LECTURE 5

Solving Systems of Linear Equations

Recall that we introduced the notion of matrices as a way of standardizing the expression of systems of linear equations. In today's lecture I shall show how this matrix machinery can also be used to solve such systems. However, before we embark on solving systems of equations via matrices, let me remind you of what such solutions should look like.

1. The Geometry of Linear Systems

Consider a linear system of m equations and n unknowns:

(5.1)
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

What does the solution set look like? Let's examine this question case by case, in a setting where we can easily visualize the solution sets

- 0 equations in 3 unknowns. This would correspond to a situation where you have 3 variables x_1, x_2, x_3 with no relations between them. Being free to choose whatever value we want for each of the 3 variables, it's clear that the solutions set is just $\mathbb{R}^3$, the 3-dimensional space of ordered sets of 3 real numbers.
- 1 equation in 3 unknowns.
$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
In this case, use the equation to express one variable, say x_3, in terms of the other variables;
$$x_3 = \frac{1}{a_{13}}\left(b_1 - a_{11}x_1 + a_{12}x_2\right)$$
The remaining variables x_1, and x_2 are then unrestricted. Letting these variables range freely over $\mathbb{R}$ will then fill out a 2-dimensional *plane* in $\mathbb{R}^2$. Thus, in the case of 1 equation in 3 unknowns we have a 2-dimensional plane as a solution space.
- 2 equations in 3 unknowns.
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned}$$
As in the preceding example, the solution set of each individual equation will be a 2-dimensional plane. The solution set of the pair of equation will be the *intersection* of these two planes. (For points common to both solution sets will be points corresponding to the solutions of both equations.) Here there will be three possibilities:
  1. The two planes do not intersect. In other words, the two planes are parallel but distinct. Since they share no common point, there is no simultaneous solutoin of both equations.
  2. The intersection of the two planes is a line. This is the generic case.
  3. The two planes coincide. In this case, the two equations must be somehow redundant.

Thus we have either no solution, a 1-dimensional solution space (i.e. a line) or a 2-dimensional solution space.

- 3 equations and 3 unknowns. In this case, the solution set will correspond to the intersection of the three planes corresponding to the 3 equations. We will have four possiblities:
  (1) The three planes share no common point. In this case there will be no solution.
  (2) The three planes have one point in common. This will be the generic situation.
  (3) The three planes share a single line.
  (4) The three planes all coincide.

  Thus, we either have no solution, a 0-dimensional solution (i.e., a point), a 1-dimensional solution (i.e. a line) or a 2-dimensional solution.

We now summarize and generalize this discussion as follows.

THEOREM 5.1. *Consider a linear system of m equations and n unknowns:*

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

The solution set of such a system is either:

(1) *The empty set $\{\}$; i.e., there is no solution.*
(2) *A hyperplane of dimension greater than or equal to $(n-m)$ (= the number of unknowns minus the number of equations).*

2. Elementary Row Operations

In the preceding lecture we remarked that our new-fangled matrix algebra allowed us to represent linear systems such as (**??**) succintly as matrix equations:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{5.2}$$

For example the linear system

$$\begin{aligned} x_1 + 3x_2 &= 3 \\ x_1 + 2x_2 &= 1 \end{aligned} \tag{5.3}$$

can be represented as

$$\begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{5.4}$$

Now, when solving linear systems like (**??**) it is very common to create new but equivalent equations by, for example, mulitplying by a constant or adding one equation to another. In fact, we have

THEOREM 5.2. *Let S be a system of m linear equations in n unknowns and let S' be another system of m equations in n unknowns obtained from S by applying some combination of the following operations:*

- *interchanging the order of two equations*
- *multiplying one equation by a non-zero constant*
- *replacing a equation with the sum of itself and a multiple of another equation in the system.*

Then the solution sets of S ane S' are identical.

In particular, the solution set of (**??**) is equivalent to the solution set of

$$\begin{aligned} x_1 + 3x_2 &= 3 \\ -x_1 - 2x_2 &= -1 \end{aligned} \tag{5.5}$$

(where we have multiplied the second equation by -1), as well as the solution set of

$$\begin{aligned} x_1 + 3x_2 &= 3 \\ x_2 &= 2 \end{aligned} \tag{5.6}$$

(where we have replaced the second row in (**??**) by its sum with the first row), as well as the solution of

$$\begin{aligned} x_1 &= -3 \\ x_2 &= 2 \end{aligned} \tag{5.7}$$

(where we have replaced the first row in (**??**) by its sum with -3 times the second row). We can thus solve a linear system by means of the elementary operations described in the theorem above.

Now because there is a matrix equation corresponding to every system of linear equations, each of the operations described in the theorem above corresponds to a matrix operation. To convert these operations in our matrix language, we first associate with a linear system

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

an **augmented matrix**

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

This is just the $m \times (n+1)$ matrix that's obtained appending the column vector $\mathbf{b}$ to the columns of the $m \times n$ matrix $\mathbf{A}$. We shall use the notation $[\mathbf{A} \mid \mathbf{b}]$ to denote the augmented matrix of $\mathbf{A}$ and $\mathbf{b}$.

The augmented matrices corresponding to equations (**??**), (**??**), and (**??**) are thus, respectively

$$\left[\begin{array}{cc|c} 1 & 3 & 3 \\ 1 & 2 & 1 \end{array}\right] \quad ,$$

$$\left[\begin{array}{cc|c} 1 & 3 & 3 \\ 0 & 1 & 2 \end{array}\right] \quad ,$$

and

$$\left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 2 \end{array}\right]$$

From this example, we can see that the operations in Theorem 5.2 translate to the following operations on the corresponding augmented matrices:

- Row Interchange: the interchange of two rows of the augmented matrix
- Row Scaling: multiplication of a row by a non-zero scalar
- Row Addition: replacing a row by its sum with a multiple of another row

Henceforth, we shall refer to these operations as **Elementary Row Operations**.

3. Solving Linear Equations

Let's now reverse directions and think about how to recognize when the system of equations corresponding to a given matrix is easily solved. (To keep our discussion simple, in the examples given below we consider systems of n equations in n unknowns.)

3.1. Diagonal Matrices. A matrix equation $\mathbf{Ax} = \mathbf{b}$ is trivial to solve if the matrix $\mathbf{A}$ is purely diagonal. For then the augmented matrix has the form

$$\mathbf{A} = \left[\begin{array}{cccccc|c} a_{11} & 0 & \cdots & 0 & 0 & b_1 \\ 0 & a_{22} & \cdots & 0 & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & 0 & b_{n-1} \\ 0 & 0 & \cdots & 0 & a_{nn} & b_n \end{array}\right]$$

the corresponding system of equations reduces to

$$\begin{aligned} a_{11}x_1 &= b_1 \quad \Rightarrow \quad x_1 = \frac{b_1}{a_{11}} \\ a_{22}x_2 &= b_2 \quad \Rightarrow \quad x_2 = \frac{b_2}{a_{22}} \\ &\vdots \\ a_{nn}x_n &= b_n \quad \Rightarrow \quad x_n = \frac{b_n}{a_{nn}} \end{aligned}$$

3.2. Lower Triangular Matrices. If the coefficient matrix $\mathbf{A}$ has the form

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n,2} & \cdots & a_{n,n-1} & a_{nn} \end{bmatrix}$$

(with zeros everywhere above the diagonal from a_{11} to a_{nn}), then it is called **lower triangular**. A matrix equation $\mathbf{Ax} = \mathbf{b}$ in which $\mathbf{A}$ is lower triangular is also fairly easy to solve. For it is equivalent to a system of equations of the form

$$\begin{aligned} a_{11}x_1 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

To find the solution of such a system one solves the first equation for x_1 and then substitutes its solution b_1/a_{11} for the variable x_1 in the second equation

$$a_{21}\left(\frac{b_1}{a_{11}}\right) + a_{22}x_2 = b_2 \quad \Rightarrow \quad x_2 = \frac{1}{a_{22}}\left(b_2 - \frac{a_{21}b_1}{a_{11}}\right)$$

One can now substitute the numerical expressions for x_1, and x_2 into the third equation and get a numerical expression for x_3. Continuing in this manner we can solve the system completely. We call this method **solution by forward substitution**.

3.3. Upper Triangular Matrices. We can do a similar thing for systems of equations characterized by an upper triangular matrices $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

(that is to say, a matrix with zero's everywhere below and to the left of the diagonal), then the corresponding system of equations will be

$$\begin{aligned} a_{11}x_1 + a_{12}x2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n &= b_{n-1} \\ a_{nn}x_n &= b_n \end{aligned}$$

which can be solved by substituting the solution of the last equation

$$x_n = \frac{b_n}{a_{nn}}$$

into the preceding equation and solving for x_{n-1}

$$x_{n-1} = \frac{1}{a_{n-1,n-1}} \left(b_{n-1} - a_{n-1,n} \left(\frac{b_n}{a_{nn}} \right) \right)$$

and then substituting this result into the third from last equation, etc. This method is called **solution by back-substitution**.

3.4. Solution via Row Reduction.

3.4.1. *Gaussian Reduction.* In the general case a matrix will be neither be upper triangular or lower triangular and so neither forward- or back-substituiton can be used to solve the corresponding system of equations. However, using the *elementary row operations* discussed in the preceding section we can always convert the augmented matrix of a (self-consistent) system of linear equations into an equivalent matrix that is upper triangular; and having done that we can then use back-substitution to solve the corresponding set of equations. We call this method **Gaussian reduction**. Let me demonstate the Gaussian method with an example.

EXAMPLE 5.3. Solve the following system of equations using Gaussian reduction.

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

First we write down the corresponding augmented matrix.

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 2 \\ 2 & -1 & -1 & -3 \end{array} \right]$$

We now use elementary row operations to convert this matrix into one that is upper triangular.

Adding -1 times the first row to the second row produces

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 2 & 2 \\ 2 & -1 & -1 & -3 \end{array} \right]$$

Adding -2 times the first row to the third row produces

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 2 & 2 \\ 0 & -3 & 1 & -3 \end{array}\right]$$

Adding $-\frac{3}{2}$ times the second row to the third row produces

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 2 & 2 \\ 0 & 0 & -2 & -6 \end{array}\right]$$

The augmented matrix is now in upper triangular form. It corresponds to the following system of equations

$$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ -2x_2 + 2x_3 &= 2 \\ -2x_3 &= -6 \end{aligned}$$

which can easily be solved via back-substitution:

$$\begin{array}{rcl} -2x_3 = -6 & \Rightarrow & x_3 = 3 \\ \Rightarrow & -2x_2 + 6 = 2 & \Rightarrow \quad x_2 = 2 \\ \Rightarrow & x_1 + 2 - 3 = 0 & \Rightarrow \quad x_1 = 1 \end{array}$$

In summary, solution by Gaussian reduction consists of the following steps

(1) Write down the augmented matrix corresponding to the system of linear equations to be solved.
(2) Use elementary row operations to convert the augmented matrix $[\mathbf{A} \mid \mathbf{b}]$ into one that is upper triangular. This is acheived by systematically using the first row to eliminate the first entries in the rows below it, the second row to eliminate the second entries in the row below it, etc.
(3) Once the augmented matrix has been reduced to upper triangular form, write down the corresponding set of linear equations and use back-substitution to complete the solution.

Remark 5.4. The text refers to matrices that are upper triangular as being in **row-echelon form**.

Definition 5.5. *A matrix is in* ***row-echelon form*** *if it satisfies two conditions:*

(1) *All rows containing only zeros appear below rows with non-zero entries*
(2) *The first non-zero entry in a row appears in a column to the right of the first non-zero entry of any preceding row.*

For such a matrix. the first non-zero entry in a row is called the ***pivot*** *for that row.*

3.4.2. *Gauss-Jordan Reduction.* In the Gauss Reduction method described above, one carries out row operations until the augmented matrix is upper triangular and then finishes the problem by converting the problem back into a linear system and using back-substitution to complete the solution. Thus, row reduction is used to carry out about half the work involved in solving a linear system.

It is also possible to use only row operations to construct a solution of a linear system $\mathbf{Ax} = \mathbf{b}$. This technique is called **Gauss-Jordan reduction**. The idea is this

(1) Write down the augmented matrix corresponding to the system of linear equations to be solved.
(2) Use elementary row operations to convert the augmented matrix $[\mathbf{A} \mid \mathbf{b}]$ into one that is upper triangular. This is acheived by systematically using the first row to eliminate the first entries in

the rows below it, the second row to eliminate the second entries in the row below it, etc.

$$[\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right] \xrightarrow{\text{row operations}} \left[\begin{array}{cccc|c} a'_{11} & a'_{12} & \cdots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \cdots & a'_{2n} & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a'_{nn} & b'_n \end{array}\right]$$

(3) Continue to use row operations on the aumented matrix until all the entries above the diagonal of the first factor have been eliminated, and only 1's appear along the diagonal

$$\left[\begin{array}{cccc|c} a'_{11} & a'_{12} & \cdots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \cdots & a'_{2n} & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a'_{nn} & b'_n \end{array}\right] \xrightarrow{\text{row operations}} \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b''_1 \\ 0 & 1 & \cdots & 0 & b''_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b''_n \end{array}\right] = [\mathbf{I} \mid \mathbf{b}'']$$

(4) The solution of the linear system corresponding to the augmented matrix $[\mathbf{I} \mid \mathbf{b}'']$ is trivial

$$\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{b}''$$

Moreover, since $[\mathbf{I} \mid \mathbf{b}'']$ was obtained from $[\mathbf{A} \mid \mathbf{b}]$ by row operations, $\mathbf{x} = \mathbf{b}''$ must also be the solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Example 5.6. Solve the following system of equations using Gauss-Jordan reduction.

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

First we write down the corresponding augmented matrix.

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 2 \\ 2 & -1 & -1 & -3 \end{array}\right]$$

In the preceding example, I demonstated that this augmented matrix is equivalent to

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 2 & 2 \\ 0 & 0 & -2 & -6 \end{array}\right]$$

We'll now continue to apply row operations until the first block has the form of a 3×3 identity matrix.

Multiplying the second and third rows by $-\frac{1}{2}$ yields

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Replacing the first row by its sum with -1 times the second row yields

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Replacing the second row by its sum with the third row yields

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

The system of equations corresponding to this augmented matrix is

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 2 \\ x_3 &= 3 \end{aligned}$$

This is our solution.

To maintain contact with the definitions used in the text, we have

DEFINITION 5.7. *A matrix is in* ***reduced row-echelon form*** *if it is in row-echelon form with all pivots equal to* 1 *and with zeros in every column entry above and below the pivots.*

THEOREM 5.8. *Let* $\mathbf{Ax} = \mathbf{b}$ *be a linear system and suppose* $[\mathbf{A}' \mid \mathbf{b}']$ *is row equivalent to* $[\mathbf{A} \mid \mathbf{b}]$ *with* $\mathbf{A}'$ *a matrix in row-echelon form. Then*

(1) The system $\mathbf{Ax} = \mathbf{b}$ is inconsistent if and only if the augmented matrix $[\mathbf{A}' \mid \mathbf{b}']$ has a row with all entries 0 to the left of the partition and a non-zero entry to the right of the partition.
(2) If $\mathbf{Ax} = \mathbf{b}$ is consistent and every column of $\mathbf{A}'$ contains a pivot, the system has a unique solution.
(3) If $\mathbf{Ax} = \mathbf{b}$ is consistent and some column of $\mathbf{A}'$ has no pivot, the system has infinitely many solutions, with as many free variables as there are pivot-free columns in $\mathbf{A}'$.

4. Solving Linear Systems: The Procedure

Let me now set down in a succinct form our procedure for solving linear systems.

1. From the equations of the linear system, construct the corresponding augmented matrix

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{array}\right\} \rightarrow \quad \left\{ \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ \vdots & \ddots & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array}\right]\right.$$

2. Row reduce the augmented matrix to Reduced Row Echelon Form. Identify the variables corresponding to the columns of the RREF that **do not** contain pivots as the free variables of the solution. In what follows we shall use the following nomenclature:

variables corresponding to columns of the REF which contain pivots	$\longleftrightarrow$	"fixed variables"
variables corresponding to columns of the REF which **do not** contain pivots	$\longleftrightarrow$	"free variables"

3. Write down the equations of the RREF augmented matrix in such a way that the fixed variables are kept on the left hand side and the free variables have been moved to the right hand side (along with the constants appearing in the last column).

4. Now use the equations from Step 3 to write down a typical solution vector

$$\mathbf{x} = \left[\begin{array}{c} x_1 \\ \vdots \\ x_n \end{array}\right] = ...$$

wherein the fixed variables have been replaced by their expressions in terms of the free variables (using the equations of the preceding step).

5. Decompose the solution vector $\mathbf{x}$ obtained in Step 4 as a constant vector (corresponding to the last column of the augmented matrix) plus a sum of constant vectors with the free variables as scalar factors.

Let me the last three steps with an explicit example. Suppose the augmented matrix of a $3x4$ linear system had row reduced to

(Step 2)
$$\left[\begin{array}{cccc|c} 1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The pivots of this matrix live in columns 1 and 4. Therefore the fixed variables of the solution will be x_1 and x_4 and the free variables will be x_2 and x_3. Following the instruction of Step 3, we obtain

(Step 3)
$$\left.\begin{aligned} x_1 - 2x_3 &= 1 \\ x_4 &= 0 \\ 0 &= 0 \end{aligned}\right\} \quad \Rightarrow \quad \left\{\begin{aligned} x_1 &= 1 + 2x_3 \\ x_4 &= 0 \end{aligned}\right.$$

A solution vector will then have the form

(Step 4)
$$\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array}\right] = \left[\begin{array}{c} 1 + 2x_3 \\ x_2 \\ x_3 \\ 0 \end{array}\right]$$

We now expand the right hand side of Step 4

(Step 5)
$$\mathbf{x} = \left[\begin{array}{c} 1 + 2x_2 \\ x_2 \\ x_3 \\ 0 \end{array}\right] = \left[\begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \end{array}\right] + x_2 \left[\begin{array}{c} 0 \\ 1 \\ 0 \\ 0 \end{array}\right] + x_3 \left[\begin{array}{c} 2 \\ 0 \\ 1 \\ 0 \end{array}\right]$$

Note that the final result in Step 5 displays the solution vectors as vectors living on a hyperplane. More generally, if the RREF of the augmented matrix had k columns without pivots, there would be exactly k free parameters in the solution, and the solution vectors would correspond to points on a k-dimensional hyperplane

$$\mathbf{x} = \mathbf{x}_0 + s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_k\mathbf{v}_k$$

(each of the scalar multipliers s_i being a free parameter in the solution).

5. Elementary Matrices

I shall now show that all elementary row operations can be carried out by means of matrix multiplication. Even though carrying out row operations by matrix multiplication will be grossly inefficient from a calculation point of view, this equivalence remains an important theoretical fact. (As you'll see in some of the forthcoming proofs.)

DEFINITION 5.9. *An* ***elementary matrix*** *is a matrix that can be obtained from the identity matrix by means of a single row operation.*

THEOREM 5.10. *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix, and let* $\mathbf{E}$ *be an* $m \times m$ *elementary matrix. Then multiplication of* $\mathbf{A}$ *on the left by* $\mathbf{E}$ *effects the same elementary row operation on* $\mathbf{A}$ *as that which was performed on the* $m \times m$ *identity matrix to produce* $\mathbf{E}$.

COROLLARY 5.11. *If* $\mathbf{A}'$ *is a matrix that was obtained from* $\mathbf{A}$ *by a sequence of row operations then there exists a corresponding sequence of elementary matrices* $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_r$ *such that*

$$\mathbf{A}' = \mathbf{E}_r \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}$$