ON BEST CONDITIONED MATRICES[1]

G. E. FORSYTHE AND E. G. STRAUS

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

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

$$(1)\qquad \Lambda = \max_x \frac{x^*Ax}{x^*x}, \qquad \lambda = \min_x \frac{x^*Ax}{x^*x}$$

and hence

$$(2)\qquad P(A) = \max_{\|x\|=\|y\|=1} \frac{x^*Ax}{y^*Ay}\cdot$$

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

$$\Lambda^T = \max_x \frac{x^*T^*ATx}{x^*x} = \max_u \frac{u^*Au}{u^*Ru},$$

$$\lambda^T = \min_x \frac{x^*T^*ATx}{x^*x} = \min_u \frac{u^*Au}{u^*Ru},$$

$$(3)\qquad P(A^T) = \max_{\|u\|=\|v\|=1} \frac{u^*Au}{v^*Av}\cdot\frac{v^*Rv}{u^*Ru}\cdot$$

Thus, if we let S_Λ, S_λ be the sets of unit eigenvectors of A belonging to Λ and λ respectively, then we obtain from (2)

$$(4)\qquad P(A) = \frac{x^*Ax}{y^*Ay}, \qquad x\in S_\Lambda,\ y\in S_\lambda.$$

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)

$$P(A^T) \geqq P(A) \max_{u \in S_\Lambda, v \in S_\lambda} \frac{v^*Rv}{u^*Ru}.$$

We thus have proved:

LEMMA. *If* $\max_{u \in S_\Lambda, v \in S_\lambda} (v^*Rv/u^*Ru) \geqq 1$ *for all* $T \in \mathfrak{T}$, *then* A *is best conditioned with respect to* $\mathfrak{T}$.

It will be convenient to introduce the concept of "separability by $\mathfrak{T}$":

DEFINITION. The sets S_1, S_2 are *separable by* $\mathfrak{T}$ if there exists a $T \in \mathfrak{T}$ 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 $\mathfrak{T}$, then

$$\sup_{x \in S_1, y \in S_2} \frac{x^*Rx}{y^*Ry} \geqq 1 \qquad \text{for all } T \in \mathfrak{T}. \tag{5}$$

Combining (5) with the lemma, we have proved

THEOREM 1. *If* S_Λ *and* S_λ *are not separable by* $\mathfrak{T}$, *then* A *is best conditioned with respect to* $\mathfrak{T}$.

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

DEFINITION. A set $\mathfrak{T}$ of regular linear transformations is called *infinitesimally complete* if, for every $T \in \mathfrak{T}$, there exist arbitrarily small positive ϵ, ϵ' such that there are T_ϵ, $T_{\epsilon'} \in \mathfrak{T}$ with

$$I + \epsilon R = c(T_\epsilon^*)^{-1} T_\epsilon^{-1}, \qquad I - \epsilon' R = c'(T_{\epsilon'}^*)^{-1} T_{\epsilon'}^{-1},$$

where c, c' are (positive) numbers.

THEOREM 2. *If* $\mathfrak{T}$ *is infinitesimally complete and* S_λ, S_Λ *are separable by* $\mathfrak{T}$, *then* A *is not best conditioned with respect to* $\mathfrak{T}$.

PROOF. By the hypothesis of separability there exists a $T \in \mathfrak{T}$ and a $k > 0$ such that either

$$x^*Rx > k > y^*Ry \qquad \text{for all } x \in S_\Lambda,\ y \in S_\lambda, \tag{I}$$

or

$$x^*Rx < k < y^*Ry \qquad \text{for all } x \in S_\Lambda,\ y \in S_\lambda. \tag{II}$$

In case (I) we have $y^*Ry/x^*Rx<1$ for all $x\in S_\Lambda$, $y\in S_\lambda$. Hence there exist neighborhoods U_Λ, U_λ of S_Λ, S_λ on the unit sphere S so that for every $\epsilon>0$

$$\sup_{x\in U_\Lambda, y\in U_\lambda} \frac{y^*(I+\epsilon R)y}{x^*(I+\epsilon R)x} < 1.$$

Define U to be the Cartesian product $U_\Lambda \times U_\lambda$. Then

$$\left(\max_{(x,y)\in U} \frac{x^*Ax}{y^*Ay}\right)\cdot\left(\sup_{(x,y)\in U} \frac{y^*(I+\epsilon R)y}{x^*(I+\epsilon R)x}\right) < P(A). \tag{6}$$

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

$$\left(\max_{(x,y)\in F} \frac{x^*Ax}{y^*Ay}\right)\cdot\left(\max_{(x,y)\in F} \frac{y^*(I+\epsilon R)y}{x^*(I+\epsilon R)x}\right) < P(A). \tag{7}$$

By the infinitesimal completeness of $\mathfrak{T}$, there is a $T_\epsilon\in\mathfrak{T}$ such that

$$\frac{y^*R_\epsilon y}{x^*R_\epsilon x} = \frac{y^*(I+\epsilon R)y}{x^*(I+\epsilon R)x}, \qquad \text{where } R_\epsilon = R(T_\epsilon). \tag{8}$$

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

$$P(A^{T_\epsilon}) < P(A). \tag{9}$$

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

2. **Applications.** As examples of infinitesimally complete classes $\mathfrak{T}$ we may cite:

(i) *Quasidiagonal* matrices. These are matrices of form

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

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^*AD 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

$$x^*Rx = x^*(D^*)^{-1}D^{-1}x = \sum_{i=1}^{n} | d_{ii}|^{-2} | x_i |^2. \tag{10}$$

Thus separability by class (ii) means S_Λ and S_λ 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^Λ, x^λ belonging to* Λ, λ,

$$| \overset{\Lambda}{x_i} | = | \overset{\lambda}{x_i} | \qquad (i = 1, \cdots, n). \tag{11}$$

Moreover, if Λ, λ *are simple eigenvalues,* (11) *is also necessary.*

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

Note that (11) says that x^Λ, x^λ *are reflections of each other in some coordinate subspace.*

When Λ or λ is multiple, there are inseparable S_Λ, S_λ containing no x^Λ, x^λ which are reflections of each other. For an example of this, let

$$A = \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix}, \qquad 0 < \alpha < 1.$$

Here $\Lambda=1+2\alpha$; S_Λ consists of two points $\pm P$, where $P =(1, 1, 1)/3^{1/2}$. Also $\lambda=1-\alpha$ (double root), and S_λ is the circle $x+y+z=0$, $x^2+y^2+z^2=1$. Now we show that it is impossible to separate S_λ from S_Λ 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_λ, $P_1 =(r, -r, 0)$, $P_2=(r, 0, -r)$, and $P_3=(0, r, -r)$. Then $f(P_1) =(a+b)/2$, $f(P_2)=(a+c)/2$ 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*

$$Q = \begin{bmatrix} I_p & B \\ B^* & I_q \end{bmatrix}, \tag{12}$$

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 ν_i^2, which we number so that $0<\nu_1^2 \leqq \cdots \leqq \nu_r^2$. Let $B^*By_i=\nu_i^2 y_i$. One finds that the partitioned vectors $(By_i, \pm\nu_i y_i)$ are $2r$ linearly independent eigenvectors of Q belonging to the $2r$ eigenvalues $1\pm\nu_i$ $(i=1, \cdots, r)$. 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 $(0, v_k)$, 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 $(By_r, \nu_r y_r)$. The smallest is $\lambda=1-\nu_r$ with an eigenvector $(By_r, -\nu_r y_r)$. 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 $1-|b|\leqq d\leqq 1+|b|$, 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

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

where D_{p_2} is diagonal and Π is a permutation matrix.

References

1. John Todd, *The condition of a certain matrix*, Proc. Cambridge Philos. Soc. vol. 46 (1950) pp. 116–118.

2. David Young, *On the solution of linear systems by iteration*, Proceedings of the Sixth Symposium in Applied Mathematics, to be published.

3. ———, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Amer. Math. Soc. vol. 76 (1954) pp. 92–111.

University of California, Los Angeles