

UBIQUITY OF OPERATORS IN MATHEMATICS

DAVID SHERMAN

These notes go along with a lecture given at Linfield College on June 27, 2008 to an audience consisting mostly of undergraduates doing mathematical research (the Willamette Valley REU-RET Consortium). The emphasis is not the same as in the talk, as here I focus on the basic background which some audience members may want to review, and I give extra concepts and exercises which suggest where the theory can go.

1. BASIC DEFINITIONS

1.1. Linear algebra. Let V be a vector space. Optimistically we take the scalar field to be the complex numbers \mathbb{C} , but it may be the reals \mathbb{R} if 1) the mathematics requires it, 2) we want to draw pictures, or 3) the reader is not yet comfortable with \mathbb{C} . By an *operator* on V we mean a linear map from V to V . Recall that the linearity of $T : V \rightarrow V$ is expressed by the equation

$$T(cx + y) = cT(x) + T(y), \quad \forall x, y \in V, \forall c \in \mathbb{C}.$$

Assume that V is finite-dimensional with given basis $\{e_j\}_{j=1}^n$. Then any element $v \in V$ can be uniquely expressed as

$$v = \sum_{j=1}^n c_j e_j.$$

We call the scalars $\{c_j\}$ associated to v its *coordinates*, and we identify v with the column vector $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}$. Descartes gets some credit here.

In the presence of a chosen basis for an n -dimensional V , linear operators on V are in one-to-one correspondence with $M_n(\mathbb{C})$, the $n \times n$ matrices over \mathbb{C} . This works as follows. Any matrix operates by left multiplication on column vectors. And in the other direction, any operator T corresponds to the matrix $[T] = [T_{ij}]$ defined by

$$T(e_j) = \sum_i T_{ij} e_i.$$

It is nice that composition of linear maps corresponds to matrix multiplication, i.e. $[S \circ T] = [S][T]$.

1.2. Norms. A *norm* on a vector space V is a function from V to $[0, +\infty)$ which expresses the size of a vector. Usually we write the value of the norm at v as $\|v\|$. To match our geometric intuition, we require that

- (1) $\|cv\| = |c|\|v\|, \quad \forall v \in V, \quad \forall c \in \mathbb{C};$
- (2) $\|v + w\| \leq \|v\| + \|w\|, \quad \forall v, w \in V;$
- (3) $\|v\| = 0 \iff v = 0.$

For (very great) simplification we will assume here that the basis and norm for V satisfy

$$\left\| \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} \right\| = \sqrt{\sum |c_j|^2},$$

whether the number of basis elements is finite or infinite. This is based on the Pythagorean theorem. Some of you may know that this norm comes from an inner product, and that if V is complete in this norm, it is called a *Hilbert space*. But we won't develop that here.

Now for a linear map $T : V \rightarrow V$, we define its *operator norm* by

$$\|T\| = \sup_{\|v\|=1} \|T(v)\| = \sup_{v \neq 0} \frac{\|T(v)\|}{\|v\|}.$$

This tells us how much bigger a vector can get when T operates on it. For infinite-dimensional V we restrict attention to those T with finite operator norm – we call this class the *continuous* or *bounded* linear operators, denoted $B(V)$ (sometimes $L(V)$).

1.3. Examples. The identity operator $I \in B(V)$ is defined by $I(v) = v$. Obviously $\|I\| = 1$. When $V \simeq \mathbb{C}^2$, $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Other calculations:

$$\left\| \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\| = 1 = \left\| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\|.$$

$$\left\| \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\| = \sqrt{2}.$$

$$\left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\| = 2.$$

$$\left\| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\| = \sqrt{\frac{3+\sqrt{5}}{2}}.$$

1.4. Exercises.

1. Check the norm computations immediately above. Also show that the norm of a diagonal matrix of any size is the supremum of the absolute values of the diagonal entries.

2. Show that the operator norm is in fact a norm on the vector space $B(V)$. (This requires you to recognize the vector space structure of $B(V)$.) Show that the operator norm is *submultiplicative*: $\|ST\| \leq \|S\|\|T\|$. If you can also show that $B(V)$ is complete in the operator norm, you have shown that $(B(V), \|\cdot\|)$ is a *Banach algebra*.

3. Define instead, on \mathbb{C}^n , $\left\| \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right\| = \sum_j |c_j|$. Show that this is a norm, and give a formula for the operator norm of $[T_{ij}]$. (First try some easy examples.)

2. RANDOM WALKS ON GRAPHS

2.1. Unbiased random walk on a circle. Consider the graph of an n -cycle, which we realize as the unit circle of \mathbb{C} with vertices at the n roots of unity $\{1, \omega, \omega^2, \dots, \omega^{n-1}\}$. Let V be the n -dimensional space of functions on the vertices, and choose the basis elements to be functions which are one at a single vertex and zero elsewhere. As a result, the coordinate vector corresponding to a function f is $\begin{pmatrix} f(1) \\ f(\omega) \\ \vdots \end{pmatrix}$ – just the list of the function’s values.

Consider the operator U whose matrix is

$$\begin{pmatrix} 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

This operator is a cyclic permutation of the basis elements, rotating all functions counterclockwise one unit. Explicitly, $Uf(x) = f(\omega^{-1}x)$.

Recall that a matrix A is *normal* if it commutes with its conjugate transpose A^* . Among the normal matrices are the unitaries ($A^* = A^{-1}$) and the self-adjoints ($A^* = A$). A fundamental theorem of linear algebra says that a normal matrix is unitarily equivalent to a diagonal matrix whose diagonal is exactly its list of eigenvalues. (See a textbook for more detail.) Notice that U is unitary, with eigenvalue list $\{1, \omega, \omega^2, \dots\}$.

If f is a nonnegative function whose values add to 1, we can think of it as a *probability density*. Imagine that a man is somewhere on the graph, and $f(\omega^j)$ is the probability he is at ω^j . Supposing

that he walks one unit counterclockwise from wherever he is, his new probability density is exactly Uf .

Now suppose that he starts at 1, and every minute he flips a fair coin and moves either clockwise or counterclockwise one unit. We call this a *random walk*. Although there are many kinds of questions and variations, we take only this simple setup and ask: What is the chance he is back at 1 after an hour? You can cook up an expression using just combinatorics, and this has pros and cons, but here we want to model the problem with operators.

The point is to see that one step of this random walk is an operator W which takes any density f to $(\frac{U+U^{-1}}{2})f$. We want, then, to compute the first coordinate of

$$W^{60} \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}.$$

Recalling that the eigenvalues of U are $\{e^{\frac{2\pi ij}{n}}\}$, it follows that W is diagonalizable with eigenvalues $\{e^{\frac{2\pi ij}{n}} + e^{-\frac{2\pi ij}{n}}\} = \{\cos \frac{2\pi j}{n}\}$. A little more calculation shows that the final answer is

$$(1) \quad \frac{1}{n} \sum_{j=0}^{n-1} \left(\cos \left(\frac{2\pi j}{n} \right) \right)^{60}.$$

2.2. Unbiased random walk on a line. We next carry out a similar analysis for the unbiased random walk on the integers: start at the origin, each minute move left or right with equal probability. We have infinite versions of the operators U and W from above. Let's answer the same question regarding a return to the origin after 60 steps.

This will make more sense if we recall a theme from Fourier analysis first: in some sense, "any" function g on the unit circle can be written as $\sum_{n \in \mathbb{Z}} c_n e^{int}$. Then in the basis $\{e^{int}\}_{n \in \mathbb{Z}}$, multiplication by e^{it} is exactly U ! Similarly W is multiplication by $\frac{e^{it} + e^{-it}}{2} = \cos t$. The initial probability vector is the zeroth basis element, $e^{i0t} = 1$. We need to compute the zeroth Fourier coefficient of $(\cos t)^{60} \cdot 1$, which is

$$(2) \quad \frac{1}{2\pi} \int_0^{2\pi} (\cos t)^{60} dt.$$

Note that (1) is a Riemann sum approximation to (2).

2.3. Exercises.

1. Recalling the questions answered by (1) and (2), explain why they are equal if and only if $n > 60$. How do they compare if $n \leq 60$? Find combinatorial expressions which answer the same questions.

2. Find the operator norm of the bi-infinite matrix (entries labeled by \mathbb{Z}) which is $-\frac{1}{2}$ on the diagonal, 2 on the subdiagonal, 1 on the subsubdiagonal, and 0 elsewhere. (The shorter, more interesting part of this question is to rephrase it as a calculus problem – *most undergraduates will need some guidance in order to do this*. Then you still have to do the problem.)

3. GROUP THEORY

3.1. Cayley graphs. Let G be a discrete group. Fix a subset $S \subset G$ such that

- S is finite;
- the identity e is not in S ;
- S is *generating* (every element of G is a product of elements in S);
- S is *symmetric* ($s \in S \Rightarrow s^{-1} \in S$).

In the *Cayley graph* associated to $S \subset G$, the vertices are identified with the elements of G , and the edges connect all possible pairs of vertices (g, gs) , $g \in G$, $s \in S$.

For \mathbb{Z} , the natural choice for S is $\{\pm 1\}$, and the Cayley graph can be realized as an infinite line. For $S = \{(\pm 1, 0), (0, \pm 1)\} \subset \mathbb{Z}^2$, the Cayley graph is an infinite 2-dimensional grid. For $S = \{\pm 1\} \subset \mathbb{Z}_n$, the Cayley graph is an n -cycle.

Although precise properties of the Cayley graph of G do depend on the choice of S , certain large-scale properties do not.

3.2. Random walk operators on Cayley graphs. Let $\ell^2(G)$ be the vector space of square-summable functions on G – that is, $\|f\| = \sqrt{\sum_{g \in G} |f(g)|^2} < \infty$. For any element $h \in G$ we have a unitary (=invertible norm-preserving) operator U_h on $\ell^2(G)$ defined by

$$[U_h(f)](g) = f(h^{-1}g).$$

So U_h shifts the graph of any function by h . (Why h^{-1} instead of h ? This makes the map $h \mapsto U_h$ a group homomorphism.) The operators U in Sections 2.1 and 2.2 are nothing but U_1 on $\ell^2(\mathbb{Z}_n)$ and $\ell^2(\mathbb{Z})$, respectively. (But notice that embedding \mathbb{Z}_n in the circle group changes the role of 1: $1 \in \mathbb{Z}_n$ corresponds to $\omega \in \{1, \omega, \dots, \omega^{n-1}\}$.)

The random walks we consider on G are implemented by finite convex combinations of the operators $\{U_g\}$, i.e.

$$W = \sum_{g \in S \cup \{e\}} p_g U_g, \quad p_g \in [0, 1], \quad \sum p_g = 1.$$

This corresponds to a random walk in which each step is in the direction g with probability p_g .

In general the effect of W on a probability density is to “spread it out,” and this in turn has the effect of decreasing the norm of the density in $\ell^2(G)$. Always the operator norm of W is ≤ 1 :

$$\|W\| = \left\| \sum p_g U_g \right\| \leq \sum \|p_g U_g\| = \sum p_g = 1.$$

(Here we used that the operator norm is a norm.)

It is a fact that on \mathbb{Z}_n , \mathbb{Z} , and \mathbb{Z}^2 , any random walk operator has norm equal to 1. *Can you see why this is true???*

Let’s make up some terminology and say that a group has *property W* if any random walk operator on it has norm 1.

The free group on two generators, \mathbb{F}_2 , does not have property W. With a, b the generators, $\|\frac{1}{4}(U_a + U_{a^{-1}} + U_b + U_{b^{-1}})\| = \frac{\sqrt{3}}{2}$.

3.3. Exercises.

1. Prove that \mathbb{Z}_n , \mathbb{Z} , and \mathbb{Z}^2 have property W. (As a first step you may assume that the sets S are as above, but the full meaning of the statement includes any choice of S .) Prove that any finite group has property W.

2. Looking at the structure of your proof for \mathbb{Z} , conjecture another characterization of property W. (There are actually very many.)

3.4. Amenability (briefly). The idea behind Exercise 3.3.1 above is that there are probability vectors which are almost unchanged by a random walk operator W . For the case of \mathbb{Z} , choose large enough N and let f equal $\frac{1}{\sqrt{2N+1}}$ on $\{-N, -N+1, \dots, N-1, N\}$, zero elsewhere.

Continuing with \mathbb{Z} , we change the perspective from probability vectors to finite subsets. The important thing is that for any finite set of numbers $\{n_j\}$ (the possible jumps in a random walk), there is a finite subset of \mathbb{Z} which changes very little (proportionally to itself) if it is translated by any of the $\{n_j\}$. The extent to which you understand the preceding sentence and the succeeding paragraph is directly proportional to the extent to which you figured out Exercise 3.3.1.

Let's give a precise definition (one of many possible). Say that a discrete group G is *amenable* if for any finite subset $\{g_j\}$ and $\varepsilon > 0$, there is another finite subset F such that

$$\#(F \cap g_j F) > (1 - \varepsilon)(\#F), \quad \forall j.$$

Here “ $\#$ ” just means the number of elements in the set, while “ $g_j F$ ” is the subset obtained by multiplying each element of F by g_j (should be thought of as F translated by g_j).

It is a fact that a group is amenable if and only if it has property W.

Amenable groups are in some sense well-approximated by finite subsets. The visual content of property W says that their Cayley graphs don't spread out too quickly, so that it is possible for a probability density supported on a large region to be nearly invariant under a random walk operator.

A sentient being can easily find more information about amenability, although most approaches require some background in functional analysis. Wikipedia at least isn't misleading, a better wiki is at wiki.canisiusmath.net/index.php?title=Amenability_Equivalencies, and most everything is in the authoritative and welcoming reference *Amenability* by Alan Paterson.

Amenability is at the heart of a famous classical result, the *Banach-Tarski paradox*. This says that in dimension 3 or higher, a unit ball can be divided into finitely many pieces which can then be reassembled to form two unit balls. See the fun eponymous book by Stan Wagon.

Finally we explain what it means for a *graph* G to be amenable. Without an underlying group we cannot define translation operators, but we can still imagine an unbiased random walk. The condition for amenability is

$$\inf_{S \subset G, 0 < \#S < \infty} \frac{\#\text{edges with exactly one vertex in } S}{\#S} = 0.$$

In other words, there are subsets which are much bigger than their boundaries. The opposite condition says (spiritually) that perimeter is always at least a fixed fraction of area – the kind of thing that doesn't happen in finite dimensions. And Cayley graphs of \mathbb{F}_2 do not embed “nicely” in a finite-dimensional space, while those of \mathbb{Z}^2 do.

This is consistent with the previous definition: a discrete group is amenable if and only if all its Cayley graphs are amenable graphs.

Comments and corrections are welcome!

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VIRGINIA, P.O. BOX 400137, CHARLOTTESVILLE, VA 22904
E-mail address: dsherman@virginia.edu