MATRICES IN PERFECT CONDITION

David W. Lewis

We write $\mathrm{GL}(n, \mathbf{R})$ for the group of all non-singular $n \times n$ matrices with real entries. Let A be an element of $\mathrm{GL}(n, \mathbf{R})$ and let $\| \ \|$ be some norm on the real vector space $\mathbf{R}^n$. We define the *operator norm* of A in the usual way, as the supremum of the bounded set $S_A = \{ \|Av\|/\|v\| : v \in \mathbf{R}^n \text{ and } v \neq 0 \}$, and we denote it by $\|A\|$. The operator norm depends on the underlying norm on $\mathbf{R}^n$.

When the norm on $\mathbf{R}^n$ is the usual *euclidean norm*, that is

$$\|v\| = \left(\sum_{i=1}^{n} v_i^2 \right)^{1/2},$$

where $v = (v_i)$, then the corresponding operator norm is the *spectral norm*, so that $\|A\|$ is the square root of the largest eigenvalue of the matrix A^tA. When the norm on $\mathbf{R}^n$ is the *cartesian norm*, that is

$$\|v\| = \max_i |v_i|,$$

then the corresponding operator norm is the maximum absolute row sum norm, given by

$$\|A\| = \max_i (\sum_{j=1}^{n} |a_{ij}|).$$

When the norm on $\mathbf{R}^n$ is the *taxicab norm*, that is

$$\|v\| = \sum_{i=1}^{n} |v_i|,$$

then the corresponding operator norm is the maximum absolute column sum norm, given by

$$\|A\| = \max_j \left(\sum_{i=1}^{n} |a_{ij}| \right).$$

See [3] for proofs.

Definition. The *condition number* of the matrix A in $\mathrm{GL}(n, \mathbf{R})$ with respect to the operator norm $\| \ \|$ is the positive real number $c(A) = \|A\|\|A^{-1}\|$.

Note that $c(A)$ depends on which particular operator norm is in use and $c(A) \geq 1$ for all non-singular matrices A. (This last statement follows from the properties $\|AB\| \leq \|A\| \, \|B\|$ and $\|I\| = 1$, I denoting the identity matrix.)

We remark that condition numbers are important in perturbation theory and yield bounds for errors in numerical methods for solving systems of linear equations, inverting matrices, etc. See [1] and [3].

Definition. The matrix A in $\mathrm{GL}(n, \mathbf{R})$ is said to be *perfectly-conditioned* if $c(A) = 1$.

This definition, of course, depends on which norm is being used.

Most textbooks, including one by the author of this article, [3], say virtually nothing about perfectly-conditioned matrices beyond giving the definition and mentioning that orthogonal matrices are perfectly-conditioned for the spectral norm.

We write $G_{pc} = \{ A \in \mathrm{GL}(n, \mathbf{R}) : c(A) = 1 \}$, so that G_{pc} is the set of all perfectly-conditioned non-singular real $n \times n$ matrices. Here n is a fixed positive integer and our condition numbers are defined with respect to a fixed operator norm.

Lemma. *G_{pc} is a group under the operation of matrix multiplication, so that G_{pc} is a subgroup of* $\mathrm{GL}(n, \mathbf{R})$.

Proof: G_{pc} is a subset of $\mathrm{GL}(n, \mathbf{R})$ which contains I and which is closed under the operation of taking inverses since $c(A) = c(A^{-1})$. Thus it suffices to show that G_{pc} is closed under multiplication.

It is easy to see, via properties of operator norms, that $c(AB) \leq c(A)c(B)$ for any non-singular matrices A and B. Thus $c(AB) = 1$ whenever both $c(A) = 1$ and $c(B) = 1$, since $c(AB) \geq 1$ for all A and B. (If our norm is not an operator norm then G_{pc} need not be a group.)

We will determine the group G_{pc} in general and will specifically describe it for each of the examples of the operator norms given above.

Let $\| \ \|$ be a norm on $\mathbf{R}^n$ and write

$$G_{np} = \{ A \in \mathrm{GL}(n, \mathbf{R}) : \|Av\| = \|v\| \text{ for all } v \in \mathbf{R}^n \},$$

so that G_{np} is the group of all *norm-preserving linear operators* on $\mathbf{R}^n$. (It is an easy exercise to see that G_{np} is a subgroup of $\mathrm{GL}(n, \mathbf{R})$.)

We write $\mathbf{R}_p^*$ for the multiplicative group of all positive real numbers and we will regard $\mathbf{R}_p^*$ as a subgroup of $\mathrm{GL}(n, \mathbf{R})$ by identifying it with the set of all positive scalar multiples of the identity matrix.

Proposition. *Let $\| \ \|$ be a fixed norm on $\mathbf{R}^n$, let A be an element of $\mathrm{GL}(n, \mathbf{R})$, and let $c(A)$ be the condition number of A with respect to this norm. Then $c(A) = 1$ if and only if A is a non-zero scalar multiple of a norm-preserving linear operator on $\mathbf{R}^n$. Indeed the group G_{pc} is isomorphic to the direct product $\mathbf{R}_p^* \times G_{np}$.*

Proof: Consider the set S_A used in the definition of the operator norm $\|A\|$. Note that S_A is a closed and bounded subset of the positive real numbers. It is easy to see that $S_{A^{-1}} = \{ \alpha^{-1} : \alpha \in S_A \}$ because $w = Av$ if and only if $v = A^{-1}w$. Hence $\|A\| = \alpha_1$, where $\alpha_1 = \max S_A$ and $\|A^{-1}\| = \alpha_0^{-1}$ with $\alpha_0 = \min S_A$. It follows that $c(A) = \alpha_1/\alpha_0$, from which we see immediately that $c(A) = 1$ if and only if S_A is a singleton point set. Thus $c(A) = 1$ if and only if there exists a positive real number α such that $\|Av\| = \alpha\|v\|$ for all $v \in \mathbf{R}^n$. Writing $\alpha = \mu^2$ for some positive real number μ, we see that $(\pm\mu^{-1})A$ is a norm-preserving linear operator. It follows easily that G_{pc} is the direct product of the subgroups $\mathbf{R}_p^*$ and G_{np} of $\mathrm{GL}(n, \mathbf{R})$.

Example 1. Using the euclidean norm on $\mathbf{R}^n$, the group G_{np} is well-known to be the orthogonal group $\mathrm{O}(n) = \{ A \in \mathrm{GL}(n, \mathbf{R}) : AA^t = I \}$. Hence $G_{pc} = \mathbf{R}_p^* \times \mathrm{O}(n)$ in this case. Thus the matrices which are perfectly-conditioned with respect to the spectral norm are precisely the positive scalar multiples of the orthogonal matrices.

Example 2. Using the norm on $\mathbf{R}^n$ arising from an inner product given by some positive definite symmetric bilinear form ϕ, the group G_{np} equals $\mathrm{O}(\phi)$, the orthogonal group of ϕ, and $G_{pc} = \mathbf{R}_p^* \times \mathrm{O}(\phi)$ in this case. Note that if ϕ is represented with respect to the standard basis by the matrix B then $\mathrm{O}(\phi) = \{ A \in \mathrm{GL}(n, \mathbf{R}) : A^tBA = B \}$ and also that $\mathrm{O}(\phi)$ is isomorphic to $\mathrm{O}(n)$, because the form ϕ is isometric to the usual dot product on $\mathbf{R}^n$.

Example 3. Using the cartesian norm on $\mathbf{R}^n$, the group G_{np} turns out to be isomorphic to the wreath product $C_2 \wr S_n$, where C_2 is the cyclic group of order 2, and S_n is the symmetric group on n letters. (See [2, p.77] for the definition of wreath product.) We can see this as follows.

If $A \in G_{np}$, then $\|Av\| = \|v\|$ for all $v \in \mathbf{R}^n$. Hence writing $v = (v_i)$ and using the definition of the cartesian norm, the equation $\|Av\| = \|v\|$ becomes

$$\max\left(\left|\sum a_{1j}v_j\right|, \ldots, \left|\sum a_{nj}v_j\right| \right) = \max\left(|v_1|, \ldots, |v_n|\right)$$

for all $(v_1, \ldots, v_n) \in \mathbf{R}^n$. This equality can hold only if each row of A contains exactly one non-zero entry, this non-zero entry being equal to ± 1, and these non-zero entries are all in different columns. (Thus A is a so-called signed permutation matrix.) Examining the multiplication in the group of all such matrices we see that it yields the wreath product $C_2 \wr S_n$, which is the semi-direct product of S_n and C_2^n, where C_2^n is the direct product of n copies of C_2 and S_n acts in the obvious way on C_2^n by permuting factors.

Thus the group of matrices which are perfectly-conditioned with respect to the maximum absolute row sum norm is isomorphic to $\mathbf{R}_p^* \times (C_2 \wr S_n)$. As a set, G_{pc} consists of the positive scalar multiples of the signed permutation matrices.

Example 4. Note that the group G_{pc} in Example 3 is closed under the operation of transposition of matrices. It follows that this same group must also be the group of matrices which are perfectly-conditioned with respect to the maximum absolute column sum norm. ($\|A\|_c = \|A^t\|_r$, where $\| \ \|_c$ and $\| \ \|_r$ denote the maximum absolute column and row sum norms respectively.)

References

[1] Gene H. Golub and Charles F. Van Loan, Matrix Computations (second edition). Johns Hopkins University Press: Baltimore and London, 1989.

[2] Nathan Jacobson, Basic Algebra I. W. H. Freeman and Co.: San Francisco, 1974.

[3] D. W. Lewis, Matrix Theory. World Scientific Publishing: Singapore-New Jersey-London-Hong Kong, 1991.

D. W. Lewis,
Department of Mathematics,
University College,
Belfield,
Dublin 4.

AN ITERATION RELATED TO EISENSTEIN'S CRITERION

Eugene Gath and Thomas J. Laffey

The following question appeared in the 1994 Irish Mathematical Olympiad, the competition used to select the team to represent Ireland in the International Olympiad:
Let a, b and c be real numbers satisfying the equations:

$$b = a(4-a)$$
$$c = b(4-b)$$
$$a = c(4-c).$$

Find all possible values of $a+b+c$.

A direct approach to this problem is to write c in terms of a, and then obtain an octic polynomial in a:

$$f(a) \equiv -a(4-a)(2-a)^2((2-a)^2-2)+a=0.$$

The octic factorizes over the integers in the form

$$f(a) = a(a-3)(a^3-6a^2+9a-3)(a^3-7a^2+14a-7).$$

Observe that the factors a^3-6a^2+9a-3 and $a^3-7a^2+14a-7$ satisfy Eisenstein's irreducibility criterion for the primes 3 and 7, respectively. This, in our experience, was one of the rare occasions when polynomials satisfying the criterion arose in an uncontrived way, and we decided to investigate why they occurred here.

Put $g(x) \equiv x(4-x)$ and let $g^{(r)}(x)$ be the rth iterate $g(g(\ldots g(x)\ldots))$. Consider the polynomial $h_r(x) = x - g^{(r)}(x)$.