March 1998

Systems of Equations and Finding Bases

by
Professor Francis J. Narcowich
Department of Mathematics
Texas A&M University

§1. Solutions to linear Systems

In this section we review facts about solutions to system of linear equations. We begin by converting the system

$$\mathbf{S}: \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

into the equivalent matrix form,

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} X = b\,.$$

This of course is the equation $AX = b$, with A being the $m \times n$ coefficient matrix for S, X being the $n \times 1$ vector of unknowns, and b being the $m \times 1$ vector of b_j's. The augmented form of the system is

$$[A|b] = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

In view of the connection between row operations and operations on the individual equations comprising the system S, any matrix $[A'|b']$ that is row equivalent to the original system $[A|b]$ is the augmented matrix for a system S' equivalent (i.e., having the same solution set) to S. Let us state this formally.

THEOREM 1.1: *If $[A|b]$ and $[A'|b']$ are augmented matrices for two linear systems of equations S and S', and if $[A|b]$ and $[A'|b']$ are row equivalent, then S and S' are equivalent systems.*

For the system S, reduced-echelon matrix equivalent to the augmented matrix $[A|b]$ is very important; it tells what the set of solutions is, and it also gives information about the dimensions of the row and column spaces. We shall discuss these things in the next few paragraphs. Before doing so, we need to make one important definition:

DEFINITION: The *rank* of a matrix M is the number of nonzero rows in its reduced echelon form; we will denote the rank of M by $\mathrm{rank}(M)$.

One can use Theorem 1.1 to obtain the following result, which we state without proof.

THEOREM 1.2: *Consider the system S. We have these possibilities:*

i. *S is inconsistent if and only if* $\mathrm{rank}(A) < \mathrm{rank}([A|b])$.

ii. *S has a unique solution if and only if* $\mathrm{rank}(A) = \mathrm{rank}([A|b]) = n$.

iii. *S will have infinitely many solutions if and only if* $\mathrm{rank}(A) = \mathrm{rank}([A|b]) < n$.

We need to illustrate the use of this theorem. To do that, look at the simple systems below.

$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 3x_1 - 2x_2 &= -1 \end{aligned} \qquad \begin{aligned} 3x_1 + 2x_2 &= 3 \\ -6x_1 - 4x_2 &= 0 \end{aligned} \qquad \begin{aligned} 3x_1 + 2x_2 &= 3 \\ -6x_1 - 4x_2 &= -6 \end{aligned}$$

The augmented matrices for these systems are, respectively,

$$\left(\begin{array}{cc|c} 2 & 3 & 8 \\ 3 & -2 & -1 \end{array}\right) \qquad \left(\begin{array}{cc|c} 3 & 2 & 3 \\ -6 & -4 & 0 \end{array}\right) \qquad \left(\begin{array}{cc|c} 3 & 2 & 3 \\ -6 & -4 & -6 \end{array}\right).$$

Applying the row-reduction algorithm yields the row-reduced form of each of these augmented matrices. The result is, again respectively,

$$\left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right) \qquad \left(\begin{array}{cc|c} 1 & 2/3 & 0 \\ 0 & 0 & 1 \end{array}\right) \qquad \left(\begin{array}{cc|c} 1 & 2/3 & 1 \\ 0 & 0 & 0 \end{array}\right).$$

From each of these row-reduced versions of the augmented matrices, one can read off the rank of the coefficient matrix as well as the rank of the augmented matrix. Applying Theorem 1.2 to each of these tells us the number of solutions to expect for each of the corresponding systems. We summarize our findings in the table below.

System	$\mathrm{rank}(A)$	$\mathrm{rank}([A\|b])$	n	*# of solutions*
First	2	2	2	1
Second	1	2	2	0 *(inconsistent)*
Third	1	1	2	∞

We now turn to the discussion of homogeneous sytems, which are systems having the vector $b = 0$. By simply plugging $X = 0$ into the equation $AX = 0$, we see that every homogeneous system has at least *one* solution, the *trivial* solution $X = 0$. Are there any others? To answer this, we can use this corollary to Theorem 1.2:

COROLLARY 1.3: *A homogeneous system of equations $AX = 0$ will have a unique solution, the trivial solution $X = 0$, if and only if* $\mathrm{rank}(A) = n$. *In all other cases, it will have infinitely many solutions. As a consequence, if $n > m$—i.e., if the number of unknowns is larger than the number of equations—, then the system will have infiniely many solutions.*

PROOF: Since $X = 0$ is always a solution, case (i) of Theorem 1.2 is eliminated as a possibility. Therefore, we must always have $\text{rank}(A) = \text{rank}([A|0]) \leq n$. By Theorem 1.2, case (ii), equality will hold if and only if $X = 0$ is the only solution. When it does not hold, we are always in case (iii) of Theorem 1.2; there are thus infinitely many solutions for the system. If $n > m$, then we need only note that $\text{rank}(A) \leq m < n$ to see that the system has to have infinitely solutions. ∎

§2. Solving Linear Systems

Thus far we have discussed how many solutions a system of linear equations has. In this section, we will review finding these solutions. The key to the whole process is row-reducing the augmented matrix for the original system, S.

Let us look at an example. Suppose that we have found that our system has an augmented matrix $[A|b]$ that is row equivalent to the matrix

$$[A'|b'] = \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -3 & 7 \\ 0 & 0 & 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

We now convert this back to a system, one that is of course equivalent to whatever one we started with. The result is the following system of equations:

$$\begin{aligned} x_1 + 2x_2 &= -1 \\ x_3 - 3x_5 &= 7 \quad . \\ x_4 + 3x_5 &= 4 \end{aligned}$$

Notice that the variables corresponding to the leading columns appear in this set only once. That means that they can be solved for in terms of the other variables. Solving for these "leading" variables results in the system

$$\begin{aligned} x_1 &= -2x_2 - 1 \\ x_3 &= 3x_5 + 7 \quad . \\ x_4 &= -3x_5 + 4 \end{aligned}$$

It turns out that by assigning arbitrary values to the *non*-leading variables gives us all possible solutions to the system. It is customary to show that this assignment has been made by assigning new letters to the non-leading variables. In our example, we could set $x_2 = s$, $x_5 = t$, and then rewrite the whole solution in the column form shown below.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2s-1 \\ s \\ 3t+7 \\ -3t+4 \\ t \end{pmatrix} = s\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 0 \\ 3 \\ -3 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \\ 7 \\ 4 \\ 0 \end{pmatrix}.$$

Written in this way, we see that if we set $s = t = 0$, we get a *particular* solution to the original system. When this column is subtracted off, what is left is a solution to the corresponding *homogeneous* system. This happens in every case: A solution to $AX = b$ may be written $X_p + X_h$, where X_p is a fixed column vector satisfying $AX_p = b$, and X_h runs over all solutions to $AX_h = 0$. This is exactly analogous to what happens in the case of linear differential equations.

§3. Applications to Finding Bases

Row-reduction methods can be used to find bases. Let us now look at an example illustrating how to obtain information about bases for the *null-space* or *kernel*, the *column space* or *range*, and the *row-space* of a matrix A, given the row-reduced matrix R equivalent to it. To begin, note th at using row reduction gives

$$A = \begin{pmatrix} \mathbf{1} & \mathbf{1} & 2 & 0 \\ \mathbf{2} & \mathbf{4} & 2 & 4 \\ \mathbf{2} & \mathbf{1} & 5 & -2 \end{pmatrix} \Longleftrightarrow R = \begin{pmatrix} \mathbf{1} & 0 & 3 & -2 \\ 0 & \mathbf{1} & -1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

3.1. The Null Space. Because A and R are row equivalent, the null-spaces for A and R are identical. The equations we get from finding the null-space of R are

$$\begin{aligned} x_1 + 3x_3 - 2x_4 &= 0 \\ x_2 - x_3 + 2x_4 &= 0. \end{aligned}$$

The dependent (constrained) variables correspond to the columns containing the leading entries (these are in boldface) in R; these are x_1 and x_2. The remaining variables, x_3 and x_4, are independent (free) variables. To emphasize this, we assign them new labels, $x_3 = s$ and $x_4 = t$. Solving the system obtained above, we get

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = s \begin{pmatrix} -3 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

From this equation, it is easy to show that the vectors multiplying s and t form a basis for the null space.

3.2. The column Space. Using the fact that $AX = 0$ for arbitrary s and t, we also obtain these equations relating the four columns of A:

$$C_3 = 3C_1 - C_2 \quad \text{and} \quad C_4 = -2C_1 + 2C_2\,.$$

Thus the column space is spanned by the set $\{C_1, C_2\}$. (C_1 and C_2 are in boldface in the matrix A above.) This set is also linearly independent because the equation

$$0 = x_1C_1 + x_2C_2 = x_1C_1 + x_2C_2 + 0C_3 + 0C_4 = A\,(\,x_1 \quad x_2 \quad 0 \quad 0\,)^T$$

implies that $(\,x_1 \quad x_2 \quad 0 \quad 0\,)^T$ is in the null space of A. Matching this vector with the general form of a vector in the null space shows that the corresponding s and t are 0, and therefore so are x_1 and x_2. It follows that $\{C_1, C_2\}$ is linearly independent. Since it spans the columns as well, it is a basis for the *range* of A. Note that these columns correspond to the *dependent* variables in the problems, x_1 and x_2. This is no accident. The argument that we used can be employed to show that this is true in general:

THEOREM 3.1: *In a matrix A, the columns of A that correspond to the dependent variables in the associated homogeneous problem, $AX = 0$, form a basis for the column-space (range) of A.*

3.3. The Row Space. Row operations *preserve* the row space. They involve either taking linear combinations of rows or interchanging rows, both of which leave the span of the rows unchanged. (They do change the column space, however.) Because row operations are reversible, the orignal set of rows can be obtained from the rows of the reduced echelon form of the matrix. Thus, the rows of the latter also span the row space. Moreover, they are linearly independent.

THEOREM 3.2: *The rows of the reduced echelon form of a matrix comprise a basis for its row space.*

Returning to the example we started with, a basis for the row space of A consists of the non-zero rows of R, $\{(\,1 \quad 0 \quad 3 \quad -2\,), (\,0 \quad 1 \quad -1 \quad 2\,)\}$.