Math 350 Fall 2011
Notes about the interplay of matrices with linear maps

One of the themes of the last two lectures was about the correspondence between linear maps (between finite dimensional vector spaces) and matrices. In the following we describe how some terminologies and results about linear maps transfer to the study of matrices.

First, suppose A is a $m \times n$ matrix with real entries[1]. Then one can associate to A a linear map $T\colon \mathbb{R}^n \to \mathbb{R}^m$, via the equation

$$T(x) = Ax \quad \text{for all } x \in \mathbb{R}^n.$$

(In other words[2], one thinks of $x \in \mathbb{R}^n$ as a column vector, and let $T(x)$ be the column vector obtained by multiplying the column vector x by the matrix A.) The *range* of A is now defined to be the range of this linear map T; it is thus a subspace of $\mathbb{R}^m$. The *nullspace* of A is now by definition the nullspace of T; it is thus a subspace of $\mathbb{R}^n$.

On the other hand, given an $m \times n$ real matrix A, there are two other spaces that you may be familiar with. One is called the *column space* of A; this is by definition the vector space of all linear combinations of columns of A. In other words, this is the span of the columns of A, and it is a subspace of $\mathbb{R}^m$. Another is called the *solution space* of the equation $Ax = 0$; in fact, if x is the column vector with n components $x_1, \dots, x_n$, then $Ax = 0$ is a system of m linear equations in n unknowns $x_1, \dots, x_n$, where A is called the coefficient matrix of this system of linear equations. The space of solutions to this equation is a natural subspace of $\mathbb{R}^n$ associated to A. One can see, almost by definition, that the nullspace of A is precisely the space of solutions to the equation $Ax = 0$; on the other hand, the range of A is actually the same as the column space of A. This is because any vector in the range of A is of the form Ax for some column vector x. But Ax is just a linear combination of the columns of A, and any linear combination of the columns of A is equal to Ax for some column vector x. In fact, if $v_1, \dots, v_n$ are the columns of A, and if $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, then

$$Ax = x_1 v_1 + \cdots + x_n v_n.$$

Thus the range of A is equal to the column space of A, as desired.

Question 1. *Let*

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 2 & 1 & 5 \end{pmatrix}.$$

[1] In fact, our discussion below goes through even if the field $\mathbb{R}$ is replaced by a general field F.

[2] Alternatively, if one denotes the standard ordered basis on $\mathbb{R}^n$ and $\mathbb{R}^m$ by β and γ respectively, then this T is the unique linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that the matrix representation $[T]_\beta^\gamma$ of T with respect to β and γ is equal to A.

Find the range of A and the nullspace of A. (The latter amounts to solving the equation $Ax = 0$. Why?)

Now we define the *rank* of a matrix A to be the rank of the linear map T we associated to A. Hence the rank of a matrix A is the dimension of the range of A, or the dimension of the column space of A. On the other hand, we define the *nullity* of A to be the nullity of the linear map T. Hence the nullity of A is the dimension of the space of solutions to the equation $Ax = 0$.

Question 2. *Find the rank and the nullity of the matrix A in Question 1.*

We observe that if A is a $m \times n$ matrix with columns $v_1, \dots, v_n$, then the following are equivalent:

(a) The vectors $v_1, \dots, v_n$ are linearly independent;
(b) The equation $Ax = 0$ has only the trivial solution $x = 0$;
(c) The nullspace of A is $\{0\}$;
(d) The nullity of A is zero;
(e) The linear map associated to A is injective.

In fact, if x is a column vector with components $x_1, \dots, x_n$, then the equation $Ax = 0$ is equivalent to the condition that $x_1 v_1 + \cdots + x_n v_n = 0$. The latter equation has only the trivial solution $x = 0$ if and only if $v_1, \dots, v_n$ are linearly independent. Thus (a) and (b) are equivalent. The equivalence of (b), (c) and (d) is an easy consequence of the definitions, and the equivalence of (d) and (e) was a theorem we have proved.

On the other hand, if A is a $m \times n$ matrix, then the following are equivalent:

(i) The columns of A span $\mathbb{R}^m$;
(ii) The column space of A is the whole $\mathbb{R}^m$;
(iii) The range of A is the whole $\mathbb{R}^m$;
(iv) The linear map associated to A is surjective;
(v) $\text{rank}(A) = m$

The equivalence of (i) through (iv) is almost tautological. We just note that (iii) and (v) are equivalent since the range of A is always a subspace of $\mathbb{R}^m$.

The *rank-nullity theorem* for linear maps now says for any $m \times n$ matrix A, one has

$$\text{rank}(A) + \text{nullity}(A) = n.$$

In particular, if A is an $m \times n$ matrix, then the solution space of the equation $Ax = 0$ has dimension $n - \text{rank}(A)$. If now $m = n$, i.e. if A is a square $n \times n$ matrix, then the rank of A is n, if and only if the nullity of A is zero. Thus together with our above discussion, we see that if A is an $n \times n$ matrix, then the following are equivalent:

(1) The columns of A are linearly independent;

(2) The equation $Ax = 0$ has only the trivial solution $x = 0$;
(3) The nullity of A is zero;
(4) The rank of A is n;
(5) The range of A is the whole $\mathbb{R}^n$;
(6) The columns of A span $\mathbb{R}^n$;
(7) The linear map associated to A is injective;
(8) The linear map associated to A is surjective.

In fact we have already seen a proof of (1) $\Leftrightarrow$ (6) before. This is a second proof of the same fact, via the rank-nullity theorem. It is also very easy to see that any of the above conditions is equivalent to any of the following:

(9) The columns of A is a basis of $\mathbb{R}^n$;
(10) The linear map associated to A is bijective;
(11) The linear map associated to A is an isomorphism.

But (11) is also equivalent to the following:

(12) A is an invertible matrix.

(Recall that a matrix A is said to be invertible if there exists another matrix B such that $AB = BA = I$ where I is the identity matrix.)

In fact, more generally, we have:

Theorem 1. *Suppose $T\colon V \to W$ is a linear map between two finite dimensional vector spaces, and β, γ are ordered bases of V and W respectively. Then T is an isomorphism, if and only if the matrix representation $[T]_\beta^\gamma$ is an invertible matrix.*

Proof. Suppose $A = [T]_\beta^\gamma$ is invertible, with inverse B. Then there is a linear map $S\colon W \to V$ such that $[S]_\gamma^\beta = B$. One can check that $S \circ T$ is the identity map on V; this is because $[S \circ T]_\beta^\beta = [S]_\gamma^\beta [T]_\beta^\gamma = BA = I$ where I is the identity matrix. Similarly, one can check that $T \circ S$ is the identity map on W. Thus $T\colon V \to W$ is invertible as a map from V to W. It follows that T is bijective, and hence an isomorphism.

Conversely, suppose T is an isomorphism. Then $T\colon V \to W$ is bijective, so there is a map $S\colon W \to V$ such that $S \circ T$ is the identity map on V, and $T \circ S$ is the identity map on W. One can check that this map S is linear. Thus it has a matrix representation, say $B = [S]_\gamma^\beta$. Now if $A = [T]_\beta^\gamma$, we claim $AB = I$, the identity matrix. This is because $AB = [T]_\beta^\gamma [S]_\gamma^\beta = [T \circ S]_\gamma^\gamma = [Id_W]_\gamma^\gamma = I$, where Id_W is the identity map on W. Similarly, one has $BA = I$. Thus A is invertible, and in fact $A^{-1} = B$. $\square$

At this point, we take the chance to point out that the system of linear equations we considered above, namely $Ax = 0$, is usually called a *homogeneous* system of linear equations. The term homogeneous refers to that the right hand side of this system is zero. On the other hand, one can also consider a more general system of linear equations, namely $Ax = b$; if $b \neq 0$ then one says this is an *inhomogeneous* system of linear equations. Sometimes we also call $Ax = 0$ the homogeneous system of linear equations *associated* to the equation $Ax = b$.

To relate to what we have discussed above, given an $m \times n$ matrix A and a column vector $b \in \mathbb{R}^m$, we say that the system $Ax = b$ is *solvable*, if and only if there is some column vector $x \in \mathbb{R}^n$ that solves this equation. This is true if and only if b is in the column space of A. If the system $Ax = b$ is solvable, then the set of all solutions is equal to the set

$$\{x_0 + v \colon v \in \text{nullspace}(A)\}$$

where x_0 is any solution of the equation $Ax = b$. (Why? Prove this!) As a result, the set of all solutions of the equation $Ax = b$ is a translate of a vector space (namely the nullspace of A) by a vector (namely x_0); we sometimes call such a set an *affine space*.

Question 3. *Let*

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 2 & 1 & 5 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

(This is the same matrix A as in Question 1.)

(a) Find all solutions to the system of equations $Ax = b_1$.
(b) Find all solutions to the system of equations $Ax = b_2$.
(c) Is the system of equations $Ax = b_1$ solvable?
(d) Is the system of equations $Ax = b_2$ solvable?
(e) (i) Find all column vectors b such that the equation $Ax = b$ is solvable.
(ii) Compare your answer to the range of A that you found in Question 1. How does this relate to the column space of A?
(f) (i) Show that $x_0 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ is a solution to the equation $Ax = b_1$.
(ii) Verify that all solutions to the equation $Ax = b_1$ (which you found in part (a)) is of the form $x_0 + v$, where v is an element of the nullspace of A. (The nullspace of A was found in Question 1.)

Next, if A happens to be a square matrix, say A is $n \times n$, then (1) through (12) above are all equivalent to any of the following:

(13) The equation $Ax = b$ is solvable for any column vector $b \in \mathbb{R}^n$;
(14) The equation $Ax = b$ has exactly one solution for any column vector $b \in \mathbb{R}^n$.

In fact, if (12) holds, i.e. if A is invertible, then for any $b \in \mathbb{R}^n$, the equation $Ax = b$ is equivalent to that $x = A^{-1}b$, so the equation $Ax = b$ has exactly one solution for any $b \in \mathbb{R}^n$, and (14) follows. On the other hand, if (14) holds, then so does (13), but if (13) holds, then any $b \in \mathbb{R}^n$ is in the range of A, so (5) holds. This proves that (1) through (14) are all equivalent.

The solution of systems of linear equations is usually carried out through Gaussian elimination (or row operations). A row operation is one of the following:

(I) Interchanging two rows of a matrix;
(II) Multiplying a row of a matrix by a non-zero scalar;
(III) Subtracting from a row of a matrix a multiple of another row.

By applying these row operations successively to the augmented matrix of a system of linear equations, one can always reduce it to the row-echelon form, i.e. an upper triangular matrix so that the entries on the diagonal is a sequence of 1's followed by a sequence of 0's. One can then solve the system of linear equations by backward-substitution.

On the other hand, one of the things we'd like to point out here is that row operations can be carried out by matrix multiplications. For example, if A is a 3×3 matrix, and if you want to interchange the first two rows of the matrix, then you just need to multiply, on the left-hand side of A, the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In fact if E_0 is the above matrix, then $E_0 A$ is precisely the matrix that you would obtain from A by interchanging the first two rows of A. More generally, if A is a $m \times n$ matrix, then for any row operation you want to apply to A, there is a $m \times m$ square matrix E such that EA is the matrix that you would obtain by applying the desired row operation to A. The set of matrices E that arise this way is called the set of all *elementary matrices*; this is precisely the set of all matrices that one can obtain by applying the row operations to the $m \times m$ identity matrix. Any elementary matrix is invertible; this is because you can always undo any row operation you have done on A. There is also a notion called column operations; if A is an $m \times n$ matrix, then the column operations on A are precisely those operations on A that can be achieved by multiplying on the right-hand side of A an elementary matrix of size $n \times n$.

You should have learned from a previous course that any $m \times n$ matrix A can be brought into the a diagonal matrix D, where the diagonal entries are a sequence of 1's followed by a sequence of 0's, by applying row and column operations. For the lack of terminology, let's temporarily call D the diagonal matrix associated to A. Thus from the above discussion, $D = PAQ$ for some invertible matrices P and Q.

Now it is an easy theorem that if A is an $m \times n$ matrix, P is a $m \times m$ invertible matrix, and Q is a $n \times n$ invertible matrix, then the range of A is the same as the range of AQ, and the rank of PA is the same as the rank of A. Combining these two facts, one realizes that under the same assumptions on A, P and Q, we have $\text{rank}(A) = \text{rank}(PAQ)$. So the rank of a matrix A is equal to the rank of the diagonal matrix associated to A.

On the other hand, if A is any $m \times n$ matrix, then the diagonal matrix associated to A^t is equal to the transpose of the diagonal matrix associated of A. Since the rank of the diagonal matrix associated to A is the same as the rank of diagonal matrix associated to A^t (in fact the diagonal matrices are transposes of each other), one has

$$\text{rank}(A) = \text{rank}(A^t)$$

for any $m \times n$ matrix A.

It follows that if A is a square matrix, say A is $n \times n$, then the nullity of A is the same as the nullity of A^t (since both are equal to $n - r$ where $r = \text{rank}(A) = \text{rank}(A^t)$). In particular, if A is a square matrix, and T, T^t are the linear maps associated to A and A^t respectively, then

$$T \text{ is injective} \Leftrightarrow T^t \text{ is injective} \Leftrightarrow T^t \text{ is surjective} \Leftrightarrow T \text{ is surjective.}$$

(In fact, the first two statements are equivalent to that the nullity of $A = 0$, and the last two statements are equivalent to the fact that $\text{rank}(A) = n$, whereas for $n \times n$ matrices, $\text{nullity}(A) = 0$ if and only if $\text{rank}(A) = n$.)

Incidentally, one should remember that the analog of this last fact is not true, if T is a linear map from an infinite dimensional vector space into itself. This last statement is very much a finite dimensional phenomenon.

Finally, you should remember from a previous course that one can define the determinant of any $n \times n$ matrix A. If A is an $n \times n$ matrix, and B is the transpose of its cofactor matrix, then $AB = \det(A)I$ where I is the identity matrix. We claim that if $\det(A) \neq 0$, then A is invertible. In fact if $\det(A) \neq 0$, then first we claim that the equation $Bx = 0$ has only the trivial solution $x = 0$; this is because if $\det(A) \neq 0$ and $Bx = 0$, then $x = Ix = \frac{1}{\det(A)}A(Bx) = \frac{1}{\det(A)}A(0) = 0$. It then follows from our previous discussion that B is invertible, so $BA = BABB^{-1} = B(\det(A)I)B^{-1} = \det(A)I$ as well. It follows that $C := \frac{1}{\det(A)}B$ satisfies $CA = AC = I$. So A is invertible.

On the other hand, if A were invertible, say $C = A^{-1}$, then $AC = I$, where I is the identity matrix, and thus $\det(A)\det(C) = 1$, from which one sees that $\det(A) \neq 0$.

So now if A is an $n \times n$ matrix, we have one last statement, that is equivalent to all of (1) through (14) above:

(15) $\det(A) \neq 0$