning of this research through the National Bureau of Standards, Los Angeles, and now in connection with Project ONR 044-144. The authors wish to acknowledge helpful suggestions by their colleague Dr. T. S. Motzkin.
Hence from (3) and (4)

\[ P(A^T) \geq P(A) \max_{u \in S_A, v \in S_\lambda} \frac{v^*Rv}{u^*Ru}. \]

We thus have proved:

**Lemma.** If \( \max_{u \in S_A, v \in S_\lambda} (v^*Rv/u^*Ru) \geq 1 \) for all \( T \in \mathcal{T} \), then \( A \) is best conditioned with respect to \( \mathcal{T} \).

It will be convenient to introduce the concept of "separability by \( \mathcal{T} \)"

**Definition.** The sets \( S_1, S_2 \) are separable by \( \mathcal{T} \) if there exists a \( T \in \mathcal{T} \) and a constant \( k \) so that

\[ x^*Rx < k < y^*Ry, \]

for all \( x \) in one \( S_j \) and all \( y \) in the other.

Obviously, if \( S_1, S_2 \) are not separable by \( \mathcal{T} \), then

\[ \sup_{x \in S_j, y \in S_k} \frac{x^*Rx}{y^*Ry} \geq 1 \]

for all \( T \in \mathcal{T} \).

Combining (5) with the lemma, we have proved

**Theorem 1.** If \( S_\lambda \) and \( S_\Lambda \) are not separable by \( \mathcal{T} \), then \( A \) is best conditioned with respect to \( \mathcal{T} \).

The converse to Theorem 1 is not true without further conditions on \( \mathcal{T} \). As such a condition we introduce the following concept:

**Definition.** A set \( \mathcal{T} \) of regular linear transformations is called infinitesimally complete if, for every \( T \in \mathcal{T} \), there exist arbitrarily small positive \( \epsilon, \epsilon' \) such that there are \( T_\epsilon, T_{\epsilon'} \in \mathcal{T} \) with

\[ I + \epsilon R = c(T_\epsilon^*)^{-1}T_\epsilon^{-1}, \quad I - \epsilon' R = c'(T_{\epsilon'}^*)^{-1}T_{\epsilon'}^{-1}, \]

where \( c, c' \) are (positive) numbers.

**Theorem 2.** If \( \mathcal{T} \) is infinitesimally complete and \( S_\lambda, S_\Lambda \) are separable by \( \mathcal{T} \), then \( A \) is not best conditioned with respect to \( \mathcal{T} \).

**Proof.** By the hypothesis of separability there exists a \( T \in \mathcal{T} \) and a \( k > 0 \) such that either

(I) \[ x^*Rx > k > y^*Ry \quad \text{for all } x \in S_\lambda, y \in S_\lambda, \]

or

(II) \[ x^*Rx < k < y^*Ry \quad \text{for all } x \in S_\lambda, y \in S_\lambda. \]
In case (I) we have \( y^* R y / x^* R x < 1 \) for all \( x \in S_\lambda, y \in S_\lambda \). Hence there exist neighborhoods \( U_\lambda, U_\lambda \) of \( S_\lambda, S_\lambda \) on the unit sphere \( S \) so that for every \( \varepsilon > 0 \)

\[
\sup_{x \in U_\lambda, y \in U_\lambda} \frac{y^* (I + \varepsilon R) y}{x^* (I + \varepsilon R) x} < 1.
\]

Define \( U \) to be the Cartesian product \( U_\lambda \times U_\lambda \). Then

\[
\left( \max_{(x,y) \in U} \frac{x^* A x}{y^* A y} \right) \cdot \left( \sup_{x \in U_\lambda} \frac{y^* (I + \varepsilon R) y}{x^* (I + \varepsilon R) x} \right) < P(A).
\]  

Let \( F = S \times S - U \). Then \( \max_{(x,y) \in F} (x^* A x / y^* A y) < P(A) \). Hence we may fix \( \varepsilon \) so small that

\[
\left( \max_{(x,y) \in F} \frac{x^* A x}{y^* A y} \right) \cdot \left( \max_{(x,y) \in F} \frac{y^* (I + \varepsilon R) y}{x^* (I + \varepsilon R) x} \right) < P(A).
\]

By the infinitesimal completeness of \( \mathcal{C} \), there is a \( T_\varepsilon \in \mathcal{C} \) such that

\[
\frac{y^* R_\varepsilon y}{x^* R_\varepsilon x} = \frac{y^* (I + \varepsilon R) y}{x^* (I + \varepsilon R) x}, \quad \text{where } R_\varepsilon = R(T_\varepsilon).
\]

Putting (8) into (6) and (7), we then see from (3) that

\[
P(A T_\varepsilon) < P(A).
\]

The proof of (9) in Case (II) is entirely analogous, if we replace \( I + \varepsilon R \) in (6), (7), (8) by \( I - \varepsilon' R \). But (9) proves the theorem.

2. Applications. As examples of infinitesimally complete classes \( \mathcal{C} \) we may cite:

(i) Quasidiagonal matrices. These are matrices of form

\[
T = \begin{bmatrix}
M_1 & 0 & \cdots & 0 \\
0 & M_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & M_s
\end{bmatrix},
\]

where each \( M_i \) is a square matrix of arbitrary preassigned order. An important subclass is the following:

(ii) Diagonal (or real diagonal or positive diagonal) matrices \( D \). Here forming \( D^* A D \) is a special case of the common practice of preconditioning \( A \) by scaling rows and columns. This is used, for example, to make \( A \) more easily invertible by a numerical process. For numerical operations on a general nonsingular matrix \( C \) the condition
number \( P(A) \), where \( A = CC^* \), is often significant. Preconditioning of 
\( C \) by scaling rows alone yields a matrix \( C_1 = D^*C \), for which \( C_1C_1^* = D^*AD \). Minimizing \( P(D^*AD) \) (at least approximately) thus has 
practical importance for both Hermitian and general matrices.

If \( D \) is a regular diagonal matrix, then \( x^*Rx \) assumes the particularly 
simple form

\[
x^*Rx = x^*(D^*)^{-1}D^{-1}x = \sum_{i=1}^{n} |d_{ii}|^{-2} |x_i|^2.
\]

Thus separability by class (ii) means \( S_\Lambda \) and \( S_\lambda \) can be separated by 
an axis-oriented, origin-centered ellipsoid. From (10) we can establish

**Theorem 3.** A sufficient condition for \( A \) to be best conditioned with 
respect to class (ii) is that, for some pair of eigenvectors \( x^\Lambda, x^\lambda \) belonging 
to \( \Lambda, \lambda, \)

\[
|x_i^\Lambda| = |x_i^\lambda| \quad (i = 1, \ldots, n).
\]

Moreover, if \( \Lambda, \lambda \) are simple eigenvalues, (11) is also necessary.

**Proof.** If (11) holds, then, by (10), \( x^*Rx \) assumes the same value 
for both \( x^\Lambda \) and \( x^\lambda \). The sufficiency then follows from Theorem 1. On 
the other hand, if \( \Lambda \) and \( \lambda \) are simple then \( S_\Lambda, S_\lambda \) consist of two points 
each. We then see that (11) is necessary and sufficient for separability 
of \( S_\Lambda \) and \( S_\lambda \). This proves the necessity.

Note that (11) says that \( x^\Lambda, x^\lambda \) are reflections of each other in some 
coordinate subspace.

When \( \Lambda \) or \( \lambda \) is multiple, there are inseparable \( S_\Lambda, S_\lambda \) containing 
no \( x^\Lambda, x^\lambda \) which are reflections of each other. For an example of this, 
let

\[
A = \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix}, \quad 0 < \alpha < 1.
\]

Here \( \Lambda = 1 + 2\alpha \); \( S_\Lambda \) consists of two points \( \pm P \), where \( P = (1, 1, 1)/3^{1/2} \). Also \( \lambda = 1 - \alpha \) (double root), and \( S_\lambda \) is the circle 
\( x + y + z = 0, \quad x^2 + y^2 + z^2 = 1 \). Now we show that it is impossible to 
separate \( S_\lambda \) from \( S_\Lambda \) by any quadratic surface \( f(x, y, z) = ax^2 + by^2 
+ cz^2 = d \). First, \( f(P) = (a + b + c)/3 \). Let \( r = 1/2^{1/2} \). Take, on \( S_\lambda, P_1 
= (r, -r, 0), \quad P_2 = (r, 0, -r), \quad \text{and} \quad P_3 = (0, r, -r) \). Then \( f(P_1) 
= (a + b)/2, \quad f(P_2) = (a + c)/2 \) and \( f(P_3) = (b + c)/2 \). Hence \( f(P) = [f(P_1) 
+ f(P_2) + f(P_3)]/3, \) and \( f(P) \) must lie between the extreme values of 
the \( f(P_i) \).
Theorem 3 will be applied to prove a conjecture of Young [2]. The conjecture is significant for an iterative solution of certain systems of linear equations.

**Theorem 4.** A positive definite Hermitian matrix of form

\[
Q = \begin{bmatrix}
I_p & B \\
B^* & I_q
\end{bmatrix},
\]

where \(I_p, I_q\) are unit matrices, is always best conditioned with respect to class (ii).

**Proof.** Let \(r\) be the rank of \(B\). The semidefinite matrix \(B^*B\) has exactly \(r\) positive eigenvalues \(\nu_i^2\), which we number so that \(0 < \nu_1^2 \leq \cdots \leq \nu_r^2\). Let \(B^*B y_i = \nu_i^2 y_i\). One finds that the partitioned vectors \((B y_i, \pm \nu_i y_i)\) are \(2r\) linearly independent eigenvectors of \(Q\) belonging to the \(2r\) eigenvalues \(1 \pm \nu_i\) \((i = 1, \ldots, r)\). Since \(Q\) is definite, all \(\nu_i < 1\).

If \(p - r > 0\), there are \(p - r\) linearly independent vectors \(u_j\) with \(B^*u_j = 0\). Then \((u_j, 0)\) are \(p - r\) linearly independent eigenvectors of \(Q\) belonging to the eigenvalue 1. Similarly, if \(q - r > 0\), there are \(q - r\) linearly independent eigenvectors of \(Q\) of type \((0, \nu_k)\), which all belong to the eigenvalue 1. Here all \(Bv_k = 0\).

We have found all \(p + q\) eigenvalues of \(Q\), and see that the largest is \(\Lambda = 1 + \nu_r\), with an eigenvector \((B y_r, \nu_r y_r)\). The smallest is \(\lambda = 1 - \nu_r\), with an eigenvector \((B y_r, -\nu_r y_r)\). Theorem 3 then completes the proof.

For any scalar \(c\), \(P(cD^*QD) = P(D^*QD)\). It would be interesting to know, for the \(Q\) of (12), when the class \(cQ\) contains all the best conditioned transforms \(D^*QD\). These transforms essentially constitute the matrices with Young's Property A [3], often encountered in the numerical solution of partial differential equations.

We can show that the partitioned positive definite matrices

\[
\begin{bmatrix}
c_1 I_p & B \\
B^* & c_2 I_q
\end{bmatrix}
\]

are best conditioned if and only if \(c_1 = c_2\). On the other hand, the third order matrices

\[
Q = \begin{bmatrix}
1 & 0 & b \\
0 & d & 0 \\
b & 0 & 1
\end{bmatrix},
\]
where \(|b| < 1\) and \(1 - |b| \leq d \leq 1 + |b|\), are all best conditioned, with \(P(Q) = (1 + |b|) \cdot (1 - |b|)^{-1}\). We conjecture that, for \(Q\) as in (12), any best conditioned matrix \(D^*QD \neq cQ\) has the form

\[
\begin{bmatrix}
  cI_{p_1} & 0 & B_1 \\
  0 & D_{p_2} & 0 \\
  B_1^* & 0 & cI_q
\end{bmatrix}
\]

where \(D_{p_2}\) is diagonal and \(\Pi\) is a permutation matrix.

**References**


**University of California, Los Angeles**