MA 1B RECITATION 01/24/13

1. Big Picture: A Classification of Linear Transformations

What does a linear transformation "look like"?

It turns out that there is a one-to-one correspondence between linear transformations $T : \mathbf{R}^n \to \mathbf{R}^m$ and $m \times n$ matrices. This is one of the major conceptual milestones of this course.

For example, for linear maps $\mathbf{R}^2 \to \mathbf{R}^2$, all of them are of the form

$$\begin{aligned} T : \mathbf{R}^2 &\to \mathbf{R}^2 \\ (x, y) &\mapsto (ax + by, cx + dy) \end{aligned}$$

which corresponds to the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Thus, these linear maps have a similar shape, namely, they correspond to a scaling or "shearing" operation that you can visualize by say, fixing a unit square with lower-left corner at $(0, 0)$ and seeing what happens to the square when you let T act on $\mathbf{R}^2$.

However, this is not the complete truth. I've implicitly made an assumption about choices of bases. Namely, I've assumed that both copies of $\mathbf{R}^2$ have the same basepoint (the origin 0) and the same bases x and y, corresponding to the standard basis $e_1 = (1, 0)$ and $e_2 = (0, 1)$.

We could also suppose that the second copy of $\mathbf{R}^2$ has a different basepoint, say, $(3, 1)$ instead of the $(0, 0)$, and choose bases x' and y' *relative to the basepoint*, so x' is given by the vector $(3, 1) \to (4, 1)$ and y' is given by the vector $(3, 1) \to (3, 2)$. To distinguish this slightly modified copy, let's call this $\widetilde{\mathbf{R}^2}$. This is still a two-dimensional vector space and so is isomorphic to $\mathbf{R}^2$. Since linear maps must map the origin to origin, the linear maps $\mathbf{R}^2 \to \mathbf{R}^2$ are of the form

$$\begin{aligned} T : \mathbf{R}^2 &\to \widetilde{\mathbf{R}^2} \\ (x, y) &\mapsto (ax' + by' + 3, cx' + dy' + 1) \end{aligned}$$

We see that now we have *translation* by $(3, 1)$ as a valid map from $\mathbf{R}^2$ to a copy of $\mathbf{R}^2$. Thus, if we fix the same basepoint and bases, a translation is *not* a linear transformation $\mathbf{R}^2 \to \mathbf{R}^2$, but if we allow ourselves to change the basepoint, we allow for translations.

Now, when we usually talk about maps $\mathbf{R}^2 \to \mathbf{R}^2$, we'll assume they have the same basepoints for simplicity, but maybe you're wondering, if we give ourselves the freedom to change basepoints, does that preserve anything about, say, lines or shapes in the plane? Or is it completely wild?

It turns out that it does in fact preserve a lot of information. If we allow ourselves to change basepoints, we get a class of maps called *affine transformations* that give us all the ways we can "transform" $\mathbf{R}^2$ that preserves straight lines and ratios of distances between points lying on a straight line. Such maps don't necessarily

Date: January 24, 2013.

preserve angles or lengths, but many things do hold, like points lying on lines will still remain on the line after transformation, midpoints of line segments will remain midpoints of that line segment after transformation, and parallel lines will remain parallel after transformation.

In the language of linear algebra, we can describe affine transformations very concisely: they are merely linear transformations followed by a translation.

But why do we care about these weird invariant properties? It turns out that they can have a lot of scientific meaning. For instance, results like Noether's theorem in physics says that invariants are a sign of symmetry, and can point to the fact that seemingly complicated phenomena in the world can admit a surprisingly simple description if you look at the problem the right way.

2. Terminology

First a quick note on terminology. Given sets A and B:

- a one-to-one (1-1) *mapping* $f : A \to B$ is map where any two distinct elements in A map to distinct elements in B.
- a one-to-one (1-1) *correspondence* $f : A \to B$ is a one-to-one mapping whose image is all of B. That is, each element in A "corresponds" to a unique element in B.

Remark 1. These are called injections and bijections, respectively, by mathematicians.

3. Inductive Proofs

You've seen these before, especially if you've taken any sort of discrete math class, but it is good to review since it's a useful tool and allows for a clear solution to one of your homework problems.

This is a method for proving a statement P for all natural numbers n and higher. There are two steps involved:

(a) The *base case*, checking the lowest case (usually $n = 1$).
(b) The *induction/inductive step*, where you *assume* P is true for a given n (the *induction hypothesis*) and deduce it for the $n + 1$ case.

Remark 1. This is not circular reasoning, despite the fact that you assume the statement is true to prove it for the next case. Think of the induction step as setting up a chain of dominos, and your proof of the base case as the flick that allows you to knock them all down and prove your statement.

Here's a result that admits an argument by induction. Indeed, most non-induction proofs of this result are not usually rigorous.

Proposition 2. *Let A be an invertible matrix. For all n, $(A^n)^{-1} = (A^{-1})^n$.*

Proof. We will prove this by induction on n.

Base case: For $n = 1$, we have $(A^1)^{-1} = A^{-1} = (A^{-1})^1$ so our base case holds.

Inductive step: Assume that $(A^k)^{-1} = (A^{-1})^k$. We want to show that $(A^{k+1})^{-1} = (A^{-1})^{k+1}$. We have

$$A^{n+1}(A^{-1})^{n+1} = AA^n(A^{-1})^n A^{-1} = AIA^{-1} = I$$

where the second equality uses the induction hypothesis. Therefore, $(A^{-1})^{k+1}$ is the inverse of A^{k+1}.

By the principle of mathematical induction, our statement is true for all $n \geq 1$. $\square$

As a challenge problem, try the following crazier result, which admits a surprisingly simple proof by induction on n.

Proposition 3. *Let A be an $n \times n$ matrix whose entries are only 1 and -1. Then 2^{n-1} divides $\det(A)$.*

4. Analyzing linear maps

Since one of our homework problems involves linear maps on the polynomial space $F[x]$, let's get a little more practice with this.

Let V be the linear space of all real polynomials $p(x)$. Let R, S, T be functions which map an arbitrary polynomial $p(x) = c_0 + c_1 x + \cdots + c_n x^n$ in V onto the polynomials $r(x), s(x)$, and $t(x)$, respectively, where

$$r(x) = p(0), \quad s(x) = \sum_{k=1}^{n} c_k x^{k-1}, \quad t(x) = \sum_{k=0}^{n} c_k x^{k+1}.$$

Question 1. *Are the maps R, S, and T linear? What are their null space and range?*

To prove things are linear, we need two elements and two constants, so let $q(x) = \sum_i d_i x^i \in V$ and $a, b \in \mathbf{R}$. Write $ap + bq = \sum_i u_i x^i$ where $u_i = ac_i + bd_i$.

We have $R(p) = p(0) = c_0$, so

$$R(ap + bq) = u_0 = ac_0 + bd_0 = aR(p) + bR(q)$$

so $R \in \mathcal{L}(V)$.

Now, $S(p) = (p - c_0)/x$, so

$$\begin{aligned} S(ap + bq) &= \frac{(ap + bq) - (ac_0 + bd_0)}{x} = \frac{a(p - c_0) + b(q - d_0)}{x} \\ &= a\frac{p - c_0}{x} + b\frac{q - d_0}{x} = aS(p) + bS(q) \end{aligned}$$

so $S \in \mathcal{L}(V)$.

Finally, $T(p) = xp$, so

$$T(ap + cq) = x(ap + bq) = axp + bxq = aT(p) + bT(q),$$

so $T \in \mathcal{L}(V)$ as well.

Before we answer the question about nullspace and range, it's best to prove the following lemma.

Lemma 2. *Let U be a vector space and $f, g \in \mathcal{L}(U)$ with $f \circ g = \mathrm{id}_U$. Then $N(g) = 0$ and $f(U) = U$.*

Proof. For $u \in U$,

$$u = \mathrm{id}_U(u) = (f \circ g)(u) = f(g(u)),$$

so $u \in f(U)$ and hence $f(U) = U$. Similarly, if $u \in N(g)$ (note this is nonempty because 0 is always in the nullspace), then $g(u) = 0$, so $u = f(g(u)) = f(0) = 0$. Hence, $N(g) = 0$. $\square$

Now it turns that we can answer an associated question at the same time.

Question 3. *The map T is one-to-one on V. Can you show that $N(T) = 0$ and find an $f \in \mathcal{L}(V)$ such that $f \circ T = \mathrm{id}_V$?*

Let $f, g \in \mathcal{L}(V)$ and write fg for $f \circ g$ and I for id_V, mimicking matrix notation[1]. Then $(ST)(p) = S(T(p)) = S(xp)$ and $(xp)(0) = 0$, so

$$S(xp) = (xp - (xp)(0))/x = xp/x = p.$$

In other words, $(ST)(p) = p$ for all $p \in V$, so $ST = I$. Therefore, by the lemma, $N(T) = 0$ and $S(V) = V$.

Now let's get back to nullspaces and ranges of our operators. We have $R(p) = c_0$, so $R(V)$ is the set of polynomials of degree 0. Since $p \in N(R)$ if and only if $c_0 = 0$, we see that $N(R)$ is the set of polynomials with constant term zero.

From above, we know that $S(V) = V$. Now, $S(p) = \sum_{i=1}^{n} c_i x^{i-1} = 0$ if and only if $c_i = 0$ for all $i > 0$ if and only $\deg(p) =$, so $N(S)$ is the set of polynomials of degree 0.

From above, $N(T) = 0$. Since $T(p) = x$, we see that $T(V)$ is the set of polynomials with zero constant term as well.

Question 4. *If $n \geq 1$, can you express $(TS)^n$ and $S^n T^n$ in terms of I and R?*

We have $ST = I$ and I^n for each positive integer n. Now,

$$(TS)(p) = T\left(\frac{p - c_0}{x}\right) = \frac{x(p - c_0)}{x} = p - c_0 = I(p) - R(p) = (I - R)(p),$$

so $TS = I - R$. For $n > 1$, we have

$$(TS)^n = T(ST)^{n-1}S = TIS = TS = I - R,$$

so $(TS)^n = I - R$ for each positive n.

Finally, we claim that $S^n T^n = I$ for all positive integers n. We prove this by induction on n. Since $ST = I$, our claim holds for $n = 1$. For $n > 1$, by induction we have

$$S^n T^n = SS^{n-1}T^{n-1}T = SIT = ST = I$$

which implies our result by the principle of mathematical induction.

Example 5. Let $p(x) = 2 + 3x - x^2 + x^3$. What is the image of p under each of the following transformations: $R, S, T, ST, TS, (TS)^2$, T^2S^2, S^2T^2, TRS, RST?

We have $R(p) = c_0 = 2$, $S(p) = (p - c_0)/x = 3 - x + x^2$, and $T(p) = xp = 2x + 3x^2 - x^3 + x^4$. Since $S^n T^n = I$, $(ST)(p) = (S^2T^2)(p) = p$ and $RST(p) = (RI)(p) = R(p)$. Similarly, $(TS)(p) = p - c_0 = 3x - x^2 + x^3$ and $(TS)^2(p) = (TS)(p)$. We calculate

$$(TRS)(p) = (TR)\left(\frac{p - c_0}{x}\right) = T(c_1) = c_1 x = 3x.$$

Finally,

$$T^2S^2 = T(TS)S = T(I - R)S = TIS - TRS = TS - TRS = I - R - TRS$$

and so

$$(T^2S^2)(p) = (I - R - TRS)(p) = p - c_0 - c_1 x = -x^2 + x^3.$$

---

[1]for good reason!