**Lecture I**

**I. VECTOR SPACES**

For the moment, I'll postpone discussion of the row space and the left nullspace, until the end of the week when this discussion becomes more natural.

First, I want to remind you how to express a linear transformation as a matrix.

**Procedure:** Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, we only need to know how $T$ acts on the *basis* $\vec{e}_1, \dots, \vec{e}_n$. To wit, the desired matrix $A$ is given as

$$A = \left[ \begin{array}{ccc} | & | & | \\ T(\vec{e}_1) & \dots & T(\vec{e}_n) \\ | & | & | \end{array} \right].$$

**Example:** Suppose the transformation is $T : \mathbb{R}^2 \to R^2$ given by "reflect through the line $y = x$, then reflect through the $x$-axis." To determine the corresponding $2 \times 2$ matrix, it is enough to compute $T(\vec{e}_1)$ and $T(\vec{e}_2)$. One readily sees that

$$T(\vec{e}_1) = -\vec{e}_2, T(\vec{e}_2) = \vec{e}_1$$

So the corresponding matrix is

$$\left[ \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right]$$

which you might note is identical to the matrix representing "rotation by 90 degrees clockwise." In other words, the product of two reflections is a rotation–not so obvious geometrically.

Now, on to a sort-of definition.

**"Definition"** A *vector space* is any set of objects which can be added together and multiplied by scalars.

This is not very precise, and to make it precise would make it rather unreadable. But see exercise 2.1.5 in Strang for the precise definition. The most important thing is to have at our disposal a population of examples.

**Examples:**

- $\mathbb{R}^n$

- any subspace of $\mathbb{R}^n$

- the set of polynomials

- the set of polynomials of degree at most 2

- the set of all $2 \times 2$ matrices

The game is that all the basic notions of linear algebra apply just as well to *any* vector space as they do to $\mathbb{R}^m$. For instance, basis, dimension, nullspace, column space....

**LECTURE 2**

*Example*

Let $P_2$ be the space of polynomials of degree at most 2.

**Claim 1**: A basis for this space is $\{1, x, x^2\}$.

They are evidently linearly independent, and every polynomial of degree at most 2–more or less by definition–can be expressed as some combination of these three guys. (Write more on the board–this may be a subtle point.) This also means that $P_2$ has dimension 2.

Now let $T : P_2 \to P_2$ be the map defined by

$$T(f) = f' - f$$

**Claim 2:** $T$ is a linear transformation.

To check:

- $D(f + g) = D(f) + D(g)$

- $D(af) = aD(f)$

Both are straightforward. To check the first, for instance, just check

$$D(f + g) = (f + g)' - (f + g) = f' + g' - f - g = (f' - f) + (g' - g) = D(f) + D(g)$$

as desired.

**Claim 3:** $T$ can be expressed as a matrix.

We would like to do this just as before. But the problem is, we don't have the $\vec{e}_i$! But this only *appears* to be a problem.

The fact of the matter is: the only important thing about the $\vec{e}_i$ is that they are a basis for $\mathbb{R}^n$. And we have a basis for our space $P_2$. So we'll use that basis in place of the $\vec{e}_i$.

Let $\vec{v}_1 = 1, \vec{v}_2 = x, \vec{v}_3 = x^2$ be our basis. If $f$ is the polynomial $ax^2 + bx + c$, we can write $f = c\vec{v}_1 + b\vec{v}_2 + a\vec{v}_3$. Or, more compactly, we can define

$$[f] = \begin{bmatrix} c \\ b \\ a \end{bmatrix}$$

**MOTTO:** The choice of a basis allows us to write each element of the vector space as a column vector.

In fact, this choice allows us to make a one-to-one association between elements of $P_2$ and elements of $\mathbb{R}^3$. In the math biz, we would say we have exhibited an *isomorphism* between $P_2$ and $\mathbb{R}^3$. Note also that $\vec{v}_i$ corresponds to $\vec{e}_i$.

Now we would like to represent $T$ by a matrix; that is, we would like to find $A$ such that

$$[T(f)] = A[f]$$

(this is an equality of *vectors.*)

Just as in the case treated earlier, it is true that we can find such a matrix: it is given by

$$A = \left[ \begin{array}{ccc} | & | & | \\ [T(\vec{v}_1)] & [T(\vec{v}_2)] & [T(\vec{v}_3)] \\ | & | & | \end{array} \right].$$

And we have

$$[T(\vec{v}_1)] = [1' - 1] = [-1] = \left[ \begin{array}{c} -1 \\ 0 \\ 0 \end{array} \right]$$

and

$$[T(\vec{v}_2)] = [x' - x] = [1 - x] = \left[ \begin{array}{c} 1 \\ -1 \\ 0 \end{array} \right]$$

and

$$[T(\vec{v}_3)] = [(x^2)' - x^2] = [2x - x^2] = \left[ \begin{array}{c} 0 \\ 2 \\ -1 \end{array} \right].$$

To sum up, $T$ can be represented by the matrix

$$A = \left[ \begin{array}{ccc} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{array} \right].$$

3

Observe that $A$ has rank 3; it follows that $A$ is invertible (by Rank Theorem). It follows that

$$A\vec{x} = \vec{y}$$

has a unique solution for every $y$. Translating from vectors to polynomials, we are saying that

$$T(f) = g$$

has a unique solution for every polynomial $g \in P_2$; that is, the differential equation

$$f' - f = g$$

has a unique solution. This solution can be computed explicitly by finding the inverse matrix of $A$. So we've obtained a theorem about the uniqueness of the solution to a differential equation by means of linear algebra!

**Remark:** Even more fun, we might try to look at this problem again using a *different* basis... but that's another story.

### LECTURE 3

**I. Orthogonality**

I want to talk about the transpose now.

**Recall:** If $A$ is an $m \times n$ matrix, then $A^T$ is the $n \times m$ matrix obtained by "writing $A$ sideways."

The transpose has everything to do with the idea of dot product. For instance: let $\vec{v}$ and $\vec{w}$ be two vectors. Then

$$\vec{v}^T \vec{w} = [\vec{v} \cdot \vec{w}].$$

**Definition:** Let $\vec{v}, \vec{w} \in \mathbb{R}^n$. We say $\vec{v}$ and $\vec{w}$ are *orthogonal* if (and only if)

$$\vec{v}^T \vec{w} = 0.$$

(note abuse of notation; I write 0 instead of [0].)

The *magnitude* of $\vec{v}$ is $\sqrt{\vec{v}^T \vec{v}}$. It is denoted $|\vec{v}|$. A *unit vector* is a vector of magnitude 1.

(note abuse of notation; I write 0 instead of [0].)

**Example:** (Pythagorean Theorem)

4

Draw the right triangle with edges $\vec{v}, \vec{w}, \vec{v} + \vec{w}$. Then point out that Pythagorean Theorem says $|\vec{v} + \vec{w}|^2 = |\vec{v}|^2 + |\vec{w}|^2$. Check

$$|\vec{v} + \vec{w}|^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v} \cdot \vec{v} + 2\vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w} = \vec{v} \cdot \vec{v} + \vec{w} \cdot \vec{w}.$$

**Example:** What is the set of all vectors orthogonal to $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$?

Well, let $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ be the relevant vector. Then

$$[125] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 + 2x_2 + 5x_3$$

So the set of all vectors orthogonal to $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$ is

$$V = \{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : x_1 + 2x_2 + 5x_3 = 0 \}.$$

**Remarks:**

- $V$ is a subspace.

- In fact, $V$ is the nullspace of the matrix

$$[125].$$

**Example:** What about the set of all vectors $\vec{x}$ orthogonal to $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$ and

$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$? Suppose that $\vec{x}$ has this property. Then evidently

$$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \end{bmatrix} \vec{x} = \vec{0}$$

That is, $\vec{x}$ is in $N(A)$, where $A$ is the $2 \times 3$ matrix above. In fact, the reverse is true: so the set of vectors we're interested in is just $N(A)$.

Note that $N(A)$ is 1-dimensional, because the rank of $A$ is 2.

**Definition:** Let $V$ be a subspace of $\mathbb{R}^n$. Then the *orthogonal complement* of $V$, denoted $V^{\perp}$, is the set of vectors which are orthogonal to every vector in $V$.

**Fact:** $V^{\perp}$ is a subspace of $\mathbb{R}^n$.

**Examples:** Draw a line in $R^3$, observe that the orthogonal complement is a plane. Likewise, draw a plane and observe that the orthogonal complement is a line. Now ask: what is orthogonal complement of $\mathbb{R}^3$ itself?

Ask: how does the dimension of the space relate to the dimension of its orthogonal complement? We should get to the conjecture that the sum of the two dimensions is equal.

**Theorem.** $\dim V + \dim V^{\perp} = n$.

*Proof.* (Not so formal as many proofs I give–can you find a more precise one?)

Idea: this looks a lot like the Rank Theorem. So we should try to find a matrix that has $V^{\perp}$ as nullspace and $V$ as column space.

Consider the linear transformation "orthogonal projection onto $V$." Let $A$ be a matrix representing this transformation. Now $A\vec{x} = \vec{y}$ has a solution if and only if $\vec{y} \in V$; conclude that $\mathcal{R}(A) = V$. On the other hand, $A\vec{x} = 0$ if and only if $\vec{x} \in V^{\perp}$; thus, $N(A) = V^{\perp}$. Then the desired theorem follows by the Rank Theorem. $\square$

**Fact.** $(V^{\perp})^{\perp} = V$.

Proof left to the reader–see Strang p.139. (Hint: you'll need to use the above theorem.)

**II. The other two spaces**

Let $A$ be an $m \times n$ matrix, say

$$A = \begin{bmatrix} \cdots & \vec{w}_1 & \cdots \\ \cdots & | & \cdots \\ \cdots & \vec{w}_m & \cdots \end{bmatrix}.$$

For any $\vec{x}$, we have

$$A\vec{x} = \begin{bmatrix} \vec{w}_1 \cdot \vec{x} \\ \vdots \\ \vec{w}_m \cdot \vec{x} \end{bmatrix}.$$

Suppose $\vec{x} \in N(A)$. Then the following statements are equivalent:

- $\vec{x} \in N(A)$;

- $\vec{w}_i \cdot \vec{x} = 0$ for all $i$;

- $\vec{x}$ is in the orthogonal complement of the space $W = \text{span}\{\vec{w}_1, \ldots, \vec{w}_m\}$.

Only the last requires comment. Every vector $\vec{w} \in W$ can be written as

$$\vec{w} = c_1\vec{w}_1 + \ldots + c_m\vec{w}_m.$$

Now

$$\vec{w} \cdot \vec{x} = c_1\vec{w}_1 \cdot \vec{x} + \ldots + c_m\vec{w}_m \cdot \vec{x} = 0.$$

So we have shown that $N(A)$ consists of those vectors which are orthogonal to everything in $W$; in other words,

$$N(A) = W^{\perp}$$

The space $W$ is pretty important and gets its own name.

**Definition:** The *row space* of $A$ is the space spanned by the rows. It can also be described as $\mathcal{R}(A^T)$.

We've thus proved:

**Theorem:** The nullspace of $A$ is the orthogonal complement to the row space of $A$ (and vice versa.)

This has the following remarkable corollary.

**Corollary:** $\text{rank}(A^T) = \text{rank}(A)$.

The proof is so concise it almost seems not to be there.

- $\text{rank}(A^T) = \dim\mathcal{R}(A^T)$ (by Rank Theorem.)

- $\dim\mathcal{R}(A^T) = n - \dim N(A)$ (by above theorem, and $\dim V + \dim V^{\perp} = n$.)

- $n - \dim N(A) = \dim\mathcal{R}(A)$ (by Rank Theorem again)

- $\dim\mathcal{R}(A) = \text{rank}(A)$ (by Rank Theorem yet again!)

You might have fun trying to prove the corollary above just using the definition of rank as the number of nonzero pivots after Gaussian elimination. You can do it–but it's a mess! I hope this helps you appreciate the utility of having a good theoretical language.

**Definition:** The *left nullspace* of $A$ is defined to be $N(A^T)$.

**Fact:** The left nullspace is the orthogonal complement of $\mathcal{R}(A)$.

**Fact:** The left nullspace will not be very important for us, despite Strang's protestation on p.95.