Lecture Notes for Math 104: Fall 2010 (Week 1)

Jacob Bernstein

CHAPTER 1

First Lecture

We begin by recalling some basic definitions and facts. In the class treat both real $\mathbb{R}$ and complex $\mathbb{C}$ linear algebra. The former is probably more familiar and "physical" though the latter is more general (and becomes necessary for the correct statement of some theorems). Recall $\mathbb{C}$ is the complex numbers, we will review these next lecture. Most aspects of linear algebra that we discuss won't really depend on whether we work over $\mathbb{R}$ or $\mathbb{C}$. However, as Trefethen and Bau usually works over $\mathbb{C}$ we will do so as well. Unless otherwise stated everything we will do holds over $\mathbb{R}$ and we will often illustrate concepts over $\mathbb{R}$ as the geometry. There will be some cases were the distinction matters and we will point these out! Recall that the complex numbers are just the real numbers with an additional "imaginary" number I which we treat like a regular number except $I^2 = -1$. (Note we don't use i as we want to reserve that for other purposes).

1. Vectors

For us a *vector* will be a n-tuple of real or complex numbers. i.e. $\mathbf{v}$ is can be represented by $(v_1, \ldots, v_n)$ for $v_i \in \mathbb{C}$). We will then say $\mathbf{v} \in \mathbb{C}^n$. If all the v_i are real then we have $\mathbf{v} \in \mathbb{R}^n$ is the set of n-dimensional vectors over the reals. We will also say $\mathbf{v}$ is a *real vector*. Rather than write vectors as n-tubles we will always write them as columns.

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

While the difference between tuples and columns is notational there is a more fundamental difference between the two. An n-tuple should really be though of as just an (ordered) list of numbers while the vector has some additional geometric and algebraic meaning. It is a subtle point, but conceptually important to make this distinction.

As we know we can add vectors. Let $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ with

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

then

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}$$

Geometrically, if we have (say) $\mathbb{R}$ vectors this corresponds to laying the vectors end to end. Notice $\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v}$ and we see this geometrically as well.

We may also multiply vectors by a scalar number $\lambda \in \mathbb{C}$

$$\lambda\mathbf{v} := \begin{bmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{bmatrix}$$

Geometrically, for $\mathbb{R}$ vectors this corresponds to stetching the vector by a factor of $|\lambda|$ and reversing direction if $\lambda < 0$. We can also multiply on the left side by a scalar and

$$\mathbf{v}\lambda := \lambda\mathbf{v}.$$

Finally, scalar multiplication interactions with sums in the usual way, namely:

$$\lambda(\mathbf{v}+\mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}.$$

Let us recall some important vectors:

$$0 := \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

the additive identity. This geometrically this corresponds to "doing nothing" and $0+\mathbf{v}=\mathbf{v}+0=\mathbf{v}$ and $\lambda 0 = 0$. The "standard basis vectors"

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad i\text{th slot}$$

If $n = 3$ and we consider real vectors then $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$, $\mathbf{e}_3 = \mathbf{k}$. Note we can always write a vector $\mathbf{v} \in \mathbb{C}^n$ (uniquely) as

$$\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i = \begin{bmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{bmatrix}$$

For $v_i \in \mathbb{C}$ (or $\in \mathbb{R}$ if $\mathbf{v}$ is a real vector). That is the $\mathbf{e}_i$ are a *basis* of $\mathbb{C}^n$ (we will come back to this latter)..

2. Linear Transformations

Vectors are the basic object in linear algebra but are not all that interesting in and of themselves. More interesting is to study transformations of vectors. In the context of linear algebra we restrict attention to *Linear Transformations*. That is transformations that respect the linear structure (i.e. addition and scalar multiplication). More precisely a linear transformation is a function

$$T : \mathbb{C}^n \to \mathbb{C}^m$$

so that $T(\mathbf{v}+\mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$ and $T(\lambda\mathbf{v}) = \lambda T(\mathbf{v})$. This class will mostly consist of studying linear transformations. However, just as we think of a vector as a concrete list, we will also think of a linear transformation as a concrete object– namely as an array of numers called a matrix.

Recall a complex valued matrix is just an $m \times n$ array of complex numbers (when all the entries are real we say it is *real matrix*):

$$A := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & | \cdots & | & \mathbf{a}_n \end{bmatrix}$$

where here $\mathbf{a}_j$ are columns:

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

This suggests we think of the columns as vectors, which we will very often do. We will write $A \in \mathbb{C}^{m\times n}$ to say that A is an $m \times n$ matrix with complex entries and $A \in \mathbb{R}^{m\times n}$ if the entries are real.

How do we go between T and the matrix A that represents it? The easiest way is to see what T does to the standard basis. As $T(\mathbf{e}_j)$ is a vector in $\mathbb{C}^m$ we can expand it in standard basis vectors $\mathbf{e}_i$ of $\mathbb{C}^m$ as:

$$T(\mathbf{e}_j) = \mathbf{a}_i = \sum_{i=1}^{m} a_{ij}\mathbf{e}_i$$

but that is just

$$T(\mathbf{e}_j) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Where this is standard matrix multiplication (see below). Notice that the ith column of A just the image of $\mathbf{e}_i$. By linearity then for arbitrary $\mathbf{v} = \sum_{j=1}^{n} v_j\mathbf{e}_j$ one has

$$T(\mathbf{v}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}v_j\mathbf{e}_i$$

That is

$$T(\mathbf{v}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_m \end{bmatrix}$$

More compactly one can write:

$$T(\mathbf{v}) = \begin{bmatrix}\mathbf{a}_1 & | & \cdots & | & \mathbf{a}_n\end{bmatrix}\begin{bmatrix}v_1\\ \vdots \\ v_i \\ \vdots \\ v_m\end{bmatrix}$$
$$= v_1\mathbf{a}_1 + \ldots + v_n\mathbf{a}_n$$

here the $\mathbf{a}_j = T(\mathbf{e}_j)$ are the columns of the matrix.

As with vectors we will usually just talk directly about a $m \times n$ matrix but just as with should always remember that this is representing some sort of linear transformation.

3. Algebra of matrices

Recall that we can add matrices, multiply them by a scalar and multiply a $m \times n$ matrix on the left by a $n \times k$ matrix. Addition and multiplication by a scalar are unambiguous. But in case you're a little rusty: Set

$$A = \begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn}\end{bmatrix}, B = \begin{bmatrix}b_{11} & \cdots & b_{1n}\\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn}\end{bmatrix}.$$

Then

$$A + B = \begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn}\end{bmatrix} + \begin{bmatrix}b_{11} & \cdots & b_{1n}\\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn}\end{bmatrix} = \begin{bmatrix}c_{11} & \cdots & c_{1n}\\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn}\end{bmatrix}$$

where $c_{ij} = a_{ij} + b_{ij}$. If $\lambda \in \mathbb{C}$ then

$$\lambda A = A\lambda = \begin{bmatrix}\lambda a_{11} & \cdots & \lambda a_{1n}\\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn}\end{bmatrix}$$

Slightly more complicated is matrix multiplication: Let B now be

$$B = \begin{bmatrix}b_{11} & \cdots & b_{1k}\\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk}\end{bmatrix}$$

$$\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn}\end{bmatrix}\begin{bmatrix}b_{11} & \cdots & b_{1k}\\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk}\end{bmatrix} = \begin{bmatrix}a_{11}b_{11} + \ldots + a_{1n}b_{n1} & \cdots & a_{11}b_{1k} + \ldots + a_{1n}b_{nk}\\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \ldots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1k} + \ldots + a_{mn}b_{nk}\end{bmatrix}$$

or more compactly

$$AB = \begin{bmatrix}A\mathbf{b}_1 & | \cdots & | A\mathbf{b}_k\end{bmatrix}$$

here $\mathbf{b}_j$ are the columns of B. Notice that this (perhaps mysterious) formula for matrix multiplication comes from looking at the linear map that arises from the *composition* of two linear maps. Notice also that if we let $C = AB$ and then expanding things out we see that the columns of C can be expressed as linear combinations of the columns of B.

Most of the usual algebraic rules apply except that matrix multiplication is *not* commutative. Just to list them, let $A, B \in \mathbb{C}^{m\times n}, C, D \in \mathbb{C}^{n\times k}, E \in \mathbb{C}^{k\times l}, \lambda \in \mathbb{C}$. Then

$$\begin{aligned} A + B &= B + A \\ \lambda A &= A\lambda \\ \lambda(A + B) &= \lambda A + \lambda B \\ A(\lambda C + D) &= \lambda AC + BD \\ (A + B)C &= AC + BC \\ (AC)D &= A(CD) \end{aligned}$$

Finally, we recall some important matrices. The first is the zero matrix. This is the matrix in each $\mathbb{C}^{m\times n}$ which has all zero entries and which we denote by 0. This corresponds to the linear transformation that sends everything to 0. If $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{l\times m}$ and $C \in \mathbb{C}^{n\times k}$ and $\lambda \in \mathbb{C}$ then

$$\begin{aligned} A + 0 = 0 + A &= A \\ \lambda 0 = 0\lambda &= 0 \\ B0 = 0 &= 0C. \end{aligned}$$

We also introduce the the identity matrix $I \in \mathbb{C}^{m\times m}$ (which we will also denote by Id). This is the matrix

$$I = \begin{bmatrix} \mathbf{e}_1 | & \cdots & |\mathbf{e}_i| & \cdots & |\mathbf{e}_n \end{bmatrix}$$

whose columns are the standard basis. This matrix corresponds to the transformation that preserves all vectors. For $B \in \mathbb{C}^{l\times m}$ and $C \in \mathbb{C}^{m\times k}$ we have

$$\begin{aligned} BI &= B \\ IC &= C \end{aligned}$$

CHAPTER 2

Second Lecture

In this lecture we review some of the properties of complex numbers.

1. Complex Numbers

Let us look at the following equation:

$$x^2 + 1 = 0 \tag{1.1}$$

Naively $x = \sqrt{-1}$ would seem to be a solution. However, for any real $x\mathrm{R}$ $x^2+1 \geq 1$ so this equation can never be solved over the reals. One of the great realizations in mathematics was that in fact there is a solution if we just expand our concept of numbers. Indeed, we can introduce an "imaginary" number I. This is not a real number and is just a symbol. However, by pretending it has all the algebraic properties of a usual number along with satisfying $I^2 = -1$, will lead to a consistent theory. Formally doing so will allow us to say (up to a sign) $I = \sqrt{-1}$.

In practice what this means is that we should be able to mulitply I by any real number a to get a new "imaginary" number Ia. The set of all such numbers is usually refered to as the set of (purely) "imaginary" numbers and is denoted $I\mathbb{R}$ (it is more standard to use i but I want to avoid confusion with indices). We can also add real and complex numbers to get new numbers that (usually) are neither real nor complex that is let $a, b \in \mathbb{R}$ and then $z = a + Ib$ is (for $a, b \neq 0$) neither real nor imaginary. We denote the set of all such numbers by $\mathbb{C}$ and call them complex numbers.

Since we seek an algebraically consistent set of numbers, we must be a little careful how we define various algebraic operations. To add complex numbers we have:

$$(a + Ib) + (c + Id) = (a + c) + I(b + d)$$

and to multiply we have

$$(a + Ib)(c + Id) = (a + Ib)c + (a + Ib)Ib = ac - bd + I(ad + bc).$$

One important operation that is new is complex conjugation. The idea here is to take a complex number $z = a + ib$ and associate its complex conjugate $\bar{z} = a - ib$. The reason we do this is then that $z\bar{z} = a^2 + b^2$ is then always real (and non-negative). Notice that if $z = a + Ib$ then

$$a = \Re z = \frac{z + \bar{z}}{2}$$

and

$$b = \Im z = \frac{z - \bar{z}}{2I}.$$

The fact that $z\bar{z}$ is a non-negative real number makes it tempting to think of it as a length. This we do and define the *modulus* of the complex number z to be

$$|z| = \sqrt{z\bar{z}}.$$

Notice if $z = a + Ib$ for $a, b \in \mathbb{R}$ then $|z| = \sqrt{a^2 + b^2}$. It is straightforward to see that $z = 0$ if and only if $|z| = 0$.

The complex conjugate also gives an unambiguous way to to divide a complex number by a non-zero complex number. Indeed,

$$\frac{z_1}{z_2} = \frac{z_1\bar{z}_2}{|z_2|^2}.$$

The right hand side consists of multiplication of two complex numbers and then division by a real number all of which is straightforward.

We introduced the complex numbers in order to find roots to $x^2+1 = 0$. Indeed, we can now check that I and $-I$ are the (only) two solutions of this equation. It turns out that once one has I all polynomials (even with complex coefficients) have a solution.

THEOREM 1.1. *(Fundamental Theorem of Algebra) Let*

$$p(x) = \sum_{i=0}^{n} a_i x^i$$

be a polynomial of order n (can consider $a_i \in \mathbb{C}$) with $a_n \neq 0$. Then $p(x) = 0$ has exactly n solutions (counting multiplicity) $z_1, \dots, z_n$ in $\mathbb{C}$. That is

$$p(x) = a_n(x - z_1)(x - z_2) \cdots (x - z_n)$$

2. Geometry of Complex Numbers

Real numbers are usual represented as points on a line. What is the right way to think of representing complex numbers geometrically? Notice that $z = a + Ib$ is really just a pair of numbers (a, b). It can also be thought of as a vector:

$$\begin{bmatrix} a \\ b \end{bmatrix}.$$

That is the complex number z can be represented as a 2-dimensional real vector. Notice complex addition is then the same as vector addition, However, complex numbers can be multiplied and there is no clear way to interpret this for vectors.

This graphical representation gives another way to describe a complex number $z = a + Ib$. Namely, we can associate an angle θ and a radius $r \geq 0$ to z so that $z = r\cos\theta + i\sin\theta$. Explictily, $r = |z| = \sqrt{a^2 + b^2}$ and $\tan\theta = b/a$. The number θ is called the *argument* of z and is defined only up to 2π.

This representation is particlurly useful when combined with the imporant fact know as Euler's formula (see below). For $t \in \mathbb{R}$ one has:

$$e^{It} = \cos t + I \sin t$$

More generally, one has that

$$e^{(\mu+I\nu)t} = e^{\mu t}(\cos(\nu t) + I\sin(\nu t)) \qquad (2.1)$$

One way to justify this formual is to note that it ensures that

$$\frac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t}.$$

even when $\lambda \in \mathbb{C}$.

3. Linear Algebra of Complex Numbers

As we we say we can think of a complex number $z = a + Ib$ as a real vector

$$\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$$

We then have 1 corresponding to $\mathbf{e}_1$ and I corresponding to $\mathbf{e}_2$. Where $\mathbf{e}_1, \mathbf{e}_2$ is the standard basis of $\mathbb{R}^2$.

It turns out that many of the natural operations on z as a complex number can be interpreted as a linear transformation on $\mathbf{v}$. We will use this to illustrate some ideas from last time. Lets consider the map $z \to \bar{z}$ that is let C be the function so that $C(z) = \bar{z}$. On vectors this is the map:

$$C : \begin{bmatrix} a \\ b \end{bmatrix} \to \begin{bmatrix} a \\ -b \end{bmatrix}$$

One checks that

$$\begin{bmatrix} a \\ -b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

and so C yields a linear trasformation on $\mathbb{R}^2$ with matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Notice if we apply complex conjugation twice we get back where we started. This corresonds to the fact that the square of the associated matrix is the identity.

Fix a complex number $w = c + Id$ and consider the function $M_w(z) = wz$. As vectors this yields the map

$$M_w : \begin{bmatrix} a \\ b \end{bmatrix} \to \begin{bmatrix} ac - bd \\ ad + bc \end{bmatrix}$$

We can again check that

$$\begin{bmatrix} ac - bd \\ ad + bc \end{bmatrix} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

And hence M_w can be thought of a linear map on $\mathbb{R}^{2\times 2}$ with associated matrix:

$$\begin{bmatrix} c & -d \\ d & c \end{bmatrix}$$

The rules of complex multiplication imply that $M_{w_1+w_2} = M_{w_1} + M_{w_2}$. In particular, if A, A_1 and A_2 are the matrices in $\mathbb{R}^{2\times}$ corresponding to $M_{w_1+w_2}$, M_{w_1} and M_{w_2} (respectively) then one can verify that $A = A_1 + A_2$. Similarly, one has $M_{w_1}(M_{w_2}(z)) = M_{w_1 w_2}(z)$ and so if A, A_1 and A_2 are the matrices in $\mathbb{R}^{2\times}$ corresponding to $M_{w_1 w_2}$, M_{w_1} and M_{w_2} then one checks that $A = A_1 A_2$. An important consequence is then that we can encode a complex number $z = a + Ib$ as a 2×2 real matrix and all the complex algbra directly corresponds to matrix algebra.

Now consider Euler's formula for $w = c + Id$. That is write $w = re^{i\theta}$ with $r \geq 0$ and $\theta \in [0, 2\pi)$. As a consequence $M_w(z) = M_r(M_{e^{I\theta}}(z)) = M_{e^{I\theta}}(M_r(z))$ In particular

$$\begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

Notice we we've broken our matrix into a simple diagonal matrix and a matrix that is also very simple (it is orthogonal). We will generalize this sort factoriztion for arbitrary matrices (where the diagonal will become upper triangular). This is called the QR factorization and is at the heart of all sorts of applications of linear algebra.

4. Functions of a complex variable: NIC

The algebraic properties of complex numbers allow one to define polynomial functions of a complex variable. That is for a polynomial

$$p(x) = \sum_{i=0}^{n} a_i x^i$$

it is clear what $p(z)$ means for $z = a + Ib$ a complex number. For more general functions the same result can be accomplished by using a taylor series expansion (when one exists). For instance if a function f has a taylor series expansion

$$f(x) = \sum_{i=0}^{\infty} a_i x^i$$

on for $|x| < R$ (here R is the radius of convergence). Then the function f makes sense at complex values z so that $|z| < R$ by setting

$$f(z) := \sum_{i=0}^{\infty} a_i z^i.$$

One has to of course check this sum converges in an appropriate sense, but that is beyond the scope of these notes.

The point is that this gives a rigorous justification for the Euler formula. Namely, the Taylor series expansion for e^t is

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}$$

which has infinite radius of convergence. We then have

$$e^{It} = \sum_{n=0}^{\infty} \frac{(It)^n}{n!}$$

But $I^{4k} = 1, I^{4k+1} = I, I^{4k+2} = -1$ and $I^{4k+3} = -I$ so this gives

$$e^{It} = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} + I \sum_{n=1}^{\infty} \frac{(-1)^(n-1) t^{2n-1}}{(2n-1)!}$$

But the Taylor series expansions of of $\cos t$ and $\sin t$ are

$$\cos t = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \quad \sin t \qquad = \sum_{n=1}^{\infty} \frac{(-1)^(n-1) t^{2n-1}}{(2n-1)!}$$

So

$$e^{It} = \cos t + I \sin t$$

as claimed. The formula for a general complex number can be verified in a similar manner.

CHAPTER 3

Third Lecture

We recall the definitions of basic linear algebra concepts such as spans of vectors, linear independence, basis vectos and so on. We will also translate these concepts into corresponding properties of matrices.

1. Basic Linear Algebra

Suppose we have a set of vectors $\mathbf{v}_1, \cdots, \mathbf{v}_k$ in $\mathbb{C}^n$. One very natural question to ask is can we write every vector as a linear combination of the $\mathbf{v}_j$? If yes, one further asks "how many" different ways are there to express the same vector in terms of the$\mathbf{v}_j$. If no, which vectors fail to be so expressable and how many are there.

To start making this precise we define the *span* of a set of vectors $\{\mathbf{v}_j\}$ to be the set of all vectors. $Span(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \left\{\mathbf{w} : \mathbf{w} = \sum_j c_j \mathbf{v}_j\right\}$. This is the largest set of vectors we form from the set $\{\mathbf{v}_j\}$ by only using linear algebra operations. We say the $\mathbf{v}_j$ are *linearly independent* if

$$c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = 0 \iff 0 = c_1 = \ldots = c_k$$

Otherwise we say the $\mathbf{v}_j$ are linearly dependent.

A simple example: The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

are linearly independent in $\mathbb{R}^3$. Their span is the plane $z = 0$. To show this rigrously note that:

$$a\mathbf{v}_1 + b\mathbf{v}_2 = (a+b)\mathbf{e}_1 + (a-b)\mathbf{e}_2 = \begin{bmatrix} a+b \\ a-b \\ 0 \end{bmatrix}$$

For this to equal 0 need $a + b = 0$ and $a - b = 0$. That is $a = 0$ and $b = 0$. Hence the vectors are linearly independent. Checking the spanning property is similar. Notice all this comes down to is solving ansystem of equations.

Another example: The vectors

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

are not linearly independent. Indeed, $\mathbf{v}_3 = 3\mathbf{v}_1 - 3\mathbf{v}_2$. However any pair of the vectors is linearly independent the span of all three vectors is a plane in $\mathbb{R}^3$.

A final example: In general any collection of standard basis vectors $\mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_k}$ in $\mathbb{C}^n$ is linearly independent if $k = n$ then the set spans $\mathbb{C}^n$.

How do we determine if a given set of vectors spans some set? Is linearly independent? As we sketched in the example above this is really a question about systems of linear equations. Consequently, the best method to do this is to turn the question into an equivalent question about matrices. We will then answer the corresponding question for matrices. This will also give additional information. So how do we turn this into a question about matrices? The point is that taking the span span looks like taking lots of matrix multipications. That is let V be the $n \times k$ matrix:

$$V = \begin{pmatrix} \mathbf{v}_1 & | & \cdots & | & \mathbf{v}_k \end{pmatrix}$$

Then let

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$$

then

$$V\mathbf{c} = \sum_{j=1}^{k} c_j \mathbf{v}_j$$

I.e. $\mathbf{w}$ is in the span of the $\mathbf{v}_j$ if and only if there is a vector in $\mathbf{c} \in \mathbb{C}^k$ so that $\mathbf{w} = V\mathbf{c}$. We denote by $Range(V)$ or $R(V)$ the precisely the latter such set. That is let $A \in \mathbb{C}^{m \times n}$ be a matrix we define the range of A, $Range(A) = R(A) \subset \mathbb{C}^m$ by

$$R(A) = \{\mathbf{w} \in \mathbb{C}^m : \mathbf{w} = A\mathbf{c} \text{ for some } \mathbf{c} \in \mathbb{C}^n\}$$

We also refer to $R(A)$ this as the *Image* of A or the *Column space of* A. Notice the latter makes sense as $R(A)$ is the span of the columns of A.

Similarly, we can use V to see when the $\mathbf{v}_j$ are linearly independent. Again we have the $\mathbf{v}_j$ linearly independent if

$$0 = \sum_j c_j \mathbf{v}_j = V\mathbf{c} \iff \mathbf{c} = 0$$

That is $\{\mathbf{v}_j\}$ is linearly independent if and only if the only vectors that A multiplies against to give 0 is the zero vector. For $A \in \mathbb{C}^{m \times n}$, we define $Null(A) = N(A) \subset \mathbb{C}^n$, the null space of A, to be the set of vectors

$$N(A) = \{\mathbf{v} \in \mathbb{C}^n : A\mathbf{v} = 0\}$$

This is sometimes refered to as the *kernel* of A. Hence we see that the $\mathbf{v}_j$ are linearly indepependent when and only when $N(V) = \{0\}$.

Having transformed the problem to a question about matrices we need to discuss how to use this to actually solve the problem. The main way to do this is to use Gaussian elimination. We'll review this algorithm next lecture.

2. Spaces of Vectors and Basis Vectors

In order to make some of the preceding clearer, we introduce some further mathematical definitions as well as state some important facts. The later will be proved in a couple of lectures after we have some important tools.

For a matrix $A \in \mathbb{C}^{m \times n}$ we have defined the range of A, $R(A)$ as a subset of $\mathbb{C}^m$ and the null space, $N(A)$ as a subset of $\mathbb{C}^m$. These sets are well behaved with respect to the operations of linear algebra. More precisely, we say a set of vectors

$E \subset \mathbb{C}^n$ is a *vector space* (or *vector sub-space*) if for any pair of vectors $\mathbf{v}, \mathbf{w} \in E$ and any scalar $\lambda \in \mathbb{C}$ one has $\lambda\mathbf{v} \in E$ and $\mathbf{v} + \mathbf{w} \in E$. That is E is closed uner the operations of linear algebra. Notice by taking $\lambda = 0$ we must alwasy have $0 \in E$. Note that $\mathbb{C}^n$ is a vector space as is $\{0\}$ the set consisting only of the zero vector $0 \in \mathbb{C}^n$.

Notice that the span of any set of vectors $vV_i \in \mathbb{C}^n$, $1 \leq i \leq k$ is a vector space. To see this let $\mathbf{v}, \mathbf{w} \in Span(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ then $\mathbf{v} = \sum_{i=1}^k c_i \mathbf{v}_i, \mathbf{w} = \sum_{i=1}^k d_i \mathbf{v}_i$. Then using the algebraic properties of vectors one has $\mathbf{v} + \mathbf{w} = \sum_{i=1}^k (c_i + d_i)\mathbf{v}_i$ which is then clearly in the span. A similar argument shows $\lambda\mathbf{v}$ is in the span. In particular, as the range space of A is the span of the columns of A, $R(A) \subset \mathbb{C}^m$ is a vector space. The null space of A, $N(A) \subset \mathbb{C}^n$ is also a vector space. To see this let $\mathbf{v}, \mathbf{w} \in N(A)$. Then $A\mathbf{v} = A\mathbf{w} = 0$. Now $A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w} = 0$ and $A(\lambda\mathbf{v}) = \lambda(A\mathbf{v}) = \lambda 0 = 0$, hence $\mathbf{v} + \mathbf{w} \in N(A)$ and $\lambda\mathbf{v} \in N(A)$.

For a given vector space E we say that a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in E$ are a *basis* of E if $E = Span(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ and the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent. An easy example is that that the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{C}^n$. One important consequence of $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ being a basis of $E \subset \mathbb{C}^n$ is that any vector $\mathbf{v} \in E$ can be expressed uniquely in terms of the $\mathbf{v}_i$ that is

$$\mathbf{v} = \sum_{i=1}^k c_i \mathbf{v}_i = V\mathbf{c}$$

where here $V \in \mathbb{C}^{n \times k}$ is the matrix with columns the $\mathbf{v}_i$ and $\mathbf{c} \in \mathbb{C}^k$ are the *coefficients* of $\mathbf{v}$ with respect to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$. It is a simple excercise to see that $N(V) = \{0\}$ and $R(A) = E$. In particular, if we want to find the coefficients $\mathbf{c}$ for a given vector $\mathbf{v}$ we must solve the equation:

$$V\mathbf{c} = \mathbf{v}.$$

Which consists of n equations in k unknowns.