Linear Algebra 1

Basics of matrix manipulation and linear algebra, basics on eigenvalues and determinants

by Ivan Aidun

These are notes interspersed with exercises. The purpose of these notes is to be more fleshed out than Evan Dummit's old notes, but shorter and more focused than a textbook treatment. Halfway between Dummit and Dummit and Foote, so to speak. I hope these are helpful to you!

Conventions

In these notes, most vector spaces have been chosen to be real or complex vector spaces for the sake of concreteness. Anything stated over $\mathbb{R}$ can be done over any field, and anything stated over $\mathbb{C}$ can be done over any algebraically closed field. The one exception to this would be the definition of an inner product, which presupposes that your vector space is either real or complex (otherwise you have no inner products at all).

I freely exchange between referring to a matrix and to a linear transformation, except in cases where a single linear transformation may have several different matrices (due to a change of basis occurring).

Elementary techniques and constructions

Here are two pieces of philosophy that can be helpful when thinking about linear algebra:

- An invertible linear transformation is essentially a change of basis, and a change of basis is essentially a change of perspective. Many matrix manipulations involve "changing your perspective" several times until you arrive at a basis that is easy to read information from.
- Most (but not all) true statements about $n \times n$ matrices for n large are true for 2×2 matrices, and most (but not all) false statements about matrices have a 2×2 counterexample. So, always look for 2×2 examples/counterexamples first, since they are much easier to manipulate.

It is also good to keep in mind a list of criteria equivalent to a matrix being invertible.

Proposition. For a square matrix A, the following are equivalent.

- A is invertible,
- A has linearly independent columns (or rows),
- A has rank ($= \dim \operatorname{im} A$) n,
- A has trivial kernel,
- 0 is not an eigenvalue of A,
- $\det A \neq 0$.

Row Reduction

The most computationally important matrix manipulations are the **Elementary Row (or Column) Operations**. They come in three flavors, creatively named Type I, Type II, and Type III row operations. They are as follows:

Type I: swap two rows,

Type II: multiply a row by a nonzero scalar,

Type III: add a multiple of one row to another.

(Note, by combining type II and type III operations, we can replace one row by an arbitrary linear combination of that row with other rows.)

Exercise 1. The different kinds of elementary row operations correspond to multiplying on the left by certain matrices. What kinds of matrices give you each of the types of row operation?

These operations are employed in an algorithm called **row reduction** or **Gauss-Jordan reduction** (not the same Jordan as Jordan Normal Form, weirdly). The algorithm goes like

1. Choose a row with a nonzero entry. Swap it to the top, and multiply by an appropriate constant so the leftmost entry is 1. This entry is called the *pivot*.
2. Use Type III operations to make all the entries above and below the pivot 0.
3. Recurse onto the submatrix obtained by eliminating the first row and column, and return to step 1.

At the end of the algorithm, you get a block matrix that looks like

$$\begin{pmatrix} I_k & A \\ 0 & 0 \end{pmatrix}$$

for some matrix A. In the same vein, you could also column reduce a matrix, or both row *and* column reduce a matrix.

Here's how to interpret what the row reduction algorithm does: given a linear transformation $A\colon V \to W$ between finite-dimensional vector spaces, and a basis $\{v_1, \ldots, v_m\}$ for V and a basis $\{w_1, \ldots, w_n\}$ for W, we can form the **matrix of the transformation** (with respect to these two bases) as the matrix whose (i, j) entry is the coefficient of w_i in the expansion of $A(v_j)$. The row reduction algorithm changes the basis of W until $w_i = A(v_i)$ for as many i as possible. Conversely, column reduction changes the basis on V until $\operatorname{span}(A(v_i) - w_i)$ is as small as possible. (Notice that if $V = W$, we are keeping track of *two possibly different* bases for V throughout this algorithm.)

Both forms make it easy to read off information about both the kernel and the image, but the row reduced form is particularly conducive to information about the image, and column reduced form to information about the kernel, because of the aforementioned properties. For the next problem, try both parts using row reduction, then both parts using column reduction.

Exercise 2. The matrix

$$A = \begin{pmatrix} 1 & -5 & 5 & 2 \\ -1 & 5 & -4 & -5 \\ -1 & 5 & -3 & -8 \end{pmatrix}$$

defines a linear transformation $\mathbb{R}^4 \to \mathbb{R}^3$ given by $T(v) = Av$.

(a) Compute a basis for the kernel of this transformation.

(b) Compute a basis for the image of this transformation.

Occasionally, row reduction problems, or problems made much easier to think about with row reduction, have actually appeared on the qual.

Exercise 3. (August 2019 Problem 4) This problem involves 4×4 matrices with entries in $\mathbb{R}$. We will say that a matrix M is an **E-matrix** if M has 1's along the diagonal and a single nonzero entry off the diagonal. For instance:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 31 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is an E-matrix. Express the matrix A below as a product of E-matrices.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

Exercise 4. (January 2020 Problem 4) Let V be a finite-dimensional vector space over a field. Let $A\colon V \to V$ be a linear transformation.

(a) Show that there exists $B\colon V \to V$ such that $ABA = A$.

(b) Show that there exists $C\colon V \to V$ such that $ACA = A$ and $CAC = C$.

(**NB** a matrix C like in part (b) is known as a "pseudo-inverse" to A.)

Direct sums and complements

As in many areas of algebra, given two vector spaces V, W, you can form a new vector space whose underlying set is $V \times W$. In the case of vector spaces (and, we will later see, abelian groups and modules) this construction is referred to as the **direct sum** $V \oplus W$ rather than the direct product, because in many ways it feels more like adding the vector spaces than like multiplying them. For example, $\dim V \oplus W = \dim V + \dim W$.

There is also, as in other areas, the concept of a vector space being an *internal* direct sum. Given a vector space V and subspaces U_1, U_2, we say V is the **internal direct sum** of U_1 and U_2

if the following two things hold:

(i) every vector $v \in V$ can be written as $v = u_1 + u_2$ for $u_1 \in U_1$, $u_2 \in U_2$, and

(ii) this representation is unique.

The condition (i) is often written as $U_1 + U_2 = V$, or perhaps as $\langle U_1, U_2 \rangle = V$ or $\operatorname{span}(U_1, U_2) = V$.

Exercise 5. Let V be a vector space of dimension n, U_1, U_2 subspaces, and let $\mathcal{B}_1 = \{u_{11}, \ldots, u_{1k}\}$ be a basis for U_1 and $\mathcal{B}_2 = \{u_{21}, \ldots, u_{2\ell}\}$ be a basis for U_2.

(a) **(The direct sum test)** Prove that V is the internal direct sum of U_1 and U_2 if and only if $\dim U_1 + \dim U_2 = n$ and $U_1 \cap U_2 = \{0\}$.

(b) Prove that $U_1 + U_2 = V$ if and only if $\mathcal{B}_1 \cup \mathcal{B}_2$ contains a basis for V.

(c) Prove that V is the internal direct sum of U_1 and U_2 if and only if $\mathcal{B}_1 \cup \mathcal{B}_2$ *is* a basis for V.

(d) Notice that $V \oplus W$ has subspaces $\{(v, 0)\} \cong V$ and $\{(0, w)\} \cong W$. Show that $V \oplus W$ is the internal direct product of these two subspaces.

If U is a subspace of V, another subspace U' is called a **complement** to U if V is the internal direct sum of U and U'. In general, the complement is not unique, **Exercise 6.** find two different complements to

$$\operatorname{span}\begin{pmatrix}1\\0\end{pmatrix}$$

inside $\mathbb{R}^2$. However, in the next set of notes, the section on inner products will discuss how inner products can make one complement "better" than all the rest.

Given linear maps $A\colon V \to U$ and $B\colon W \to U$, there is a (unique) linear map $A \oplus B\colon V \oplus W \to U$ given by $(A \oplus B)(v, w) = A(v) + B(w)$. I mention this (1) because I reference it in Exercise 9 and (2) because in future notes the theme of constructions of objects being associated to constructions of homomorphisms will be explored much more in-depth.

The rank-nullity theorem

The rank-nullity theorem says that if $A\colon V \to W$ is a linear transformation where $\dim V = n$, then $\operatorname{rank} A + \operatorname{null} A = n$. Here, $\operatorname{null} A = \dim \ker A$, and as mentioned before $\operatorname{rank} A = \dim \operatorname{im} A$. The rank-nullity theorem is the fundamental theorem of elementary linear algebra. In a sense, the Isomorphism Theorems in other parts of algebra can be thought of as the best approximations to the rank-nullity theorem.

Eigen-stuff

This section talks about some basic facts about eigenvalues, eigenvectors, and some associated constructions. Many of these facts make frequent appearance in Qual problems, so you should be comfortable with them.

An **eigenvector** for a square matrix A is a *nonzero* vector v so that A acts on v by scaling it. In an equation, $Av = \lambda v$, for some scalar λ. The number λ is called an **eigenvalue** for A. Eigenvalues turn out to be the most important features of linear transformations of finite-dimensional vector spaces.

(It's important to insist that eigenvectors be nonzero, because $A\ 0 = \lambda \cdot 0$ is true for literally every λ. Whenever a problem involves eigenvectors, be on alert that your purported eigenvectors are nonzero.)

Exercise 7. Prove that the eigenvalues of an upper (or lower) triangular matrix are just the diagonal entries.

Exercise 8. *Without* using its characteristic polynomial, find the eigenvectors and eigenvalues of

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

(Set up the eigenvector problem as a linear system with a variable parameter, and solve that system.)

Exercise 9. Suppose V, W are finite-dimensional vector spaces, and $A\colon V \to V$ and $B\colon W \to W$ are linear transformations. Describe the eigenvalues and eigenvectors of $A \oplus B\colon V \oplus W \to V \oplus W$.

Exercise 10. Two matrices A and A' are **similar** if there exists an invertible matrix B such that $A' = BAB^{-1}$.

(a) Prove that similar matrices have the same eigenvalues.

(b) Use this to show that if B is invertible, then AB has the same eigenvalues as BA. (This holds in general, but I couldn't think of a slick proof for the general case. If you know one, tell me!)

Comment. In the context of group theory, two group elements g and g' are said to be **conjugate** if there exists an element h so that $g' = hgh^{-1}$. Thus, if A, A' are invertible, and we're in the context of talking about the general linear group GL_n, saying A and A' are similar (in the linear algebra sense) is the same as saying they are conjugate (in the group theory sense.) Unfortunately, sometimes people also use the term "conjugate" to mean other things when talking about matrices (such as being complex conjugates). Also, if you're talking about a group smaller than GL_n, like SL_n, two matrices may be similar in the linear algebra sense, but may fail to be conjugate in the smaller group. **Exercise 11.** Find an example. (There are examples in $\mathrm{SL}_2(\mathbb{R})$.)

It's a confusing world, so be careful out there.

The nicest possible matrices are diagonal matrices, because they are very easy to work with. It turns out if a matrix has "enough" eigenvectors, then it is similar to a diagonal matrix. We call such matrices **diagonalizable**.

Exercise 12.

(a) Prove that if v_1 and v_2 are eigenvectors for A corresponding to eigenvalues $\lambda_1 \neq \lambda_2$, then v_1 and v_2 are linearly independent. Imagine in your mind the induction argument that shows that any collection of eigenvectors belonging to different eigenvalues is linearly independent.

(b) Suppose an $n \times n$ matrix A has n linearly independent eigenvectors $v_1, \ldots, v_n$. Let T be the matrix whose columns are the v_i. Prove that $T^{-1}AT$ is a diagonal matrix. (Note that T is invertible since it has linearly independent columns by assumption.)

(c) Write the matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

as $T\Lambda T^{-1}$ for a diagonal matrix Λ and invertible matrix T.

(d) Give an example of a 2×2 matrix which cannot be diagonalized because it only has 1 eigenvalue, and a 1-dimensional eigenspace for that eigenvalue.

Determinants

There are three different definitions of the determinant (that I know of):

1. The determinant of the matrix A is the product of the eigenvalues of A (with multiplicity).
2. The determinant is the unique function $(\mathbb{R}^n)^n \to \mathbb{R}$ (i.e. it takes in n vectors, each of length n, interpreted as the columns of your matrix) that is
   (i) linear in each of the inputs individually;
   (ii) **alternating**, that is, if you switch any two inputs, you multiply the output by -1;
   (iii) and sends the standard basis (identity matrix) to 1.
3. The determinant of the matrix $A = (a_{i,j})_{i,j=1}^n$ is a sum over permutations $\sigma \in S_n$,

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{i,\sigma(i)},$$

   where $\varepsilon : S_n \to \{\pm 1\}$ is the sign homomorphism, which takes value 1 if σ can be written as a product of an even number of transpositions, and value -1 otherwise. (We will see this homomorphism again in group theory.)

If you have time and inclination, you can prove that these are equivalent, but I'm not sure it's an enlightening exercise.

Definition 2 actually is a sort of proposition-definition, because it asserts (1) that there is a function satisfying those three criteria, and (2) that such a function is unique. This follows from the fact that the set of multilinear (condition (i)), alternating functions $(\mathbb{R}^n)^n \to \mathbb{R}$ forms a 1-dimensional vector space. (More generally, the set of multilinear, alternating functions $(\mathbb{R}^n)^k \to \mathbb{R}$ forms a vector space of dimension $\binom{n}{k}$. It's not too hard to prove this, tho there is some technology which can make it easier.)

Exercise 13.

(a) Prove that similar matrices have the same determinant. (Which definition makes this proof one line?)

(b) Fix a matrix A. How does the determinant of A relate to the determinant of matrices obtained from A by doing one of the elementary column operations? (Which definition makes this proof one line?)

(c) Prove that $\det(A) = \det(A^T)$. (Which definition makes this proof one line?)

(d) Prove that a matrix with two identical columns has determinant 0. (Which definition...)

(e) (Important fact to know, but tricky proof. Skip this one and come back to it another time.) Prove that $\det(AB) = \det(A)\det(B)$. (Try thinking about the second definition.)

Exercise 14. Compute the determinant of

$$\begin{pmatrix} 1 & 3 & 4 \\ -1 & -1 & 2 \\ 2 & 6 & 2 \end{pmatrix}.$$

Characteristic polynomials

Saying that $Av = \lambda v$ is the same as saying that v lies in the kernel of $\lambda I - A$. So, to find all the eigenvalues of A, we can look at $\det(xI - A)$. This will be a polynomial of degree n called the **characteristic polynomial** of A. By the preceding, λ is an eigenvalue of A if and only if $\lambda I - A$ has nontrivial kernel, which happens if and only if $\det(\lambda I - A) = 0$, i.e. , if and only if λ is a root of the characteristic polynomial.

Exercise 15. Find the characteristic polynomial of

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

There's one caveat to the above discussion: a linear transformation may not have any eigenvalues/vectors in the field it is defined over, because the characteristic polynomial may fail to factor over that field. Exercise 16 gives an example where a matrix defined over $\mathbb{R}$ (in fact, over $\mathbb{Q}$) fails to have any real eigenvectors/values, while Exercise 17 asks you to show (among other things) that this perverse behavior cannot occur over algebraically closed fields like $\mathbb{C}$.

Exercise 16. Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

(a) Compute the characteristic polynomial of A.

(b) Compute the eigenvalues of A.

(c) Compute the eigenvectors of A inside $\mathbb{C}^2$.

Exercise 17. (August 2014 Problem 3 (a)-(c))

(a) Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $T: V \to V$ be a linear transformation. Show that T has an eigenvector.

(b) Give an example of a nonzero finite-dimensional vector space V over $\mathbb{R}$ and a linear transformation $T: V \to V$ such that T does not have an eigenvector.

(c) Does a linear transformation of an infinite-dimensional vector space over $\mathbb{C}$ have to have an eigenvector? Either prove this is the case, or give an example of a linear transformation of an infinite-dimensional vector space with no eigenvector.

An important fact about the characteristic polynomial is the **Cayley-Hamilton Theorem:** a matrix always satisfies its characteristic polynomial. That means that if I take the characteristic polynomial $x^n + \cdots + c_1x + c_0$, and I "plug in" A , I will get the 0 matrix: $A^n + \cdots + c_1A + c_0I = 0$.

Exercise 18. Try it with the matrices from Exercise 15 and Exercise 16.

Exercise 19. Let V be a finite-dimensional vector space over $\mathbb{C}$, and $A: V \to V$ a nonzero linear transformation.

(a) Give an example of a pair (V, A) such that the only eigenvalue of A is 0.

(b) Prove that the only eigenvalue of A is 0 if and only if A is nilpotent.

Trace

There are two definitions of the trace of a matrix (actually, I know a third, but it is only fit to be known by Masters of the Matrix like Josh Mundinger):

1. The trace of A is the sum of the diagonal entries of A.
2. The trace of A is the sum of the eigenvalues of A.

Again, time and energy permitting, you can choose to prove the equivalence of these two.

Exercise 20.

(a) Prove that the trace is a linear function $\mathbb{R}^{n\times n} \to \mathbb{R}$.

(b) Prove that the trace satisfies the cyclic relation $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA)$.

One neat thing to note is that $\mathrm{Tr}(A)$ is (up to a sign) the coefficient of x^{n-1} in the characteristic polynomial, and $\det(A)$ is (again, up to a sign) the constant coefficient. Among other things, this provides a way to compute the characteristic polynomial of a 2×2 matrix pretty quickly:

$$\mathrm{char}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - (a+d)\lambda + (ad - bc).$$