MA 1B RECITATION 01/22/15

1. Feedback

We have been using the following terminology for linear transformations and such without explanation, but it seems like not everybody is familiar with it, so let's quickly recall the definitions.

We say that a map of sets $f : A \to B$ is

- **one-to-one** or **injective** if $f(x) = f(y)$ implies that $x = y$;
- **onto** or **surjective** if $f(A) = B$, that is, for all $b \in B$, there exists an $a \in A$ that $f(a) = b$;
- a **one-to-one correspondence** or **bijective** if f is both injective and surjective.

Since vectors spaces are also sets, we use these terms to talk about linear transformations. However, the connection is slightly deeper when we have the structure of linear maps, in particular, we have the following result.

Fact 1. If $T : V \to W$ is a linear transformation between vector spaces of the same *finite* dimension, then the following are equivalent:

(a) T is injective
(b) T is surjective
(c) T is bijective
(d) T has nullspace $\{0\}$
(e) T is invertible.

In other words, one property holding implies all the other properties.

Remark 2. As I mentioned in the first recitation, there is really only one big theorem in the whole course. The fact above is just a part of that big theorem.

It seems like many people are having some trouble with the more abstract arguments in the homeworks, so we're going to go over some examples of how to reason with such things.

For example, one thing that tripped a lot of people up on the homework was how to show that vectors are linearly independent, so we'll go through an example of such a proof.

Proposition 3. *If $v_1, \ldots, v_n$ are linearly independent elements a vector space V and $T : V \to W$ is a one-to-one linear transformation, then $T(v_1), \ldots, T(v_n)$ are linearly independent in W.*

Proof. Suppose that we have $a_i \in \mathbf{R}$ such that

$$a_1 T(V_1) + \cdots + a_n T(v_n) = 0.$$

Date: January 22, 2015.

We want to show that $a_1 = a_2 = \cdots = a_n = 0$. We have

$$\begin{aligned} 0 &= a_1T(v_1) + \cdots + a_nT(v_n) \\ &= T(a_1v_1) + \cdots + T(a_nv_n) \\ &= T(a_1v_1 + \cdots + a_nv_n). \end{aligned}$$

Since T is one-to-one, the only element of V that maps to $0 \in W$ must be $0 \in V$, and so

$$a_1v_1 + \cdots + a_nv_n = 0.$$

Since $v_1, \ldots, v_n$ are linearly independent, this implies that $a_1 = a_2 = \cdots = a_n = 0$, as desired. □

We've done a lot of work with deducing properties of specific linear transformations, and with abstract maps, but haven't really talked much about existence or nonexistence results. Here's an example of such a result that we can prove with the tools we currently have available.

Example 4. There does not exist a linear map $\mathbf{R}^5 \to \mathbf{R}^2$ whose nullspace equals

$$\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^2 \mid x_1 = 3x_2, x_3 = x_4 = x_5\}.$$

Why? It is spanned by $(3, 1, 0, 0, 0)$ and $(0, 0, 1, 1, 1)$. Thus, the rank-nullity theorem does not hold. This is the idea, but how would one write a proof of this? The easiest way is via proof by contradiction.

Proof. Suppose that there exists such a map $T : \mathbf{R}^5 \to \mathbf{R}^2$. Since the image is in $\mathbf{R}^2$, the rank of T is at most 2. The nullspace is spanned by $(3, 1, 0, 0, 0)$ and $(0, 0, 1, 1, 1)$. (Self-test: How can you show that the nullspace is the span of these vectors?) Thus, the nullity is 2. However, if T is a linear map, by the rank-nullity theorem, we must have

$$5 = \text{rank}(T) + \text{nullity}(T).$$

However, $\text{rank}(T) + \text{nullity}(T) \leq 2 + 2 < 5$, which contradicts the rank-nullity theorem. Thus, there cannot exist such a map T. □

2. Linear Transformations in Matrix Form

Linear transformations can be represented and analyzed using matrices. Indeed, this was why matrices were developed in the first place!

We have done some of this already, but let's get a little more practice.

Example 1. Let $X = \{x_1, x_2\}$ a basis for a 2-dimensional vector space V and $T \in \mathcal{L}(V)$ with

$$\begin{aligned} T(x_1) &= x_1 + 2x_2 \\ T(x_2) &= 2x_1 - x_2. \end{aligned}$$

For $f \in \mathcal{L}(V)$, let $m(f) = m_X(f)$ be the matrix of f with respect to X.

We first calculate $m(T)$ and $m(T^2)$:

$$m(T) = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

and since $m(T^2) = m(T)^2$ (Self-check: Why?), we have

$$m(T^2) = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}.$$

Now, consider a different basis $Y = \{y_1, y_2\}$, where

$$y_1 = x_1 + x_2$$
$$y_2 = x_1 + 2x_2.$$

Q: What is $m_Y(T)$ and $m_Y(T^2)$?

The formula for the matrix representation is given by

$$m_Y(T) = B^{-1}m(T)B$$

where $B = m(g)$ for g the unique (Why?) linear map such that $g(x_i) = y_i$. This B is often called the *base change matrix*, for obvious reasons. Since $g(x_1) = y_1 = x_1+x_2$ and $g(x_2) = x_1 + 2x_2$, we have

$$B = m(g) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

What about B^{-1}? Usually it's a quite involved process to calculate B^{-1}, but luckily we're in the simple 2×2 case, so there is a simple formula:

$$B^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} m_Y(T) = B^{-1}m(T)B &= \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 5 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 10 \\ -2 & -5 \end{bmatrix} \end{aligned}$$

and

$$m_Y(T)^2 = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

Lastly, let's check that the calculation is correct, on say, $y_1 = x_1 + x_2$: with respect to Y we have

$$\begin{bmatrix} 5 & 10 \\ -2 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

and with respect to X we have

$$\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

and indeed

$$T(y_1) = 5y_1 - 2y_2 = 5(x_1 + x_2) - 2(x_1 + 2x_2) = 3x_1 + x_2.$$

Remark 2. One way to remember how to represent the base change matrix is to remember the columns correspond to things in this form, so if $X = \{x_1, \dots, x_n\}$ is your basis,

$$m(T) = m_X(T) = [T(x_1), T(x_2), \dots, T(x_n)].$$

where the $T(x_i)$ are column vectors. To remember this, think of T acting on the "column vector" given by $[x_1, \dots, x_n]^T$. But remember this is only a mnemonic!

Also, the matrix representative necessarily depends on the basis, so always state explicitly which basis you are using to represent the transformation, even if it's just a short remark saying "under the standard basis."

2.1. **Preview: Some bases are better than others.** We had something interesting happening in the last example. For instance, we saw that the matrix representations of T^2 with respect to both bases X and Y were the same. It turns out this is not just a coincidence and there is actually something deeper going on.

This raises the following question. Suppose we have a linear transformation T. Is there any way to choose the "best" basis?

What do we mean by best? Just the basis that makes the problem easy to analyze.

For instance, suppose we have a transformation T represented by the following matrix under the standard basis as

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & -4 & 2 \end{bmatrix}.$$

Looks pretty un-uniform and general, and I can't analyze its properties too easily. For instance, if I was captured by an evil mathematician that told me to calculate, say, A^{76} by hand, or die, my life would probably start flashing before my eyes, since I'd inevitably make some arithmetic mistake.

But it turns out that this isn't so tough. Consider the basis $Y = \{y_1, y_2, y_3\}$ given by

$$\begin{aligned} y_1 &= -x_1 + -x_2 + 2x_3 \\ y_2 &= x_3 \\ y_3 &= -x_1 + 2x_3. \end{aligned}$$

Then

$$B = \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix}, \text{ and so } B^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}.$$

However, we then have

$$B^{-1}AB = \begin{bmatrix} 0 & -1 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & -4 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So calculating A^{76} is actually quite easy! Who says linear algebra has no practical application? It can save your life!

But you might say, hey, I guess you were quite lucky to have an math angel on your shoulder tell you this magical basis. But what if I told you that you, too, can do this, without the interference of ethereal beings?

Finding such a basis is something that we will learn about in the second part of the course, but it's worth thinking a bit by yourself how you might go about finding such a basis. It's actually related to the idea of invariants that I mentioned in the last recitation.

What are the "intrinsic" properties of the transformation? That is, which properties of the transformation are the same regardless of the matrix representation we choose for it? It turns out that this will be the secret to finding the "best" basis.

3. Invariants of Linear Transformations

We will delve into this topic more deeply in the next couple of weeks, but we will begin with some easy properties.

3.1. **Rank of a matrix.** We know what the rank of a linear transformation is. What is the rank of a matrix?

For instance, what's the rank of

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}?$$

They are 1, 3, and 2, respectively. We could have calculated them by looking at the linear transformations they define, but there is an even faster way.

Recall the following definitions and results from matrix theory. The **row rank** of a matrix is the dimension of its **row space**, the span of its rows. Similarly, the **column rank** of a matrix is the dimension of the **column space**, the span of its columns. A fundamental fact is that

$$\text{column rank} = \text{row rank},$$

and so we can just refer to the **rank of a matrix** as this number.

Fact 1. The rank of a linear transformation equals the (matrix) rank of any of its matrix representations.

In other words, the rank of a linear transformation is independent of the matrix representation that you choose and so is an "intrinsic" property.