æ

REVIEW NOTES ON MATH 250

This note is prepared by Prof. Gene Speers, and is a brief review of a small portion of Math 250, covering:

- A few basic facts about the nature of the set of solutions of a system of linear equations;
- Computation of the reduced row-echelon form of a matrix;
- The process of solving a system of linear equations using Gaussian elimination, which here means through the reduced row echelon form (RREF) of the augmented matrix;
- Interpreting the RREF of the augmented matrix: Do the equations have a solution? If so, is the solution unique? If it is not unique, how many free parameters are there?
- Interpreting the RREF of the coefficient matrix: Do the equations have a solution for every possible right hand side? If so, how many free parameters are there in a solution?
- The rank and nullity of a matrix.

You are responsible for knowing this material at the beginning of the semester and throughout the course. There are some exercises in Section 6 of these notes through which you can test your knowledge. If the notes are not a sufficient review, you should consult the Math 250 text [1] or chapter 3 of the Math 350 text [2].

There will be an in-class quiz on this material on Monday, January 26. **You must pass this quiz within the first two weeks of the course**; if you don't pass the version given on the 26$^{\text{th}}$ you may retake it during my office hours.

In Section 5 we cover an additional topic:

- Linear independence.

You should know this material, but it will not be on the review quiz.

SOLVING SYSTEMS OF LINEAR EQUATIONS

1. General properties of solutions

Suppose we are given a system of m linear equations in n unknowns $x_1, \ldots, x_n$:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{1}$$

To write this in a more compact form we introduce the *coefficient matrix* A, the vector $\mathbf{b}$ giving the terms on the right hand side of the equations, and the vector $\mathbf{x}$ of unknowns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix},$$

so that (1) becomes $A\mathbf{x} = \mathbf{b}$. **(In these notes, the term *vector* always refers to a column vector of real numbers.)** In the next section we turn to the problem of determining whether or not the system has any solutions and, if so, of finding them all. Before that, however, we make some general comments on the nature of solutions.

The homogeneous problem. Suppose first that the system (1) is *homogeneous*, that is, that the right hand side is zero, or equivalently that $b_1 = b_2 = \cdots = b_m = 0$ or $\mathbf{b} = \mathbf{0}$. Suppose further that we have found, by some method, two solutions $\mathbf{x}_1$ and $\mathbf{x}_2$ of the equations. Then for any constants c and d, $\mathbf{x} = c\mathbf{x}_1 + d\mathbf{x}_2$ is also a solution, since

$$A\mathbf{x} = A(c\mathbf{x}_1 + d\mathbf{x}_2) = cA\mathbf{x}_1 + dA\mathbf{x}_2 = c \cdot \mathbf{0} + d \cdot \mathbf{0} = \mathbf{0}.$$

The key step is at the second equality: we are using the fact that matrix multiplication is linear, which means exactly that $A(c\mathbf{x}_1 + d\mathbf{x}_2) = cA\mathbf{x}_1 + dA\mathbf{x}_2$. The argument extends to any number of solutions, and we have the

Theorem 1: The principle of superposition. *If $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_k$ are all solutions of $A\mathbf{x} = \mathbf{0}$, and c_1, c_2, ..., c_k are constants, then*

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k \tag{2}$$

is also a solution of $A\mathbf{x} = \mathbf{0}$.

The name of this principle comes from the fact that (2) is called a *linear combination* or *linear superposition* of the solutions $\mathbf{x}_1, \ldots, \mathbf{x}_k$. We will see later (see Theorem 3 (iii)) that there is a special value of k such that (i) we can find a set of solutions $\mathbf{x}_1, \ldots, \mathbf{x}_k$ with the property that *every* solution of $A\mathbf{x} = \mathbf{0}$ can be built as a linear combination of these solutions, and (ii) k different solutions are really needed for this to be true.

To verify the principle, that is, to see that (2) is a solution of $A\mathbf{x} = \mathbf{0}$, we just plug the putative solution into the equation and again use linearity of matrix multiplication:

$$A\mathbf{x} = A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) = c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2 + \cdots + c_kA\mathbf{x}_k = c_1 \cdot \mathbf{0} + \cdots + c_k \cdot \mathbf{0} = \mathbf{0}.$$

Notice also that **the homogeneous system always has at least one solution, the zero solution $\mathbf{x} = \mathbf{0}$**, since $A\mathbf{0} = \mathbf{0}$. This is the *trivial* solution. The system may or may not also have nozero solutions, which are called *nontrivial*.

The inhomogeneous problem. Consider now the case in which the system (1) is *inhomogeneous*, that is, $\mathbf{b}$ is arbitrary. Suppose again that we are given two solutions, which we will now call $\mathbf{x}$ and $\mathbf{X}$. Then $\mathbf{x}_h = \mathbf{x} - \mathbf{X}$ is a solution of the *homogeneous* system, since

$$A\mathbf{x}_h = A(\mathbf{x} - \mathbf{X}) = A\mathbf{x} - A\mathbf{X} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

What this means is that if we know *one* solution of our equations, X, then every other solution has the form $\mathbf{x} = \mathbf{X} + \mathbf{x}_h$ with $A\mathbf{x}_h = \mathbf{0}$; with (2), this means that

$$\mathbf{x} = \mathbf{X} + c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k. \tag{3}$$

with $\mathbf{x}_1, \ldots, \mathbf{x}_k$ solutions of $A\mathbf{x} = \mathbf{0}$. We have the

Theorem 2: Inhomogeneous linear equations. *Every solution* $\mathbf{x}$ *of the system of inhomogeneous equations (1) is of the form* $\mathbf{x} = \mathbf{X} + \mathbf{x}_h$*, where* $\mathbf{X}$ *is some particular solution of the system, and* $\mathbf{x}_h$ *is a solution of the corresponding homogeneous system, that is,* $A\mathbf{x}_h = \mathbf{0}$ *and* $\mathbf{x}_h$ *has the form (2).*

One can check directly that $\mathbf{X} + \mathbf{x}_h$ is a solution just by plugging it into the equation $A\mathbf{x} = \mathbf{b}$:

$$A\mathbf{x} = A(\mathbf{X} + \mathbf{x}_h) = A\mathbf{X} + A\mathbf{x}_h = \mathbf{b} + \mathbf{0} = \mathbf{0}.$$

2. Row reduction and reduced row-echelon form

The key technique that we will use for solving linear equations, and also for investigating general properties of the solutions, is the reduction of a matrix to *reduced row-echelon form* by the use of *elementary row operations*, a procedure often called *row reduction* or *Gaussian elimination*. Symbolically, if A is a matrix, we have

$$A \xrightarrow[\text{operations}]{\substack{\text{elementary}\\ \text{row}}} R$$

where R is in reduced row-echelon form. What does this all mean?

Reduced row-echelon form: The matrix R is in reduced row-echelon form (RREF) if it satisfies four conditions:

(i) All nonzero rows (that is, rows with at least one nonzero entry) are above any zero rows (rows with all zeros).

(ii) The first nonzero entry in any nonzero row is a 1. This entry is called a *pivot*.

(iii) Each pivot lies to the right of the pivot in the row above it.

(iv) All matrix entries above a pivot are zero.

Here is a matrix in reduced row-echelon form:

$$R = \begin{pmatrix} 0 & \mathbf{1} & 3 & 0 & -1 & 0 & 0 & 2 \\ 0 & 0 & 0 & \mathbf{1} & -2 & 0 & 0 & 65 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 28 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The pivots are the entries, all with value 1, shown in boldface.

Remark: Reduced row-echelon form is a special case of *row-echelon form*. Row-echelon form is important for some computational purposes, but in this course we will simplify our life by working only with the reduced row-echelon form.

One fact with which you may not be familiar is that the RREF of a matrix A is unique—whatever sequence of row operations is used to go from A to R, with R in RREF, the resulting R will be the same. The matrix R is called *the reduced row-echelon form of A*.

Example 1: Row reduction

Here we carry out the reduction of a 3×4 matrix to reduced row-echelon form. We indicate the row operations used by a simple notation: $\mathbf{r}_i$ denotes the i^{th} row of the matrix, and the row operations are denoted by $\mathbf{r}_i \leftrightarrow \mathbf{r}_j$ (interchange rows i and j), $\mathbf{r}_i \to c\,\mathbf{r}_i$ (multiply row i by the scalar c), and $\mathbf{r}_i \to \mathbf{r}_i + c\,\mathbf{r}_j$ (add c times row j to row i). Notice that in the first step we *must* switch the first row with another: because the first column is not identically zero, the first pivot must be in the upper left corner, and we need a nonzero entry there to get started.

$$\begin{pmatrix} 0 & -3 & -1 & 1 \\ 1 & 2 & 3 & 0 \\ 2 & 2 & 5 & -3 \end{pmatrix} \xrightarrow{\mathbf{r}_1 \leftrightarrow \mathbf{r}_2} \begin{pmatrix} \mathbf{1} & 2 & 3 & 0 \\ 0 & -3 & -1 & 1 \\ 2 & 2 & 5 & -3 \end{pmatrix}$$

$$\xrightarrow{\mathbf{r}_3 \to \mathbf{r}_3 - 2\,\mathbf{r}_1} \begin{pmatrix} \mathbf{1} & 2 & 3 & 0 \\ 0 & -3 & -1 & 1 \\ 0 & -2 & -1 & -3 \end{pmatrix}$$

$$\xrightarrow{\mathbf{r}_2 \to -(1/3)\,\mathbf{r}_2} \begin{pmatrix} \mathbf{1} & 2 & 3 & 0 \\ 0 & \mathbf{1} & 1/3 & -1/3 \\ 0 & -2 & -1 & -3 \end{pmatrix}$$

$$\xrightarrow[\mathbf{r}_3 \to \mathbf{r}_3 + 2\,\mathbf{r}_2]{\mathbf{r}_1 \to \mathbf{r}_1 - 2\,\mathbf{r}_2} \begin{pmatrix} \mathbf{1} & 0 & 7/3 & 2/3 \\ 0 & \mathbf{1} & 1/3 & -1/3 \\ 0 & 0 & -1/3 & -11/3 \end{pmatrix}$$

$$\xrightarrow{\mathbf{r}_3 \to -3\,\mathbf{r}_3} \begin{pmatrix} \mathbf{1} & 0 & 7/3 & 2/3 \\ 0 & \mathbf{1} & 1/3 & -1/3 \\ 0 & 0 & \mathbf{1} & 11 \end{pmatrix}$$

$$\xrightarrow[\mathbf{r}_2 \to \mathbf{r}_2 - (1/3)\,\mathbf{r}_3]{\mathbf{r}_1 \to \mathbf{r}_1 - (7/3)\,\mathbf{r}_3} \begin{pmatrix} \mathbf{1} & 0 & 0 & -25 \\ 0 & \mathbf{1} & 0 & -4 \\ 0 & 0 & \mathbf{1} & 11 \end{pmatrix}$$

Elementary row operations: There are three elementary row operations on matrices:

R1. Interchange of two rows.

R2. Multiplication of a row by a nonzero scalar.

R3. Addition of a multiple of one row to another row.

By using these operations repeatedly we can bring any matrix into row echelon form. The procedure is illustrated in Example 1, and there are also worked out examples in the Math 250 text [1], Section 1.4.

Rank and nullity: The number of nonzero rows in R, the reduced row-echelon form of A, is called the *rank* of A, and written $\mathrm{rank}(A)$ (it is also the rank of R, since R is already in RREF). This is of course also the number of pivots of A, or the number of rows containing pivots, or the number of columns containing pivots. The *nullity* of A, written $\mathrm{nullity}(A)$, is the number of columns *without* pivots; if A has n columns then $\mathrm{nullity}(A) + \mathrm{rank}(A) = n$.

Remark: Row operations on a matrix may be implemented by multiplying the matrix on the left by *elementary matrices*. These are invertible, and as a consequence we know that if the matrix B is obtained from A by row operations then $B = SA$ for some invertible $m \times m$ matrix S. In particular, if R is the RREF of A, then $R = SA$ and $A = S^{-1}R$ for an invertible S.

3. Solving systems of linear equations

Suppose now that we are given the system of linear equations (1) and want to determine whether or not it has any solutions and, if so, to find them all. The idea is to solve (1) by doing elementary operations on the equations, corresponding to the elementary row operations on matrices: interchange two equations, multiply an equation by a nonzero constant, or add a multiple of one equation to another. What is important is that these operations do not change the set of solutions of the equations, so that we can reduce the equations to simpler form, solve the simple equation, and know that we have found the all solutions of the original equations, but no extraneous ones. Moreover, instead of working with the equations, we can work with the *augmented matrix*:

$$(\,A\,|\,\mathbf{b}\,) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

(It's not necessary to write the vertical bars here, but they remind us that the last column plays a special role.) Simplifying the original set of equations is equivalent to reducing the augmented matrix to RREF. Once this is done, we can easily find the solutions explicitly, if there are any. Equally important, just by looking at the RREF we can determine whether solutions exist and, if so, many of their properties. We will write this symbolically as

$$(\,A\,|\,\mathbf{b}\,) \xrightarrow[\text{operations}]{\substack{\text{elementary}\\ \text{row}}} (\,R\,|\,\mathbf{e}\,)$$

The entire new augmented matrix $(\,R\,|\,\mathbf{e}\,)$ is supposed to be in RREF; this means that we have also reduced A to the RREF matrix R.

Example 2: Suppose we want to solve the equations

$$\begin{aligned} -3x_2 - x_3 &= 1 \\ x_1 + 2x_2 + 3x_3 &= 0 \\ 2x_1 + 2x_2 + 5x_3 &= -3 \end{aligned} \tag{4}$$

The augmented matrix is the one we studied in the example in Example 1, so we already know the reduced row-echelon form for it:

$$(\,A\,|\,\mathbf{b}\,) = \begin{pmatrix} 0 & -3 & -1 & 1 \\ 1 & 2 & 3 & 0 \\ 2 & 2 & 5 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 0 & -25 \\ 0 & \mathbf{1} & 0 & -4 \\ 0 & 0 & \mathbf{1} & 11 \end{pmatrix} = (\,R\,|\,\mathbf{e}\,).$$

The RREF corresponds to the equations

$$\begin{aligned} x_1 &= -25 \\ x_2 &= -4 \\ x_3 &= 11 \end{aligned} \tag{5}$$

This is the solution; notice that it is unique.

In the next examples we will omit the step of row reduction and start with a matrix in reduced row-echelon form.

Example 3: Suppose that the RREF form of the augmented matrix is

$$(\,R\,|\,\mathbf{e}\,) = \begin{pmatrix} \mathbf{1} & 2 & 0 & 1 & | & 5 \\ 0 & 0 & \mathbf{1} & 3 & | & 2 \\ 0 & 0 & 0 & 0 & | & \mathbf{1} \end{pmatrix}.$$

The last equation here is $0 = 1$, which clearly has no solutions: it expresses a contradiction. This is the signal that our original equations have no solutions. Notice that one way to say what has happened here is that the rank of R, which is 2, is less than the rank of $(\,R\,|\,\mathbf{e}\,)$, which is three. In general, we will have no solution precisely if $\operatorname{rank}(R) < \operatorname{rank}(\,R\,|\,\mathbf{e}\,)$.

Example 4: Suppose that the RREF form of the augmented matrix is

$$(\,R\,|\,\mathbf{e}\,) = \begin{pmatrix} 0 & \mathbf{1} & 2 & 0 & 1 & | & 5 \\ 0 & 0 & 0 & \mathbf{1} & 3 & | & 2 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}.$$

Now the idea is to solve for the variables x_2 and x_4, the variables for the columns containing pivots, in terms of the other variables, which are treated as parameters. To remind us that we are treating these variables as parameters, we will give them new names: $x_1 = c_1$, $x_3 = c_2$, and $x_5 = c_3$. Then our solution is

$$x_1 = c_1, \quad x_2 = 5 - 2c_2 - c_3, \quad x_3 = c_2, \quad x_4 = 2 - 3c_3, \quad x_5 = c_3.$$

In vector form,

$$\mathbf{x} = \begin{pmatrix} c_1 \\ 5 - 2c_2 - c_3 \\ c_2 \\ 2 - 3c_3 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 0 \\ 2 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ -1 \\ 0 \\ -3 \\ 1 \end{pmatrix}. \tag{6}$$

Here we have three parameters, one for each column of R which does not contain a pivot: *the number of free parameters in the solution is* $\mathrm{nullity}(A)$. There are $n = 5$ unknowns and $r = \mathrm{rank}(R) = 2$ pivots, and subtracting these numbers indeed gives $n - r = 3$ free parameters.

The pattern here is quite general. A solution will exist if $\mathrm{rank}(R) = \mathrm{rank}(\,R\,|\,\mathbf{e}\,)$, and it will have the general form

$$\mathbf{x} = \mathbf{X} + c_1\mathbf{x}_1 + c_2\mathbf{x}_x + \cdots + c_k\mathbf{x}_k.$$

The free parameters $c_1, \ldots c_k$ are just the original unknowns corresponding to the columns without pivots. Since there are $r = \mathrm{rank}(R) = \mathrm{rank}(A)$ pivots there will be $n - r = \mathrm{nullity}(A)$ free parameters in the solution (that is, $k = n - r$). Since we can choose the parameters freely, we can take $c_1 = c_2 = \cdots = c_k = 0$ and we thus find that $\mathbf{X}$ itself a solution. This is the *particular* solution we discussed in Section 1. If we consider now the homogeneous problem—the same equations, but with $\mathbf{b} = 0$—then we will also have $\mathbf{e} = 0$, and by looking at (6) we can see that we will have $\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_x + \cdots + c_k\mathbf{x}_k$ with the *same* vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$; this means that we have recovered (3).

We summarize:

Theorem 3: Solving linear equations. *Suppose that the augmented matrix* $(\,A\,|\,\mathbf{b}\,)$ *is reduced to the RREF* $(\,R\,|\,\mathbf{e}\,)$. *Then:*

(i) *If* $\mathrm{rank}(R) < \mathrm{rank}(\,R\,|\,\mathbf{e}\,)$, *so that the last nonzero equation is* $0 = 1$, *then the equations have no solutions. This cannot happen if the system is homogeneous.*

(ii) *If* $\mathrm{rank}(R) = \mathrm{rank}(\,R\,|\,\mathbf{e}\,)$ *then the equations have at least one solution. Write* $r = \mathrm{rank}(R) = \mathrm{rank}(A)$*; then the solution is unique if* $n = r$, *i.e., if every column in* R *has a pivot. Otherwise, the equations have a family of solutions with* $k = n - r = \mathrm{nullity}(A)$ *free parameters. The general solution may be written in the form*

$$\mathbf{x} = \mathbf{X} + c_1\mathbf{x}_1 + c_2\mathbf{x}_x + \cdots + c_k\mathbf{x}_k, \tag{7}$$

where X *is a particular solution,* $c_1, \ldots, c_k$ *are the parameters, and* $\mathbf{x}_1, \ldots, \mathbf{x}_k$ *are solutions of the homogeneous equations* $A\mathbf{x} = \mathbf{0}$. *The specific solutions are found by solving the reduced equations for the variables corresponding to the columns with pivots in terms of the other variables, which become the parameters.*

(iii) *The homogeneous system always has at least one solution:* $\mathbf{x} = \mathbf{0}$. *This is the trivial solution. The system has nontrivial solutions if and only if there are columns in* R *which do not contain pivots, that is, if and only if* $r < n$. *The general solution of the homogeneous equation is of the form*

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_x + \cdots + c_k\mathbf{x}_k, \tag{8}$$

with $k = r - n$.

If we know the RREF R of the cofficient matrix A then we can draw some conclusions from Theorem 3 about what may happen for various right hand sides $\mathbf{b}$. If $\mathrm{rank}(A) = m$ then the equations $A\mathbf{x} = \mathbf{b}$ will have a solution for *every* $\mathbf{b}$; if $\mathrm{rank}(A) < m$, so that the bottom row (at least) of R is identically zero, then there will be some $\mathbf{b}$ for which

$A\mathbf{x} = \mathbf{b}$ has no solution. Whenever a solution of $A\mathbf{x} = \mathbf{b}$ exists it will contain nullity(A) free parameters; in particular, there will be a unique solution if and only if nullity$(A) = 0$ or rank$(A) = n$.

4. The case of n equations in n unknowns

Probably the most common systems of linear equations have the same number of equations as unknowns—say n equations in n unknowns. The coefficient matrix A is then *square*, with n rows and n columns. In this case there is a connection between the questions of whether a solution *exists*, and whether a solution which does exist is *unique*. As we shall see, one of two things may happen. Suppose that the augmented matrix has been reduced to RREF $(R\,|\,\mathbf{e})$.

Case 1: rank$(A) = n$**.** Since R is an $n \times n$ matrix in RREF with no zero rows, it must be the identity matrix, so that $(R\,|\,\mathbf{e}) = (I\,|\,\mathbf{e})$. The corresponding equations $x_1 = e_1$, $x_2 = e_2$, ..., $x_n = e_n$ will have a solution $\mathbf{x} = \mathbf{e}$ no matter what $\mathbf{e}$ is, and hence no matter what the original $\mathbf{b}$ was; moreover, the solution is clearly always unique.

Case 2: rank$(A) < n$**.** In this case, the last row of R is a zero row. This means that for some choices of $\mathbf{b}$, the right hand side of the original equations, the vector $\mathbf{e}$ can have one more nonzero component than there are nonzero rows in R, i.e., that the equations will have no solution for some $\mathbf{b}$. On the other hand, if a solution does exist, then because there is a column without a pivot, our solution method will lead to a solution with at least one free parameter—that is, any solution that does exist will not be unique. We have the

Theorem 4: n equations in n unknowns:. *If A is a square matrix then the system of equations $A\mathbf{x} = \mathbf{b}$* **either** *has a unique solution for every $\mathbf{b}$ (Case 1),* **or** *fails to have a solution for some $\mathbf{b}$, and never has a unique solution (Case 2).*

Note, for example, that if we know that for some $\mathbf{b}$ the system $A\mathbf{x} = \mathbf{b}$ has a unique solution, then we must be in Case 1 and we immediately know that it has a solution, and in fact a unique solution, for *every* $\mathbf{b}$. Note also that the homogeneous system $A\mathbf{x} = \mathbf{0}$ can have a nontrivial solution only in Case 2, that is, if and only if rank$(A) = 0$.

There is another way to distinguish between Case 1 and Case 2 which we will use but not prove: **we are in Case 1, that is,** rank$(A) = n$**, only if the determinant of A,** det(A)**, is not zero**.

Much more can be said in Case 1. Suppose that we are in this case, i.e., that rank$(A) = n$. Let us define the vectors $\mathbf{u}_1$, ..., $\mathbf{u}_n$ to be the columns of the $n \times n$ identity matrix:

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{u}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

We know that the system $A\mathbf{x} = \mathbf{u}_i$ has a unique solution, which we will call $\mathbf{v}_i$, that is, $A\mathbf{v}_i = \mathbf{u}_i$. Now consider a matrix B with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$: $B = (\mathbf{v}_1\ \mathbf{v}_2 \cdots\ \mathbf{v}_n)$.

Because of the definition of matrix multiplication, if we compute AB we just multiply each column of B by the matrix A: thus

$$AB = (A\mathbf{v}_1 \; A\mathbf{v}_2 \; A\mathbf{v}_3 \; \cdots \; A\mathbf{v}_m) = (\mathbf{u}_1 \; \mathbf{u}_2 \; \mathbf{u}_3 \; \cdots \; \mathbf{u}_n) = I.$$

Since $AB = I$, A has an inverse, and it is B. (One has to show that also $BA = I$, which is not very hard.)

These ideas also tell us how to compute A^{-1}. First, how do we find $\mathbf{v}_i$? We do Gaussian elimination on the augmented matrix $(\,A \,|\, \mathbf{u}_i\,)$, and $\mathbf{v}_i$, the solution, will just be the last column of the result, that is, the row reduction will be $(\,A \,|\, \mathbf{u}_i\,) \to (\,I \,|\, \mathbf{v}_i\,)$. Doing all these different problems to find all the $\mathbf{v}_i$ is a terrible duplication of effort, however, so we do them all at once:

$$(A|\mathbf{u}_1 \, \mathbf{u}_2 \, \cdots \mathbf{u}_n) \to (I|\mathbf{v}_1 \, \mathbf{v}_2 \, \cdots \mathbf{v}_n) \quad \text{or equivalently} \quad (\,A \,|\, I\,) \to (\,I \,|\, A^{-1}\,).$$

This method of computing A^{-1} is discussed in [2], page 161.

We can conclude that if A is a square matrix then any one of the following conditions is enough to guarantee that we are in Case 1, and hence that in fact all the conditions hold:

C1: The system $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$.

C2: Whenever the system $A\mathbf{x} = \mathbf{b}$ has a solution, the solution is unique.

C3. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

C4: $\text{rank}(A) = n$.

C5: $\text{nullity}(A) = 0$.

C6: A has an inverse matrix A^{-1} satisfying $AA^{-1} = A^{-1}A = I$.

C7: The reduced row-echelon form of A is the identity matrix I.

C8: The determinant of A is not zero.

5. Linear independence

We will spend a lot of time in Math 350 discussing the concept of linear independence, but here we review some of the ideas which were already discussed in Math 250. **The material in this section will not be included in the review quiz.**

Suppose we are given k vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, these might be either row or column vectors, but they are all one or the other, and they all have the same number of components. We then ask the question: can any one of these vectors be expressed as a linear combination of the remaining ones? If so, the vectors are *linearly dependent*, if not, they are *linearly independent*.

Example 5: (a) The vectors $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and $\mathbf{x}_3 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$ are linearly *dependent*, since $\mathbf{x}_3$ can be expressed as a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$: $\mathbf{x}_3 = 3\mathbf{x}_1 - 2\mathbf{x}_2$,

(b) The vectors $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $\mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, are linearly *independent*. For example, we cannot write $\mathbf{x}_1 = a\mathbf{x}_2 + b\mathbf{x}_3$ no matter how we choose a and b, since $a\mathbf{x}_2 + b\mathbf{x}_3 = \begin{pmatrix} 0 \\ a \\ b \end{pmatrix}$ has first component 0, and $\mathbf{x}_1$ has first component 1.

(c) The vectors $\mathbf{x}_1 = \begin{pmatrix} 1 & 5 & -3 & 2 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$, and $\mathbf{x}_3 = \begin{pmatrix} 7 & -1 & 2 & 0 \end{pmatrix}$ are linearly *dependent*, since $\mathbf{x}_2 = 0\mathbf{x}_1 + 0\mathbf{x}_3$. Clearly, a set of vectors in which one vector is $\mathbf{0}$ must be linearly dependent, by the same reasoning.

There is another way to describe linear dependence: the vectors $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_k$ are linearly dependent if there exist scalars $c_1, \ldots, c_k$, **not all zero**, such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0}. \tag{9}$$

The restriction that not all the c_i be zero is important, since we could always make (9) true by taking $c_1 = c_2 = \cdots = c_k = 0$. This new definition of linear dependence is the same as our original definition. For if the vectors are linearly dependent according to our first definition then one of them, say $\mathbf{x}_1$, can be expressed as a linear combination of the others: $\mathbf{x}_1 = d_2\mathbf{x}_2 + d_3\mathbf{x}_3 + \cdots + d_k\mathbf{x}_k$; but then

$$\mathbf{x}_1 - d_2\mathbf{x}_2 - d_3\mathbf{x}_3 - \cdots - d_k\mathbf{x}_k = \mathbf{0},$$

which shows that (9) holds with the coefficients c_i not all zero (since $c_1 = 1$). Conversely, if (9) holds with some coefficient not zero—say, $c_1 \neq 0$—then we can solve the equation for $\mathbf{x}_1$, expressing it as a linear combination of the others:

$$\mathbf{x}_1 = -\left(\frac{c_2}{c_1}\right)\mathbf{x}_2 - \cdots - \left(\frac{c_k}{c_1}\right)\mathbf{x}_k,$$

so that the vectors are linearly dependent by our first definition.

How can we determine if the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly dependent or linearly independent? Here is one way. Suppose that these are column vectors with n components, and build a matrix A with these vectors as columns: $A = (\mathbf{x}_1\,\mathbf{x}_2\,\ldots\,\mathbf{x}_n)$. The matrix A is $n \times k$. To say that (9) holds is just to say that $A\mathbf{c} = \mathbf{0}$, where $\mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$. This means that $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly dependent—i.e., that (9) holds with the c_i not all zero—if the system of equations $A\mathbf{c} = \mathbf{0}$ has a nontrivial solution for $\mathbf{c}$. We know how to use Gaussian elimination to determine whether or not this is true , as discussed above.

Finally, suppose we have n vectors, each with n components, and want to know if they are linearly independent. Then the matrix A is an $n \times n$ square matrix, and we can study it via the ideas of the Section 4. The system $A\mathbf{c} = \mathbf{0}$ has no nontrivial solution if and only if we are in Case 1 (this is condition **C3** for being in case 1), i.e., if the matrix A satisfies any of the conditions **C1–C7**. Note that this means that we could add another condition to the list **C1–C7**, equivalent to all the rest:

C8. The columns of A are linearly independent vectors.

We summarize:

Theorem 5: Linear independence:. *The column vectors* $\mathbf{x}_1, \ldots, \mathbf{x}_k$ *are linearly dependent if the system of equations* $A\mathbf{c} = \mathbf{0}$*, where* $A = (\mathbf{x}_1\, \mathbf{x}_2\, \ldots\, \mathbf{x}_k)$*, has a nontrivial solution. The vectors are linearly independent if the system* $A\mathbf{c} = \mathbf{0}$ *has only the trivial solution* $\mathbf{c} = \mathbf{0}$*. If* $k = n$ *and the vectors are column vectors with* n *components, then they are linearly independent if and only if the matrix* A *satisfies any of the conditions* **C1–C7** *of Section 4.*

Remark: In Theorem 3 we found that $k = \text{nullity}(A) = n - \text{rank}(A)$ vectors are needed to express every solution of the equations $A\mathbf{x} = \mathbf{b}$, and observed that row reduction produced the needed vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$. We want to observe here that these k vectors are **linearly independent**. To see this, consider Example 4 and form a linear combination of the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ produced there:

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ -1 \\ 0 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ -2c_2 - c_3 \\ c_2 \\ -3c_3 \\ c_3 \end{pmatrix}.$$

By looking at the first, third, and fifth components of the final form of this vector we see that if $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = 0$ then necessarily $c_1 = c_2 = c_3 = 0$, and this is precisely linear independence of $\mathbf{x}_1, \mathbf{x}_2$, and $\mathbf{x}_3$. The pattern is the same for any system $A\mathbf{x} = \mathbf{b}$.

6. Exercises

1. [2] Section 3.2, problems 5(a), (c), and (e).

2. [2] Section 3.4, problems 1 and 2. Do problem 2 specifically by the methods used in these notes, that is, by introducing the augmented matrix and then reducing it to reduced row-echelon form. Write the solutions in in the form (7) (as we did in (6)).

3. In (a)-(c) below we suppose that we have been given a system of equations $A\mathbf{x} = \mathbf{b}$ and that we have already reduced the augmented matrix $(\,A \,|\, \mathbf{b}\,)$ to the reduced row-echelon form $(\,R \,|\, \mathbf{e}\,)$ given. In each case, determine (i) whether the original equations have a solution; (ii) if they do have a solution, whether or not it is unique; and (iii) if it is not unique, on how many free parameters there are in the solution. Then write the solution explicitly in the form (7).

(a) $$(\,R \,|\, \mathbf{e}\,) = \left(\begin{array}{ccccc|c} 1 & 5 & -3 & 2 & 8 & 2 \\ 0 & 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

(b) $(R\,|\,\mathbf{e}) = \begin{pmatrix} 1 & 0 & 0 & 0 & | & 2 \\ 0 & 1 & 0 & 0 & | & -1 \\ 0 & 0 & 1 & 0 & | & 3 \\ 0 & 0 & 0 & 1 & | & 4 \end{pmatrix}$

(c) $(R\,|\,\mathbf{e}) = \begin{pmatrix} 0 & 1 & 2 & 0 & -2 & 0 & | & 2 \\ 0 & 0 & 0 & 1 & 3 & 0 & | & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & | & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$

4. In each part below, give a $m \times n$ matrix R *in reduced row-echelon form* satisfying the given condition, or explain briefly why it is impossible to do so.

(a) $m = 3$, $n = 4$, and the equation $R\mathbf{x} = \mathbf{e}$ has a solution for all $\mathbf{e}$.

(b) $m = 3$, $n = 4$, and the equation $R\mathbf{x} = \mathbf{0}$ has a unique solution.

(c) $m = 4$, $n = 3$, and the equation $R\mathbf{x} = \mathbf{e}$ has a solution for all $\mathbf{e}$.

(d) $m = 4$, $n = 3$, and the equation $R\mathbf{x} = \mathbf{0}$ has a unique solution.

(e) $m = 4$, $n = 4$, and the equation $R\mathbf{x} = \mathbf{0}$ has no solution.

(f) $m = 4$, $n = 4$, and the equation $R\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

(g) $m = 4$, $n = 4$, and for every $\mathbf{e}$ the equations $R\mathbf{x} = \mathbf{e}$ have a solution containing a free parameter.

5. Let A be an $m \times n$ matrix of rank r. What can you conclude about m, n, and r (other than $r \leq m$ and $r \leq n$, always true) if the equation $A\mathbf{x} = \mathbf{b}$ has

(a) exactly one solution for some $\mathbf{b}$ and no solution for some other $\mathbf{b}$?

(b) infinitely many solutions for all $\mathbf{b}$?

(c) exactly one solution for every $\mathbf{b}$?

(d) infinitely many solutions for some $\mathbf{b}$ and no solutions for some other $\mathbf{b}$?

6. Suppose that $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions of $A\mathbf{x} = \mathbf{0}$ and that $\mathbf{X}$ is a solution of $A\mathbf{x} = \mathbf{b}$. *Without looking at the these notes or the book*, show that for any constants c_1 and c_2, $c_1\mathbf{x}_1 + c_2\mathbf{x}_2$ is a solution of $A\mathbf{x} = \mathbf{0}$ and that $\mathbf{X} + c_1\mathbf{x}_1 + c_2\mathbf{x}_2$ is a solution of $A\mathbf{x} = \mathbf{b}$.

Some brief answers:

1,2. See the "Answers to Selected Exercises" in [2].

3. (a) no solution, (b) unique solution, (d) solution with 3 parameters.

4. (b), (c), (e), and (g) are impossible.

5. (a) $r = n < m$; (b) $r = m < n$; (c) $r = n = m$; (d) $r < n$ and $r < m$.

References

[1] Spence, L. E., Insel, A. J. and Friedberg, S. H., *Elementary Linear Algebra: A Matrix Approach, 2nd Edition.* Upper Saddle River, Prentice-Hall, 2007.

[2] Friedberg, S. H., Insel, A. J. and Spence, L. E., *Linear Algebra, 4th Edition.* Upper Saddle River, Prentice-Hall, 2007.