Math 413/513 Chapter 3 (from Friedberg, Insel, & Spence)

David Glickenstein

October 16, 2015

1 Matrices

1.1 Main results of this section

The main theorem of this section is the following:

Theorem 1 *Let A be an $m \times n$ matrix of rank r. Then $r \leq m$, $r \leq n$, and A can be transformed by row and column operations into a matrix*

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

where O_1, O_2, O_3 are zero matrices. Thus $D_{ij} = 1$ if $i = j \leq r$ and 0 otherwise.

Corollary 2 *There exist invertible matrices $B \in F^{m \times m}$ and $C \in F^{n \times n}$ such that $D = BAC$.*

Corollary 3 *Let A be an $m \times n$ matrix. Then*

1. $\operatorname{rank} A = \operatorname{rank} A^T$.
2. *The rank of A is equal to the maximum number of linear independent rows, which is equal to the dimension of the row space.*
3. *The rows and columns of A generate subspaces of the same dimension, each with dimension equal to the rank of A.*

Corollary 4 *Every invertible matrix is the product of elementary matrices.*

1.2 Explanation and proof of the corollaries

In order to make sense of these we need to know (1) what rank of a matrix is, (2) what row and column operations are, (3) what elementary matrices are, and (4) what the row and column spaces are.

Definition 5 *If $A \in F^{m\times n}$, then the* $\operatorname{rank} A = \operatorname{rank} L_A$*, where L_A is the linear transformation $F^n \to F^m$ given by $x \to Ax$. The column space is the span of the columns of A (so it is the span of n vectors in F^m) and the row space is the span of the rows of A (so it is the span of m vectors in F^m).*

Proposition 6 *The column space of a matrix A is equal to $R(L_A)$. In particular, the rank of A is equal to the dimension of the column space.*

Proof. The ith column is simply $Ae_i \in R(L_A)$, thus the column space is in $R(L_A)$. If $y \in R(L_A)$, then there exists $x \in F^n$ such that $y = Ax$, and so if the columns of A are labeled $A_1, \ldots, A_n$, we have that

$$y = x_1 A_1 + \cdots + x_n A_n.$$

■

We already know that row operations are important. It turns out that row operations correspond to multiplication by certain matrices.

Definition 7 *If $A \in F^{m\times n}$, the elementary row [column] operations are the following:*

1. *interchanging any two rows [columns] of A;*
2. *multiplying any row [column] of A by a nonzero scalar;*
3. *adding any scalar multiple of a row [column] of A to another row [column].*

Definition 8 *An $n \times n$ elementary matrix is a matrix obtained by performing one elementary row operation on I_n.*

So examples of elementary matrices are

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Theorem 9 *Let $A \in F^{m\times n}$ and suppose B is obtained from A by performing an elementary row [column] operation. Then there exists an $m \times m$ [$n \times n$] elementary matrix E such that $B = EA$ [$B = AE$], where E is the elementary matrix obtained from I in the same way. Conversely, multiplication by elementary matrices correspond to performing row operations.*

Proof. The proof relies on verifying this on each type of row/column operation and each type of elementary matrix. It helps that we know that

$$Be_i = B_i$$

where B_i is the ith column, and we can see similarly that $e_i^T B$ is the ith row. The details are left as an exercise. ■

It now follows that elementary matrices are invertible (this can also be verified directly as well) since each row operation is reversible. Thus doing a sequence of row (column) operations is the same as multiplying on the left (right) by a sequence of elementary matrices. Thus, the only thing left to show that Theorem 1 implies Corollary 2 is that the product of invertible matrices is invertible.

Theorem 10 *Let $A \in F^{m\times n}$, $P \in F^{m\times m}$, and $Q \in F^{n\times n}$. If P and Q are invertible, then*

1. $\operatorname{rank}(AQ) = \operatorname{rank} A$
2. $\operatorname{rank}(PA) = \operatorname{rank} A$
3. $\operatorname{rank}(PAQ) = \operatorname{rank} A$

Proof. We can use what we know about linear transformations to see that

$$R(L_{AQ}) = R(L_A L_Q) = L_A L_Q(F^n)$$

But since Q is invertible, it is onto, so

$$L_A L_Q(F^n) = L_A(F^n) = R(A).$$

Thus the ranks of AQ and A are equal. Similarly, since P invertible, L_P is an isomorphism, so

$$\dim L_P(L_A(F^n)) = \dim L_A(F^n),$$

or $\operatorname{rank}(PA) = \operatorname{rank} A$. The last one follows. ∎

Corollary 11 *The product of elementary matrices is invertible, and elementary row and column operations and their compositions are rank preserving.*

This also allows us to prove Corollary 4.

Proof of Corollary 4. We have just shown that the product of elementary matrices is invertible. Now suppose a matrix is invertible. By Corollary 2, we can convert the matrix into the form of D. Since this is rank preserving and the original matrix is invertible, D must be the identity matrix. But then we have that $I = BAC$ where B and C are products of elementary matrices. It follows that $A = B^{-1}C^{-1}$, which is also a product of elementary matrices (if $B = E_1E_2\cdots E_k$ and $C = E_{k+1}E_{k+1}\cdots E_\ell$ then $B^{-1}C^{-1} = E_k^{-1}\cdots E_1^{-1}E_\ell^{-1}\cdots E_{k+1}^{-1}$, which is a product of elementary matrices since the inverse of an elementary matrix is an elementary matrix). ∎

Finally, we can prove Corollary 3.

Proof of Corollary 3. We first note that if E is an elementary matrix, then E^T is also an elementary matrix (check this!) and also that D^T has the same form as D (though the dimensions of the matrix may differ). So if $D = BAC$ then we have that $D^T = C^T A^T B^T$, and the rank of D^T is clearly equal to $r = \operatorname{rank} A$. Since multiplication by C^T and B^T is rank preserving (it is the product of elementary matrices) we have that $\operatorname{rank} A^T = \operatorname{rank} D^T = \operatorname{rank} D = \operatorname{rank} A$. The others follow from the fact that $\operatorname{rank} A$ is the dimension of the column space of A and $\operatorname{rank} A^T$ is the dimension of the row space of A (which is the column space of A^T). ∎

1.3 Proof of Theorem 1

We phrase this as a sequence of exercises:

We will induct on m, the number of rows. Before looking at base cases, we actually think about the inductive step:

1) If $m > 1$ and $n > 1$ and A is not the zero matrix, show that by a finite number of row and column operations, A can be transformed into a matrix of the form:

$$B = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{pmatrix}$$

where B' is a matrix of dimension $(m-1) \times (n-1)$. Note that $\operatorname{rank} B = 1 + \operatorname{rank} B' \leq m$ and $\leq n$ by the inductive hypothesis.

We now can use the inductive hypothesis on B'. We need only prove the cases of $n = 1$, $m = 1$, and for the zero matrix. Note that the zero matrix is already in the proper form, and has rank equal to zero (less than n and m).

2) Show that if $n = 1$ and A is not the zero matrix, we can transform A to the appropriate form by a finite sequence of column operations, and $\operatorname{rank} A = 1$.

3) Show that if $m = 1$ and A is not the zero matrix, we can transform A to the appropriate form by a finite sequence of row operations and $\operatorname{rank} A = 1$.

This completes the proof. I suggest you go through the process on an actual example. The one from the book is:

$$A = \begin{pmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix}.$$

You should find that this matrix has rank 3 and that you can transform it by row and column operations to the matrix

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

2 Inverse matrix

We already know that an $n \times n$ matrix is invertible if and only if it has rank n, and if and only if it is a product of elementary matrices. We can compute the inverse using augmented matrices.

Definition 12 *Let A and B be $m \times n$ and $m \times p$ matrices, respectively. The augmented matrix $(A|B)$ is the $m \times (n+p)$ matrix $(A\ B)$, i.e., the matrix whose first n columns are the columns of A and the last p columns are the columns of B.*

Proposition 13 *Suppose A and B both have n rows and suppose M is an $m \times n$ matrix. Then*

$$M\,(A|B) = (MA|MB)\,.$$

Proof. Exercise. ■

Now we can use this idea to find the inverse, since we see that

$$A^{-1}\,(A|I) = \left(A^{-1}A|A^{-1}\right) = \left(I|A^{-1}\right).$$

Furthermore, we know that if A is invertible, we can use row and column operations to transform A into I, but in particular, since A is invertible, we do not need column operations. Thus there is an invertible matrix B, the product of elementary matrices, such that $BA = I$. If we apply this matrix to $(A|I)$, we get

$$B\,(A|I) = (BA|B) = (I|B)\,.$$

Since $BA = I$, we have that $B = A^{-1}$ and so we can construct the inverse by converting $(A|I)$ to $\left(I|A^{-1}\right)$ via row operations.

3 Problems

- FIS Section 3.1 exercises 2, 3, 5, 7-11
- FIS Section 3.2 exercises 2-7, 11, 14, 15, 18, 21. Read the Theorem 3.7 and its proof.

4 Systems of linear equations

Recall a system of m linear equations in n unknowns can be written as

$$Ax = b$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

and

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

are given and

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

is the unknown. The solution set S of the equation $Ax = b$ is the set of all $x \in F^n$ that satisfy this equation and each element of S is said to be a solution of the equation.

Definition 14 *The system $Ax = b$ is said to be* consistent *if its solution set S is nonempty, otherwise it is said to be* inconsistent*.*

Definition 15 *If $b = \vec{0}$, then we say the system is* homogeneous*, otherwise it is* nonhomogeneous*.*

Note that each equation $Ax = b$ has an associated homogeneous equation $Ax = \vec{0}$.

Notice that every homogeneous system has a solution, namely $x = \vec{0}$, so every homogeneous linear system is consistent.

Theorem 16 *Let $Ax = \vec{0}$ be a homogeneous system of m equations with n unknowns over a field F, and let K denote the solution set. Then $K = N(L_A)$ and hence K is a subspace of F^n of dimension $n - \operatorname{rank} A$.*

Corollary 17 *If $m < n$ then $Ax = 0$ has a nonzero solution.*

Proof. We know that $\operatorname{rank} A = \operatorname{rank} L_A \leq m$, and so the dimension of the solution set K is $n - \operatorname{rank} A \geq n - m > 0$. ■

Theorem 18 *Let K be the solution set to the linear equation $Ax = b$ and let K_H be the solution set to the associated homogeneous equation $Ax = \vec{0}$. Then for any solution $s \in K$,*

$$K = \{s\} + K_H = \{s + k : k \in K_H\}.$$

Note that this says that K is not a vector space, but an affine space (we did not define this, but it is a translation of a vector space). K_H is a vector space.
Proof. Suppose $s \in K$. If $k \in K_H$, then $A(s + k) = As + Ak = b + \vec{0} = b$, and so $s + k \in K$, and we get $\{s\} + K_H \subseteq K$. Conversely, if $x \in K$, then $A(s - x) = As - Ax = b - b = \vec{0}$ and so $s - x \in K_H$, and $x = s + (x - s) \in \{s\} + K_H$. ■

Theorem 19 *Let $Ax = b$ be a system of linear equations in n equations and n unknowns. If A is invertible, then the system has exactly one solution, $A^{-1}b$. Conversely, if the system has exactly one solution, then A is invertible.*

Proof. Suppose A is invertible. Then clearly if $x = A^{-1}b$ then $Ax = A\left(A^{-1}b\right) = \left(AA^{-1}\right)b = Ib = b$ so x is a solution. Suppose s is any solution, then

$$\begin{aligned} As &= b \\ A^{-1}As &= A^{-1}b \end{aligned}$$

and we can conclude that $s = A^{-1}b$.

Conversely, suppose $Ax = b$ has exactly one solution, s. If $k \in K_H$, then $s + k$ is also a solution, so if s is the only solution, we must have that $K_H = \left\{\vec{0}\right\}$. That implies that $N(L_A) = \left\{\vec{0}\right\}$ and looking at the dimsensions we can conclude that A is invertible. ■

Theorem 20 *Let $Ax = b$ be a system of linear equations. Then the system is consistent if and only if* $\operatorname{rank} A = \operatorname{rank}(A|b)$.

Proof. Notice $Ax = b$ is consistent if and only if $b \in R(L_A)$ if and only if b is in the column space (span of the columns of A). This is true if and only if the ranks of $(A|b)$ is the same as the rank of A. ■

5 Problems

FIS section 3.3, exercises 2-6, 8-10

6 Solving systems of linear equations

Definition 21 *Two systems of linear equations are* equivalent *if they have the same solution set.*

Theorem 22 *Let $Ax = b$ be a system of m linear equations in n unknowns, and let $C \in F^{m\times m}$ be invertible. Then the system*

$$(CA)x = Cb$$

is equivalent to $Ax = b$.

Proof. Suppose x is a solutoin to $Ax = b$. Then $C(Ax) = Cb$ and hence $(CA)x = Cb$. Conversely, suppose $(CA)x = Cb$. Then $Ax = C^{-1}Cb = b$. ■

Corollary 23 *If $(A'|b')$ is obtained from $(A|b)$ by a finite number of elementary row operations, then $A'x = b'$ is equivalent to the original system.*

Proof. The row operations can be represented by an invertible matrix C, and so $(A'|b') = C(A|b) = (CA|Cb)$. The result now follows from the Theorem. ■

We now describe an algorithm for easily solving linear systems called Gaussian Elimination. We will find a sequence of row operations to produce zeroes in the matrix:

1) In the leftmost nonzero column, create a 1 in the first row.

2) Use row operations of the third type to obtain zeroes in the other positions in the first column.

3) Create a 1 in the next row in the leftmost possible column without using the previous rows.

4) Obtain zeroes below that 1.

5) Repeat steps 3 and 4 until there are no nonzero rows.

Steps 1-5 are called the forward pass and result in a matrix that could be pretty easily solved using back substitution. To make it easier, however, we will do a backward pass to get more zeroes.

6) Working upward, beginning with the last nonzero row, add multiples of each row to the rows above to create zeroes.

7) Repeat step 6 until is done on the second row.

The result is a matrix in reduced row echelon form.

Definition 24 *A matrix is said to be in* reduced row echelon form *if all of the following are satisfied:*

1. *The rows containing all zeroes are at the bottom (if there are any).*
2. *The first nonzero entry in each row is the only nonzero entry in its column. (These are called* pivots.*)*
3. *The first nonzero entry in each row is 1 and it soccurs in a column to the right of the first nonzero entry in the preceding row.*

Theorem 25 *Let $Ax = b$ be a system of r nonzero equations in n unknowns. Suppose* $\operatorname{rank} A = \operatorname{rank}(A|b)$ *and $(A|b)$ is in reduced row echelon form. Then*

1. $\operatorname{rank} A = r$
2. *Each of the variables that does not appear as a pivot is a free variable, and the others are determined by solving the equation. In this way, we can write the solution to $Ax = b$ as $s_0 + \sum_{j=1}^{n-r} t_j u_j$.*

As an example, consider

$$\begin{pmatrix} 2 & 3 & 1 & 4 & -9| & 17 \\ 1 & 1 & 1 & 1 & -3| & 6 \\ 1 & 1 & 1 & 2 & -5| & 8 \\ 2 & 2 & 2 & 3 & -8| & 14 \end{pmatrix}$$

which has reduced row echelon form

$$\begin{pmatrix} 1 & 0 & 2 & 0 & -2| & 3 \\ 0 & 1 & -1 & 0 & 1| & 1 \\ 0 & 0 & 0 & 1 & -2| & 2 \\ 0 & 0 & 0 & 0 & 0| & 0 \end{pmatrix}$$

Theorem 26 *Let A be an $m \times n$ matrix of rank r, where $r > 0$ and let B be the reduced row echelon form of A. Then:*

1. *The number of nonzero rows of B is r.*
2. *B contains columns that look like $e_1, \ldots, e_r$ (and possibly other columns) where e_j is equal to 1 in the jth slot and zero otherwise. We call these columns $b_{j_1}, \ldots, b_{j_r}$.*
3. *The columns of A that are in the same slots as the columns of B corresponding to $e_1, \ldots, e_r$ are linearly independent, i.e., the columns $a_{j_1}, \ldots, a_{j_r}$ are linearly independent.*

4. *if a column k of B is $d_1e_1 + \cdots + d_re_r = d_1b_{j_1} + \cdots + d_rb_{j_r}$, then the column k of A is $d_1a_{j_1} + \cdots + d_ra_{j_r}$ (notice the same coefficients).*

Proof. The first statement is clear since if there are more, then the rank of B would be less than r, but since there exists an invertible matrix C such that $B = CA$, we must have that $\operatorname{rank} B = \operatorname{rank} A = r$. The second follows from the definition of reduced row echelon form and the fact that the rank is r. The last two follow from the fact that the columns satisfy $b_j = Ca_j$, so

$$\sum_{i=1}^{r} d_i b_{j_i} = \sum_{i=1}^{r} d_i C a_{j_i} = C \sum_{i=1}^{r} d_i a_{j_i}.$$

■

Corollary 27 *The reduced row echelon form is unique.*

Proof. Exercise ■

7 Problems

FIS Section 3.4, exercises 2, 3, 5, 6, 9, 10, 12, 14, 15