Honors 1: Undergraduate Math Lab*

Peter Sarnak[†], Steven Miller[‡], Alex Barnett[§]

September 25th, 2002

Courant Institute of Mathematical Sciences
New York University
New York, NY

Abstract

We review some basic facts about Linear Algebra and Matrix Groups, and give an introduction to Random Matrix Theory. Lecture by Steven Miller; notes by Steven Miller and Alex Barnett.

1 Linear Algebra Review

Matrices can either be thought of as rectangular (often square) arrays of numbers, or as linear transformations from one space to another (or possible to the same space). The former picture is the simplest starting point, but as Professor Sarnak emphasized, it is the latter, geometric view that gives a deeper understanding.

To connect with the simpler vector case, a vector can be thought of as a list of real numbers which change in a certain way when the coordinate system changes, or as a geometric object with length and direction. The latter object is *coordinate-independent*, and has different representations in different choices of coordinate axes. Try to keep the geometric picture in mind for matrices.

---

*Homepage: `http://www.math.nyu.edu/~millerj/`

[†]E-mail: `sarnak@math.princeton.edu`

[‡]E-mail: `millerj@cims.nyu.edu` or `sjmiller@math.princeton.edu`

[§]E-mail: `barnett@nmr.mgh.harvard.edu`

1.1 Definitions

Given an $n \times m$ matrix A (where n is the number of rows and m is the number of columns), the **transpose of** A, denoted A^T, is the $m \times n$ matrix where the rows of A^T are the columns of A (or, equivalently, the columns of A^T are the rows of A).

Lemma 1.1 *$(AB)^T = B^T A^T$ and $(A^T)^T = A$.*

We leave the proof to the reader.

If an $n \times n$ matrix (also called a **square** matrix) A satisfies $A^T = A$, then we say A is **symmetric**

Example 1.2 *Let A be the matrix*

$$\begin{pmatrix} 2 & 2 & 4 & 2 \\ 1 & 1 & -2 & 2 \\ -2 & 0 & 0 & 1 \\ 1 & 1 & 2 & 1 \end{pmatrix} \tag{1}$$

Then A^T is

$$\begin{pmatrix} 2 & 1 & -2 & 1 \\ 2 & 1 & 0 & 1 \\ 4 & -2 & 0 & 2 \\ 2 & 2 & 1 & 1 \end{pmatrix} \tag{2}$$

Note the above matrix is not symmetric.

The number of *degrees of freedom* in a symmetric matrix (*i.e.* independent real numbers needed to completely specify the matrix) is $n(n+1)/2$. Why? There are n^2 entries, n on the diagonal. If you specify all entries above the diagonal and all entries on the diagonal, then you know the symmetric matrix.

There are $n^2 - n$ non-diagonal entries (half above the diagonal, half below). Thus, one needs to specify $\frac{n^2-n}{2} + n = \frac{n^2+n}{2}$ entries.

Exercise 1.3 *If A and B are symmetric, show AB is symmetric.*

Matrix multiplication. We call the element in the i^{th} row and j^{th} column a_{ij}. Think of $i = 1 \cdots n$ going down the left side, and $j = 1 \cdots M$ going across the top. A vector v we represent as a column of elements with the i^{th} being v_i. A nice way to see matrix-vector multiplication is that the v_i give the *coefficients* by which the columns of A are linearly mixed together. For the product $w = Av$ to make sense, the length (dimension) of v must equal m, and the dimension of w will be n. A is therefore a linear transform from m-dim space to n-dim space.

Multiple transformations appear written backwards: if we apply A then B then C to a vector, we write

$$w \;=\; CBAv. \tag{3}$$

Note that taking the product of two $n \times n$ matrices requires $O(n^3)$ effort.

Exercise 1.4 *Show that there are two ways to evaluate triple matrix products of the type CBA. The slow way involves $O(n^4)$ effort. How about the fast way? How do these results scale for the case of a product of k matrices?*

Definition 1.5 (Invertible Martices) *A is invertible if a matrix B can be found such that $BA = AB = I$. The inverse is then written $B = A^{-1}$. Invertibility requires A to be square.*

Transformations of a matrix. Just as with vectors, we can find out how the components of a square matrix A change under transformation. Say we have a scalar quantity $x = w^T Av$. We transform our coordinate system linearly such that the vector v has components $v' = Mv$, where M is some invertible matrix representing the transformation. Therefore also $w' = Mw$. The only way that x can remain unchanged by the transformation (as any scalar must), for all choices of v and w, is if the transformed matrix is written $A' = M^{-T}AM^{-1}$. Check this via

$$x' = w'^T A'v' = (Mw)^T(M^{-T}AM^{-1})(Mv) = w^T IAIv = w^T Av = x. \tag{4}$$

This is called a *similarity transformation*, or a *conjugation*. Really we have one object, the transformation A, but it may have different representations by a matrix of numbers, depending on the choice of basis.

Definition 1.6 (Orthogonal Matrices) *Q is an orthogonal $n \times n$ matrix if it has real entries and $Q^TQ = QQ^T = I$.*

Q is invertible, with inverse Q^T. The geometric meaning of Q is a *rotation*: the vector $w = Qv$ is just v rotated (about the origin).

The number of degrees of freedom in an orthogonal matrix is $n(n-1)/2$.

Exercise 1.7 *In 3 dimensions a general rotation involves 3 angles (for example, azimuth, elevation, and 'roll'). How many angles are needed in 4 dimensions? In 3d you rotate about a line-like axis (the set of points which do not move under rotation); what object do you rotate about in 4d?*

Exercise 1.8 *Show that the identity matrix I, always has representation $I_{ij} = \delta_{ij}$ regardless of the choice of basis. Hint: perform orthogonal tranformation on the matrix δ_{ij}.*

The set of orthogonal matrices of order n forms a *continuous* (or *topological*) group, which we call $O(n)$. (Not to be confused with "of order N"). Group properties:

- Associativity follows from that of matrix multiplication.
- The identity matrix acts as an identity element, since it is in the group.
- Inverse is the transpose (see above): $Q^{-1} = Q^T$.
- Closure is satisfied because any product QR of orthogonal matrices is itself orthogonal.

Exercise 1.9 *Prove the last assertion.*

However, not all the elements of $O(n)$ can 'talk' to each other, *i.e.* you cannot reach all the elements by continuous transformation from the identity I.

Example for $n = 2$: a general order-2 orthogonal matrix can be written

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}, \tag{5}$$

where $0 \le \theta < 2\pi$ is a real angle. The first has determinant $+1$ and defines the 'special' (*i.e.* unit determinant) group $SO(2)$ which is a subgroup of $O(2)$ with identity I. The second has determinant -1 and corresponds to rotations with a reflection; this subgroup is disjoint from $SO(2)$, and has the weird (reflecting) identity can be written in some basis as

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6}$$

Note that $SO(2)$, alternatively written as the family of planar rotations $R(\theta)$, is *isomorphic* to the unit length complex numbers under the multiplication operation:

$$R(\theta) \;\longleftrightarrow\; e^{i\theta}. \tag{7}$$

Therefore we have $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$. This commutativity relation does *not* hold in higher $n > 2$.

Orthogonal transformations. If an orthogonal matrix Q is used for conjugation of a general square matrix A, then the rule Eq. 4 for transformation becomes,

$$A' \;=\; QAQ^T. \tag{8}$$

This tells you how to 'rotate' a (square) matrix.

Definition 1.10 (Complex Conjugate Transpose) *Let A be an $n \times m$ matrix. Then the complex conjugate transpose of A, denoted A^*, is obtained by the following: (1) take the complex conjugate of A; ie, replace every entry $a_{jk} = x_{jk} + iy_{jk}$ with $\overline{a_{jk}} = x_{jk} - iy_{jk}$, and call this matrix A_1; (2) take the transpose of A_1.*

Exercise 1.11 *Prove that $(AB)^* = B^*A^*$.*

Definition 1.12 (Dot or Inner Product) *The dot (or inner) product of two real vectors v and w is defined as $v^T w$; if the vectors are complex, we instead use $v^* w$.*

Exercise 1.13 *Show that the dot product is invariant under orthogonal transformation. That is, show that given two vectors, transforming them using the same orthogonal matrix leaves their dot product unchanged.*

Definition 1.14 (Length of a vector) *The length of a real vector v is $|v|^2 = v^T v$; for a complex vector, we have $|v|^2 = v^* v$.*

Definition 1.15 (Orthogonality) *Two real vectors are orthogonal (also called perpendicular) if $v^T w = 0$; for two complex vectors, the equivalent condition is $v^* w = 0$.*

Definition 1.16 (Eigenvalue, Eigenvector) *Recall λ is an* ***eigenvalue*** *and v is an* ***eigenvector*** *if $Av = \lambda v$ and v is not the zero vector.*

Exercise 1.17 *If v is an eigenvector of A with eigenvalue λ, show $w = av$, $a \in \mathbb{C}$, is also an eigenvector of A with eigenvalue λ.*

Exercise 1.18 *Show that given an eigenvalue λ and an eigenvector v, you can always find an eigenvector w with the same eigenvalue, but $|w| = 1$.*

To find the eigenvalues, we solve the equation $\det(\lambda I - A) = 0$. This gives a polynomial $p(\lambda) = 0$. We call $p(\lambda)$ the **characteristic polynomial**.

The **trace** of a matrix A, denote $\mathrm{Tr}(A)$ is the sum of the diagonal entries of A:

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}. \tag{9}$$

Lemma 1.19 *$Tr(A) = \sum_{i=1}^{n} \lambda_i$.*

The proof relies on writing out the characteristic equation and comparing powers of λ with the factorized version. By the fact that the polynomial has roots λ_i we can write

$$\det(\lambda I - A) \;=\; p(\lambda) \;=\; \prod_{i=1}^{n} (\lambda - \lambda_i). \tag{10}$$

Note the coefficient of λ^n is 1, thus we have $\prod_i (\lambda - \lambda_i)$ and not $c \prod_i (\lambda - \lambda_i)$ for some constant c.

By expanding out the RHS, the coefficient of λ^{n-1} is $-\sum_{i=1}^{n} \lambda_i$, which we will show is $-\mathrm{Tr}(A)$. Expanding the LHS, we want to find the corresponding coefficient in

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & & \\ \vdots & & \ddots & \\ -a_{n1} & & & \lambda - a_{nn} \end{vmatrix}.$$

We have to remember the expansion of the determinant. Taking the top-left-most 2×2 block, we see its determinant is $(\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21})$. The determinant of the top-left-most 3×3 block is then formed by $(\lambda - a_{33})$ times the above 2×2 determinant, plus two other multiples of determinants which can give only a highest power of λ of λ^1. Thus we see that the coefficient in λ^2 is $-(a_{11} + a_{22} + a_{33})$. Repeating this argument for 4×4 block up to $n \times n$ gives us the coefficient of λ^{n-1} in the full determinant is $-\sum_{i=1}^{n} a_{ii}$. Since the LHS and RHS must be equal $\forall \lambda$, the LHS and RHS coefficients in λ^{n-1} are equal. $\Box$

Corollary 1.20 *Tr(A) is invariant under rotation of basis.*

The proof follows immediately from the invariance of the eigenvalues under rotation of basis. We need the following:

Lemma 1.21 $\det(AB) = \det(A)\det(B)$. *Further, by induction one can show* $\det(AB \cdots Z) = \det(A)\det(B)\cdots\det(Z)$. *Further,* $\det(I) = 1$.

Proof of Corollary: Let $A = Q^T BQ$. We show A and B have the same trace by showing A and B have the same eigenvalues. To find the eigenvalues of A we must solve:

$$\begin{aligned}
\det(\lambda I - A) &= \det(\lambda I - Q^T BQ) \\
&= \det(\lambda Q^T Q - Q^T BQ) \\
&= \det(Q^T \lambda IQ - Q^T BQ) \\
&= \det\left(Q^T(\lambda I - B)Q\right) \\
&= \det(Q^T)\det(\lambda I - B)\det(Q) \\
&= \det(Q^T)\det(Q)\det(\lambda I - B) \\
&= \det(Q^T Q)\det(\lambda I - B) = \det(I)\det(\lambda I - B) = \det(\lambda I - B).
\end{aligned} \tag{11}$$

As the eigenvalues of A and B satisfy the same equation, they are equal. $\Box$

1.2 Spectral Theorem for Real Symmetric Matrices

The main theorem we will prove is

Theorem 1.22 (Spectral Theorem) *Let A be a real symmetric $n \times n$ matrix. Then there exists an orthogonal $n \times n$ matrix Q and a diagonal matrix Λ such that $Q^T AQ = \Lambda$. Moreover, the n eigenvalues of A are the diagonal entries of Λ.*

This result is remarkable: it tells you that any real, symmetric matrix is diagonal when rotated into an appropriate basis (recall the rotation effect of conjugation using Q). In other words, the operation of matrix A on a vector v can be broken down into three steps:

$$Av = Q\Lambda Q^T v = (\text{undo the rotation})(\text{stretch along coord axes})(\text{rotation})v. \tag{12}$$

Recall the ordering of transformations is read like Hebrew, right to left. The rotation is just the rotation into the basis of eigenvectors.

Furthermore, the eigenvalues λ_i (= diag els of Λ) are a set of numbers invariant under rotations of A. In other words, if $A' = PAP^T$ is an orthogonally-conjugated (*i.e.* P is orthogonal) version of A, then A' has the same $\{\lambda_i\}$ as A. Of course the ordering of the λ_i has to be chosen the same.

For the Spectral Theorem we prove a sequence of needed lemmas:

Lemma 1.23 *The eigenvalues of a real symmetric matrix are real.*

Let A be a real symmetric matrix with eigenvalue λ and eigenvector v. Note that we do not yet know that v has only real coordinates!

Therefore, $Av = \lambda v$. Take the dot (or inner) product of both sides with the vector v^*, the complex conjugate transpose of v:

$$v^* A v = \lambda v^* v. \tag{13}$$

But the left hand side is real. The two sides are clearly complex numbers (ie, 1-dimensional matrices). Taking the complex conjugate transpose of the LHS gives

$$\Big(v^*(Av)\Big)^* = (Av)^*(v^*)^* = v^* A v. \tag{14}$$

Therefore, the LHS is real, implying the RHS is real. But clearly $v^* v$ is real (similar calculation). Thus, λ is real. $\Box$

We will only prove the Spectral Theorem when all the eigenvalues are distinct. Henceforth, we shall always assume A is a real symmetric matrix.

Lemma 1.24 *The eigenvectors of a real symmetric matrix are real.*

The eigenvalues solve the equation $(\lambda I - A)v = 0$. Let λ be an eigenvalue. Then $\det(\lambda I - A) = 0$. Therefore the matrix $B = \lambda I - A$ is not invertible. Therefore it send a vector to 0 (standard linear algebra calculation).

Lemma 1.25 *If λ_1 and λ_2 are two distinct eigenvalues of a real symmetric matrix A, then their corresponding eigenvectors are perpendicular.*

We study $v_1^T A v_2$. Now

$$v_1^T A v_2 = v_1^T (A v_2) = v_1^T (\lambda_2 v_2) = \lambda_2 v_1^T v_2. \tag{15}$$

Also,

$$v_1^T A v_2 = v_1^T A^T v_2 = (v_1^T A^T) v_2 = (A v_1)^T v_2 = (\lambda_1 v_1)^T v_2 = \lambda_1 v_1^T v_2. \tag{16}$$

Therefore

$$\lambda_2 v_1^T v_2 = \lambda_1 v_1^T v_2 \quad \text{or} \quad (\lambda_1 - \lambda_2) v_1^T v_2 = 0. \tag{17}$$

As $\lambda_1 \neq \lambda_2$, $v_1^T v_2 = 0$. Thus, the eigenvectors v_1 and v_2 are perpendicular. $\Box$

We can now prove the Spectral Theorem for real symmetric matrices **if there are n distinct eigenvectors**.

Let λ_1 to λ_n be the n distinct eigenvectors, and let v_1 to v_n be the corresponding eigenvectors chosen so that each v_i has length 1.

Consider the matrix Q, where the first column of Q is v_1, the second column of Q is v_2, all the way to the last column of Q which is v_n:

$$Q = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \cdots & v_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \tag{18}$$

The transpose of Q is

$$Q^T = \begin{pmatrix} \leftarrow & v_1 & \rightarrow \\ & \vdots & \\ \leftarrow & v_n & \rightarrow \end{pmatrix} \tag{19}$$

Exercise 1.26 *Show that Q is an orthogonal matrix. Use the fact that the v_i all have length one, and are orthogonal (perpendicular) to each other.*

Consider $Q^T A Q$. This is a matrix, call it B. To find its entry in the i^{th} row and j^{th} column, we look at

$$e_i^T B e_j \tag{20}$$

where the e_k are column vectors which are 1 in the k^{th} position and 0 elsewhere:

$$e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{21}$$

Thus, we need only show that $e_i^T B e_j = 0$ if $i \neq j$ and equals λ_j if $i = j$.

Exercise 1.27 *Show $Qe_i = v_i$ and $Q^T v_i = e_i$.*

We calculate

$$\begin{aligned} e_i^T B e_j &= e_i^T Q^T A Q e_j \\ &= (e_i^T Q^T) A (Q e_j) \\ &= (Q e_i)^T A (Q e_j) \\ &= v_i^T A v_j \\ &= v_i^T (A v_j) \\ &= v_i^T \lambda_j v_j = \lambda_j v_i^T v_j. \end{aligned} \tag{22}$$

As $v_i^T v_j$ equals 0 if $i \neq j$ and 1 if $i = j$, this proves the claim.

Thus, the off-diagonal entries of $Q^T A Q$ are zero, and the diagonal entries are the eigenvalues λ_j. This shows that $Q^T A Q$ is a diagonal matrix whose entries are the n eigenvalues of A. $\square$

Note that, in the case of n distinct eigenvalues, not only can we write down the diagonal matrix, we can easily write down what Q should be. Further, by reordering the columns of Q, we see we reorder the positioning of the eigenvalues on the diagonal.