Math 110 - Fall 05 - Lectures notes # 8 - Sep 16 (Friday)

Homework due Thursday, Sep 22:
Sec 2.1: 1 (justify your answers), 3, 4, 5, 6, 7, 13, 14, 15, 16, 17, 18, 19,
20, 21, 22, 25, 35, 38

Start reading chapter 2. Just as Chapter 1 generalized ideas
about vectors familiar from math 54, Chapter 2 will generalize
ideas about matrices:

```
               Ma54          => Ma110
    real numbers             => fields
    vectors of real numbers  => vector spaces
       [ Ex: (x1,x2)         => also P_n(F), Func(R,R), etc ]
    lines and planes through 0=> subspaces
       [ Ex: a*(x1,x2)       => span({s1,s2,...,sn}) ]
    lines are 1D, planes 2D...=> dimension of any vector space

    matrices                 => linear transformations
       [ Ex: [ 1 2 ]         => also differentiation, integration etc ]
             [ 3 4 ]
    multiplying matrix*vector => applying a linear transformation to a vector
       [ Ex: [ 1 2 ]*[x1]=[  x1+2*x2]  => also T(f(x)) = f'(x)   ]
             [ 3 4 ] [x2] [3*x1+4*x2]
    multiplying matrix*matrix => composing linear transformations
    forming inv(Q)*M*Q       => changing the basis of a linear transformation
```

```
Def: Let V and W be vector spaces over F. A function T: V -> W is a
     linear transformation from V to W (or just linear)
     if for all x,y in V and c in F
     (a) T(x+y) = T(x) + T(y)
     (b) T(c*x) = c*T(x)

     (note: (a) => (b) if F = Q, but in general need both)

Lemma:
    T(0_V) = 0_W
    T linear if and only if T(c*x+y) = c*T(x) + T(y)
    T linear implies T(x-y)=T(x) - T(y)
    T linear if and only if T(sum_i c_i*v_i) = sum_i c_i*T(v_i)

Ex: V = R^2, W = R, T((x,y)) = x+2*y
```

```
ASK & WAIT: Why is T linear?

Ex: V = W = R^2, T((x;y)) = (x+y; -3*x+2*y)
ASK & WAIT: Why is T linear?
This is also written as the matrix-vector muliplication
     T([ x ]) = [ 1  1 ] * [ x ] = [    x+y   ]
      ([ y ])   [-3  2 ]   [ y ]   [ -3*x+2*y ]
Indeed, T: F^n -> F^m where T is multiplying
an n-vector by an m x n matrix to get an m-vector
is a linear transformation.
In section 2.2, we will systematically write *all* linear
transformations from F^n to F^m as multiplying an n-vector
by an m x n  matrix, to get an m-vector.
But these are not the only linear transformations.

Ex: V = {differentiable functions from [0,1] to R}
    W = {functions from [0,1] to R}
    T(f(x)) = f'(x) + integral_0^x t^2*f(t) dt
ASK & WAIT: Why is T linear? Is T: V -> V?

Ex: V = W = R^2, T_theta((x,y)) = vector (x,y) rotated clockwise by theta
ASK & WAIT: Why is T linear?


Def: V = W, T(x) = x called identity transformation, written I_V or I

Def: T(x) = 0_W called zero transformation

Def: Let T: V -> W be linear. Then the null space of T, N(T), is
     the set of all vectors v in V such that T(v) = 0_W.
     The range space (or just range) of T, R(T), is the set of
     all vectors w = T(v) for all v in V

Thm 1: N(T) and R(T) are subspaces of V and W, resp.
    Proof: Consider N(T): clearly 0_V in N(T), since T(0_V) = 0_W.
           Since T is linear, v1 and v2 in N(T) implies
             T(a*v1 + b*v2) = T(a*v1)+T(b*v2) = a*T(v1) + b*T(v2)
                            = a*0_V + b*0_V = 0_V,
           so N(T) is closed under + and *, and so a subspace.
           Next consider R(T): clearly 0_W in R(T), since T(0_V) = 0_W.
           If w1 and w2 in R(T), then there are v1 and v2 so that
           w1=T(v1) and w2=T(v2) and so
```

```
            a*w1+b*w2 = a*T(v1)+b*T(v2) = T(a*v1)+T(b*v2) = T(a*v1+b*v2)
           is in R(T) too, so R(T) is closed under + and *, and so a subspace.

Ex:  V = R^2, W = R, T((x;y)) = x+2*y
     N(T) = {(x;y): x+2*y = 0} = {f*(2, -1) for all reals f}
     R(T) = {w: w=x+2*y for some x,y} = R
ASK & WAIT: Why is R(T) = R?

Ex:  V = W = R^2, T((x;y)) = (x+y; -3*x+2*y)
     N(T) = {(x;y): x+y=0 and -3*x + 2*y = 0}
     solving these 2 equations in 2 unknowns => x = y = 0,
     so N(T) = {0_V}
     R(T) = {(w,z): w = x+y and z = -3*x+2*y for some (x;y)}
     solving these 2 equations in 2 unknowns => y = (1/5)*z + (3/5)*w
     and x = (-1/5)*z + (2/5)*w, so any (w;z) in range

Thm 2: If S = {v_1,...,v_n} is a basis for V, then
       R(T) = span(T(v_1),...,T(v_n)}
   Proof: if S is a basis all v in V are of the form
     sum_i a_i*v_i, so all w in R(T) are of the form
        T(sum_i a_i*v_i) = sum_i T(a_i*v_i) = sum_i a_i*T(v_i),
     so {T(v_1),...,T(v_n)} is a spanning set.

ASK & WAIT: Is {T(v_1),...,T(v_n)} necessarily a basis?

Ex:  V = R^2, W = R^3, T((x;y)) = (x+y; -3*x+2*y, -x + 4*y)
     To compute N(T), note x+y=0 and -3*x+2*y=0 implies x=y=0 from before,
     so N(T) = {0_V} as before.
     R(T) = span(T((1;0)), T((0;1))) = span((1;-3;-1),(1;2;4))

Def: Let T: V -> W be a linear transformation with null space N(T) and
     range R(T).
     If N(T) is finite dimensional, we call its dimension the nullity of T,
     written "nullity(T)".
     If R(T) is finite dimensional, we call its dimension the rank of T,
     written "rank(T)".

Thm 3 (Dimension Theorem): Let T: V -> W be a linear transformation.
     If V is finite dimensional, then dim(V) = rank(T) + nullity(T)

Ex: Just as in chapter 1 where we asked what our theorems about vector
    spaces meant for R^3 or R^n, in chapter 2 we can ask what our
```

```
    theorems about linear transformation mean for matrices, especially
    simple matrices like diagonal ones.

    Consider V = R^n, W = R^m and T is multiplication by an m x n matrix.
    Suppose T is diagonal with 0s and 1s on the diagonal,
    i.e. only some T_{ii} can be nonzero, such as
    T = [ 1 0 0 0 ]; we will show that rank(T)    = # nonzero columns of T
        [ 0 1 0 0 ]                    nullity(T) = # zero columns of T
        [ 0 0 0 0 ]
    so rank(T) + nullity(T) = # nonzero columns of T + # zero columns of T
                            = # columns of T
                            = n = dim(V) as claimed by Dimension Theorem

    For simplicity we suppose T_11 = T_22 = ... = T_rr = 1,
    and the rest are 0. So T has r nonzero columns and n-r zero columns.
    Then it is easy to see
      T(x) = T([   x_1   ])  =  [ x_1 ]
              ([   ...   ])     [ ... ]
              ([   x_r   ])     [ x_r ]
              ([ x_{r+1} ])     [  0  ]
              ([   ...   ])     [ ... ]   ... there are m-r zeros
              ([   x_n   ])     [  0  ]
    where there is one zero in the result vector for every zero row in T.
    So we can see that that R(T) is the space spanned by all vectors of the
    form on the right above, which has dimension r = # nonzero columns of T,
    and N(T) is all vectors of the form [ 0; ...; 0; x_{r+1};...; x_n ],
    i.e a space of dimension n-r = # zero columns of T as desired.

Proof of Dimension Theorem:
    Since N(T) is a suspace of the finite dimensional space V, it is
    also finite dimensionsal, and has a basis {v_1,...,v_k}.
    This basis can be extended to a basis of V (Replacement Theorem),
    call it {v_1,...,v_n}. We claim {T(v_{k+1}),...,T(v_n)} is a basis
    of R(T). Assuming this for a moment, since it contains n-k vectors, we get
       dim(V) = n = k + (n-k) = dim(N(T)) + dim(R(T)) as desired.
    To see {T(v_{k+1}),...,T(v_n)} is a basis, we have to show it spans R(T)
    and is independent. But since {v_1,...,v_n} is a basis of V,
     R(T) = span(T(v_1),...,T(v_k),T(v_{k+1}),...,T(v_n))
          = span( 0_W  ,..., 0_W  ,T(v_{k+1}),...,T(v_n))
          = span(T(v_{k+1}),...,T(v_n))
    so it spans R(T). To see it is independent, suppose it is not, and
    seek a contradiction: Write
```

0_W = sum_{i=k+1 to n} a_i*T(v_i)   where not all a_i = 0
    = T(sum_{i=k+1 to n} a_i*v_i) since T is linear
so sum_{i=k+1 to n} a_i*v_i is a vector in N(T), i.e.
sum_{i=k+1 to n} a_i*v_i = sum_{j=1 to k} b_j*v_j
where at least one coefficient (a_i or b_j) is nonzero.
But this contradicts the independence of the basis {v_1,...,v_n}.

Natural questions to ask about any function T, not just linear ones, are
(1) Is T one-to-one, i.e. does T(x)=T(y) imply x=y ?
(2) Is T onto, i.e. for all w in W, is there a v in V such that w=T(V)?

These are important ideas because T is one-to-one and onto if and only if
T is invertible, i.e. an inverse function inv(T): W -> V exists.

These are easy questions to answer for linear T, given the rank and nullity:

Thm: Let T: V -> W where V and W are finite dimensional. Then
(1) T is one-to-one if and only if nullity(T) = 0, i.e. N(T) = {0_V}
(2) T is onto if and only if rank(T) = dim(W)
In particular, since rank + nullity = dim(V), T is invertible if and only
if nullity(T)=0 and dim(V)=dim(W).

Ex: Consider T: F^n -> F^m to be multiplication by an m x n matrix,
where T is diagonal with all T_ii = 1. We ask whether T is one-to-one
and/or onto. There are 3 cases, depending on m and n:
(1) m < n. Then T*(x_1;...;x_n) = (x_1;...;x_m). So R(T) = F^m
and T is onto. But T*(0;...;0;x_{m+1};...;x_n) [m leading zeros]
= 0_W, so nullity(T) = n-m > 0, and T is not one-to-one
We will see that no T can be one-to-one if m < n.
(2) m = n. So T is onto for the same reason as above and
T*(x_1;...;x_n) = (x_1;...;x_n) = 0_W if and only if x=0_V,
so nullity(T)=0 and so T is one-to-one. T is the
identify function.
(3) m > n. Then T*(x_1;...;x_n) = (x_1;...;x_n;0;...;0) [m-n trailing zeros]
So rank(T) = n < m = rank(W) and T is not onto.
But T(v)=0_W only if v=0_V, so nullity(T) = 0 and T is one-to-one.
We will see that no T can be onto if m > n.

Proof: (1) Suppose N(T) = {0_V}. Then
T(x)=T(y) => T(x-y)=0 ... since T is linear
=> x-y=0_V ... since x-y in N(T) = {0_V}
so T is one-to-one.

```
            Conversely, T one-to-one => T(v)=0_W only if v=0_V => N(T) = {0_V}
        (2) T onto => R(T) = W => rank(T) = dim(W).
            Conversely rank(T) = dim(R(T)) = dim(W), so since
            R(T) is a subspace of W, we must have R(T) = W, i.e. T is onto

Corollary: Suppose T: V -> W and dim(V) = dim(W) is finite. Then the following
are equivalent:
      (1) T is one-to-one
      (2) T is onto
      (3) T is invertible
      (4) rank(T) = dim(V)
Proof: To prove them "equivalent", we need to show that if any one of
them is true, then all of them are true. To do this we will prove that
(1) <=> (2),  ((1) and (2)) <=> (3), and (4) <=> (1).

(1) <=> (2) because (1) => nullity(T)=0 =>
    rank(T) = dim(V) - nullity(T) = dim(V) = dim(W) => (2);
    now note that all the implications work in the opposite direction too.

((1) and (2)) <=> (3) by the definition of invertibility

(1) <=> nullity(T)=0 <=> rank(T) = dim(V) - nullity(T) = dim(V)

We need that the dimensions are finite for this to be true, as you
will see on homework.

Ex: T: P_2(R) -> R^3 is defined by T(a2*x^2 + a1*x + a0) = (a2+a1,a1+a0,a0).
Now dim(P_2(R)) = dim(R^3) = 3, so we can apply the Corollary.
Now (a2+a1,a1+a0,a0)=(0,0,0) implies a0=0 => 0 = a1+a0 = a1 => 0 = a2+a1 = a2
so T is one-to-one, and hence is onto and invertible.
ASK & WAIT: What is its inverse?

The next theorem says that a linear transformation T is uniquely determined
if we know what it does to a basis:

Thm: V, W be vectors spaces over F, and let {v_1,..,v_n} be a basis for V.
     Given any subset {w_1,...,w_n} of n vectors from W, there is exactly
     one linear transformation T: V -> W such that T(v_i) = w_i.

Proof: Since {v_1,...,v_n} is a basis, for any v in V there is a unique
       linear combination v = sum_{i=1 to n} a_i*v_i. We can then define
       T(v) = sum_{i=1 to n} a_i*w_i. We need to prove 3 things about T:
```

```
(1) T is linear:
     T( vb + va ) = T(vb) + T(va) because
       T(sum_i b_i*v_i + sum_i a_i*v_i) = T( sum_i (b_i+a_i)*v_i )
                                 = sum_i (b_i+a_i)*w_i
                                 = sum_i b_i*w_i + sum_i a_i*w_i
                                 = T( sum_i b_i*v_i) + T(sum_i a_i*v_i)
     T(c * vb) = c*T(vb) because
       T(c* sum_i b_i*v_i) = T( sum_i (c*b_i)*v_i )
                           = sum_i (c*b_i)*w_i
                           = c* sum_i b_i*w_i
                           = c* T( sum_i b_i*v_i)
(2) T(v_k) = w_k because
      T(v_k) = T(sum_i a_i*v_i) where a_k=1 and the rest are zero
             = sum_i a_i*w_i  where a_k=1 and the rest are zero
             = w_k
(3) T is unique because if U: V -> W also satisfies U(v_i) = w_i then
      for all v = sum_i a_i*v_i we get
      U(v) = U( sum_i a_i*v_i)
           = sum_i U(a_i*v_i)   by linearity of U
           = sum_i a_i*U(v_i)   by linearity of U
           = sum_i a_i*w_i      by def of U(v_i)
           = sum_i a_i*T(v_i)   by def of T(v_i)
           = sum_i T(a_i*v_i)   by linearity of T
           = T( sum_i a_i*v_i)  by linearity of T
           = T( v )
```