eNote 6

Systems of Linear Equations

(Updated 24.9.2021 David Brander)

6.1 Linear Equations

Remark 6.1 The Common Notion $\mathbb{L}$

Defnitions and rules in this eNote are valid both for the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$. The set of real numbers and the set of complex numbers are examples of *fields*. Fields have common calculation rules concerning elementary arithmetic rules (the same rules as those for $\mathbb{C}$ described in Theorem 1.12 in eNote 1). In the following when we use the symbol $\mathbb{L}$ it means that the notion is valid both for the set of real numbers and for the set of complex numbers.

A ***linear equation*** with n unknowns $x_1, x_2, \ldots x_n$ is an equation of the form

$$a_1 \cdot x_1 + a_2 \cdot x_2 + \ldots + a_n \cdot x_n = b\,. \tag{6-1}$$

The numbers $a_1, a_2, \ldots, a_n$ are called the *coefficients* and the number b is, in this context, called *the right hand side*. The coefficients and the right hand side are considered known in contrast to the unknowns. The equation is called *homogeneous* if $b = 0$, else *inhomogeneous*.

|||| Definition 6.2 Solution to a Linear Equation

By a *solution* to the equation

$$a_1 \cdot x_1 + a_2 \cdot x_2 + \ldots + a_n \cdot x_n = b\,. \tag{6-2}$$

we shall understand an n-tuple $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{L}^n$ that by substitution into the equation makes the left hand side of the equation equal to the right hand side.

By the ***general solution*** or just the ***solution set*** we understand the set of all solutions to the equation.

|||| Example 6.3 The Equation for a Straight Line in the Plane

An example of a linear equation is the equation for a straight line in the (x, y)-plane:

$$y = 2\,x + 5\,. \tag{6-3}$$

Here y is isolated on the left hand side and the coefficients 2 and 5 have well known geometrical interpretations. But the equation could also be written

$$-2\,x_1 + 1\,x_2 = 5 \tag{6-4}$$

where x and y are substituted by the more general names for unknowns, x_1 and x_2, and the equation is of the form (6-1).

The solution set for the equation (6-3) is of course the coordinate set for all points on the line - by substitution they will satisfy the equation in contrast to all other points!

|||| Example 6.4 Trivial and Inconsistent Equations

The linear equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 0 \;\Leftrightarrow\; 0 = 0 \tag{6-5}$$

where all coefficients and the right hand side are 0, is an example of a ***trivial*** equation. The solution set of the equation consists of all $\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathbb{L}^4$.

If all the coefficients of the equation are 0 but the right hand side is non-zero, the equation is an ***inconsistent*** equation, that is, an equation without a solution. An example is the equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 1 \;\Leftrightarrow\; 0 = 1\,. \tag{6-6}$$

When you investigate linear equations, you can use the usual ***rule of conversion*** for equations: The set of solutions for the equation is not changed if you add the same number to both sides of the equality sign, and you do not change the solution set if you multiply both sides of the equality sign by a non-zero constant.

All linear equations that are not inconsistent and which contain more than one solution, have infinitely many solutions. The following example shows how the solution set in this case can be written.

|||| Example 6.5 Infinitely Many Solutions in Standard Parameter Form

We consider an inhomogeneous equation with three unknowns:

$$2\,x_1 - x_2 + 4\,x_3 = 5\,. \tag{6-7}$$

By substitution of $x_1 = 1$, $x_2 = 1$ and $x_3 = 1$ into the equation (6-7) we see that $\mathbf{x} = (1,1,1)$ is a solution. But by this we have not found the general solution, because $\mathbf{x} = (\frac{1}{2},0,1)$ is also a solution. How can we describe the complete set of solutions?

First we isolate x_1:

$$x_1 = \tfrac{5}{2} + \tfrac{1}{2}\,x_2 - 2\,x_3\,. \tag{6-8}$$

To every choice of x_2 and x_3 corresponds exactly one x_1. For example, if we set $x_2 = 1$ and $x_3 = 4$, then $x_1 = -5$. This means that the 3-tuple $(-5,1,4)$ is a solution. Therefore we can consider x_2 and x_3 *free parameters* that together determine the value of x_1. Therefore we **rename** x_2 and x_3 to the parameter names s and t, respectively: $s = x_2$ and $t = x_3$. Then x_1 can be expressed as:

$$x_1 = \tfrac{5}{2} + \tfrac{1}{2}\,x_2 - 2\,x_3 = \tfrac{5}{2} + \tfrac{1}{2}\,s - 2\,t\,. \tag{6-9}$$

Now we can write the general solution to (6-7) in the following ***standard parameter form***:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ 0 \\ 0 \end{bmatrix} + s \cdot \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + t \cdot \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \text{ with } s,t \in \mathbb{L}\,. \tag{6-10}$$

Note that the parameter form of the middle equation $x_2 = 0 + s \cdot 1 + t \cdot 0$ only expresses the renaming $x_2 \to s$. Similarly, the last equation only expresses the renaming $x_3 \to t$.

If we consider the equation (6-7) to be an equation for a plane in space, then the equation (6-10) is a *parametric representation* for the same plane. The first column on the right hand side is the *initial point* in the plane, and the two last columns are *directional vectors* for the plane. This is elaborated in the eNote 10 *Geometric Vectors*.

6.2 A System of Linear Equations

A ***system of linear equations*** consisting of m linear equations with n unknowns is written in the form

$$\begin{aligned} a_{11} \cdot x_1 + \, a_{12} \cdot x_2 + \ldots + \, a_{1n} \cdot x_n &= b_1 \\ a_{21} \cdot x_1 + \, a_{22} \cdot x_2 + \ldots + \, a_{2n} \cdot x_n &= b_2 \\ &\vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \ldots + a_{mn} \cdot x_n &= b_m \end{aligned} \tag{6-11}$$

The system has m *rows*, each of which contains an equation. The n unknowns, denoted $x_1, x_2, \ldots x_n$, are present in each of the m equations (unless some of the coefficients are zero, and we choose not to write down the zero terms). The coefficient of x_j in the equation in row number i is denoted a_{ij}. The system is termed *homogeneous* if all the m right hand sides b_i are equal to 0, otherwise *inhomogeneous*.

|||| Definition 6.6 Solution of System of Linear Equations

By a *solution* to the the system of linear equations

$$\begin{aligned} a_{11} \cdot x_1 + \, a_{12} \cdot x_2 + \ldots + \, a_{1n} \cdot x_n &= b_1 \\ a_{21} \cdot x_1 + \, a_{22} \cdot x_2 + \ldots + \, a_{2n} \cdot x_n &= b_2 \\ &\vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \ldots + a_{mn} \cdot x_n &= b_m \end{aligned} \tag{6-12}$$

we understand an n-tuple $\mathbf{x} = (x_1, x_2, \ldots x_n) \in \mathbb{L}^n$ which by substitution into all of the m linear equations satisfies the equations, i.e. makes the left hand side of each equal to the right hand side.

By the ***general solution*** or just the ***solution set*** we understand the set of all solutions to the system. A single solution is often termed a ***particular*** solution.

|||| Example 6.7 A Homogeneous System of Linear Equations

A homogeneous system of linear equations consisting of two equations with four unknowns is given by:

$$\begin{aligned} x_1 + x_2 + 2\,x_3 + x_4 &= 0 \\ 2\,x_1 - x_2 - x_3 + x_4 &= 0 \end{aligned} \tag{6-13}$$

We investigate whether the two 4-tuples $\mathbf{x} = (1,1,2,-6)$ and $\mathbf{y} = (3,0,1,-5)$ are particular solutions to the equations (6-13). Substituting $\mathbf{x}$ into the left hand side of the system we get

$$\begin{aligned} 1 + 1 + 2 \cdot 2 - 6 &= 0 \\ 2 \cdot 1 - 1 - 2 - 6 &= -7 \end{aligned} \tag{6-14}$$

Because the left hand side is equal to the given right hand side 0 in the first of these equations, $\mathbf{x}$ is only a solution to the first of the two equations. Therefore $\mathbf{x}$ is not a solution to the system.

Substituiting $\mathbf{y}$ we get

$$\begin{aligned} 3 + 0 + 2 \cdot 1 - 5 &= 0 \\ 2 \cdot 3 - 0 - 1 - 5 &= 0 \end{aligned} \tag{6-15}$$

Since in both equations the left hand side is equal to the right hand side 0, $\mathbf{y}$ is a solution to both of the equations. Therefore $\mathbf{y}$ is a particular solution to the system.

The solution set to a system of linear equations is the ***intersection*** of the solution sets for all the equations comprising the system.

6.3 The Coefficient Matrix and the Augmented Matrix

When we investigate a system of linear equations it is often convenient to use *matrices*. A ***matrix*** is a rectangular array consisting of a number of rows and columns. As an example the matrix $\mathbf{M}$ given by

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 5 \\ 8 & 3 & 2 \end{bmatrix}, \tag{6-16}$$

has two rows and three columns. The six elements are termed the ***elements*** of the matrix. The ***diagonal*** of the matrix consists of the elements with equal row and column numbers. In $\mathbf{M}$ the diagonal consists of the elements 1 and 3.

By the *coefficient matrix* $\mathbf{A}$ to the system of linear equations (6-11) we understand the matrix whose first row consists of the coefficients in the first equation, whose second row consists of the coefficients in the second equation, etc. In short, the following matrix with m rows and n columns:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{6-17}$$

The *augmented matrix* $\mathbf{T}$ of the system is constructed by adding a new column to the coefficient matrix consisting of the right hand sides b_i of the system. Thus $\mathbf{T}$ consists of m rows and $n+1$ columns. If we collect the right hand sides b_i into a column vector $\mathbf{b}$, which we denote *the right hand side of the system*, $\mathbf{T}$ is composed as follows, where the vertical line symbolizes the equality sign of the system:

$$\mathbf{T} = \left[\,\mathbf{A} \mid \mathbf{b}\,\right] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \tag{6-18}$$

The vertical line in front of the last column in (6-18) has only the didactical function to create a clear representation of the augmented matrix. One can chose to leave out the line if in a given context this does not lead to misunderstandings.

|||| Example 6.8 Coefficient Matrix, Right Hand Side and Augmented Matrix

In the following system of linear equations with 3 equations and 3 unknowns

$$\begin{aligned} -x_2 + x_3 &= 2 \\ 2x_1 + 4x_2 - 2x_3 &= 2 \\ 3x_1 + 4x_2 + x_3 &= 9 \end{aligned} \tag{6-19}$$

we have

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 \\ 2 & 4 & -2 \\ 3 & 4 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 9 \end{bmatrix} \quad \text{and} \quad \mathbf{T} = \left[\begin{array}{ccc|c} 0 & -1 & 1 & 2 \\ 2 & 4 & -2 & 2 \\ 3 & 4 & 1 & 9 \end{array}\right] \tag{6-20}$$

Notice that the 0 that is placed in the top left position in $\mathbf{A}$ and $\mathbf{T}$, denotes that the coefficient of x_1 in the uppermost row of the system is 0.

The clever thing about a coefficient matrix (and an augmented matrix) is that we do not need to write down the unknowns. The unique position of the coefficients in the matrix means that we are sure of which of the unknowns any single particular coefficient belongs to. Thus we have removed redundant symbols!

6.4 Row Reduction of Systems of Linear Equations

Systems or linear equations can be reduced, that is, made simpler using a method called Gaussian elimination. The method has several versions, and the special variant used in these eNotes goes by the name ***Gauss-Jordan elimination*** . The algebraic basis for all variants is that you can reshape a system of linear equations by so-called ***row operations*** without thereby changing the solution set for the system. When a system of equations is reduced as much as possible it is usually easy to read it and to evaluate the solution set.

|||| Theorem 6.9 Row Operations

The solution set of a system of linear equations is not altered if the system is transformed by any of the following three ***row operations***:

ro_1: Let two of the equations swap rows.

ro_2: Multiply one of the equations by a non-zero constant.

ro_3: To a given equation add one of the other equations multiplied by a constant.

Here we introduce a short notation for each of the three row operations:

ro_1:	$R_i \leftrightarrow R_j$:	The equation in row i is swapped with the equation in row j.
ro_2:	$k \cdot R_i$:	The equation in row i is multiplied by k.
ro_3:	$R_j + k \cdot R_i$:	Add the equation in row i, multiplied by k, to the equation in row j.

In the following example we test the three row operations.

|||| Example 6.10 Row Operations

An example of ro_1: Consider the system of equations below to the left. We swap two equations in the two rows thus performing $R_1 \leftrightarrow R_2$.

$$\begin{aligned} x_1 + 2x_2 &= -3 \\ x_1 + x_2 &= 0 \end{aligned} \quad \rightarrow \quad \begin{aligned} x_1 + x_2 &= 0 \\ x_1 + 2x_2 &= -3 \end{aligned} \tag{6-21}$$

The system to the right has the same solution set as the system on the left.

An example of ro_2: Consider the system of equations below to the left. We multiply the equation in the second row by 5, thus performing $5 \cdot R_2$:

$$\begin{aligned} x_1 + 2x_2 &= -3 \\ x_1 + x_2 &= 0 \end{aligned} \quad \rightarrow \quad \begin{aligned} x_1 + 2x_2 &= -3 \\ 5\,x_1 + 5\,x_2 &= 0 \end{aligned} \tag{6-22}$$

The system to the right has the same solution set as the system on the left.

An example of ro_3: Consider the system of equations below to the left. To the equation in the second row we add the equation in the first row multiplied by 2, thus performing $R_2 + 2 \cdot R_1$:

$$\begin{aligned} x_1 + 2x_2 &= -3 \\ x_1 + x_2 &= 0 \end{aligned} \quad \rightarrow \quad \begin{aligned} x_1 + 2x_2 &= -3 \\ 3x_1 + 5x_2 &= -6 \end{aligned} \tag{6-23}$$

The system to the right has the same solution set as the system on the left.

The arrow, $\rightarrow$, which is used in the three examples indicates that one or more row operations have taken place.

|||| Proof

The first part of the proof of 6.9 is simple: Since the solution set of a system of equations is equal to the *intersection* F of the solution sets for the various equations comprising the system, F is not altered by the order of the equations being changed. Therefore ro_1 is allowed.

Since the solution set of a given equation is not altered when the equation is multiplied by a constant $k \neq 0$, F will not be altered if one of the equations is replaced by the equation multiplied by a constant different from 0. Therefore ro_2 is allowed.

Finally consider a system of linear equations A with n unknowns $\mathbf{x} = (x_1, x_2, \ldots x_n)$. We write the left hand side of an equation in A as $L(\mathbf{x})$ and the right hand side as b . Now we perform an arbitrary row operation of the type ro3 in the following way: An arbitrary equation $L_1(\mathbf{x}) = b_1$ is multiplied by an arbitrary number k and is then added to an arbitrary different equation $L_2(\mathbf{x}) = b_2$. This produces a new equation $L_3(\mathbf{x}) = b_3$ where

$$L_3(\mathbf{x}) = L_2(\mathbf{x}) + k\,L_1(\mathbf{x}) \;\text{ and }\; b_3 = b_2 + k\,b_1 \,.$$

We now show that the system of equations B that emerges as a result of replacing $L_2(\mathbf{x}) = b_2$ in A by $L_3(\mathbf{x}) = b_3$ has the same solution set as A, and that ro_3 thus is allowed. First, assume that $\mathbf{x}_0$ is an arbitrary solution to A . Then it follows from the transformation rules for a linear equation that

$$k\,L_1(\mathbf{x}_0) = k\,b_1$$

and further that

$$L_2(\mathbf{x}_0) + k\,L_1(\mathbf{x}_0) = b_2 + k\,b_1 \,.$$

From this it follows that $L_3(\mathbf{x}_0) = b_3$, and that $\mathbf{x}_0$ is a solution to B. Assume vice versa that $\mathbf{x}_1$ is an arbitrary solution to B . Then it follows that

$$-k\,L_1(\mathbf{x}_1) = -k\,b_1$$

and further that

$$L_3(\mathbf{x}_1) - k\,L_1(\mathbf{x}_1) = b_3 - k\,b_1 \,.$$

This means that $L_2(\mathbf{x}_1) = b_2$, and that $\mathbf{x}_1$ also is a solution to A. In sum we have shown that ro_3 is allowed.

■

From 6.9 follows directly:

|||| Corollary 6.11

The solution set of a system of linear equations is not altered if the system is transformed an arbitrary number of times, in any order, by the three row operations.

We are now ready to use the three row operations for the row reduction of systems of linear equations. In the following example we follow the principles of *Gauss-Jordan elimination*, and a complete description of the method follows in subsection 6.5.

|||| Example 6.12 Gauss-Jordan Elimination

Consider below, to the left, a system of linear equations, consisting of three equations with the three unknowns x_1, x_2 and x_3. On the right the *augmented matrix* for the system is written:

$$\begin{aligned} -x_2 + x_3 &= 2 \\ 2x_1 + 4x_2 - 2x_3 &= 2 \\ 3x_1 + 4x_2 + x_3 &= 9 \end{aligned} \qquad \mathbf{T} = \left[\begin{array}{rrr|r} 0 & -1 & 1 & 2 \\ 2 & 4 & -2 & 2 \\ 3 & 4 & 1 & 9 \end{array}\right] \tag{6-24}$$

The purpose of reduction is to achieve, by means of row operations, the following situation: x_1 is the only remaining part on left hand side of the upper equation , x_2 is the only one on the left hand side of the middle equation and x_3 is the only one on the left hand side of the lower equation. *If* this is possible then the system of equations is not only reduced but also solved! This is achieved in a series of steps taken in accordance with the Gauss-Jordan algorithm. Simultaneously we look at the effect the row operations have on the augmented matrix.

First we aim to have the topmost equation comprise x_1, and to have the coefficient of this x_1 be 1. This can be achieved in two steps. We swap the two top equations and multiply the equation now in the top row by $\frac{1}{2}$. That is,

$$R_1 \leftrightarrow R_2 \quad \text{and} \quad \frac{1}{2} \cdot R_1 \ :$$

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ -x_2 + x_3 &= 2 \\ 3x_1 + 4x_2 + x_3 &= 9 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 2 \\ 3 & 4 & 1 & 9 \end{array}\right] \tag{6-25}$$

Now we remove all other occurrences of x_1. In this example it is only one occurrence, i.e. in row 3. This is achieved as follows: we multiply the equation in row 1 by the number -3 and add the product to the equation in row 3, in short

$$R_3 - 3 \cdot R_1 \ :$$

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ -x_2 + x_3 &= 2 \\ -2x_2 + 4x_3 &= 6 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 2 \\ 0 & -2 & 4 & 6 \end{array}\right] \tag{6-26}$$

We have now achieved that x_1 only appears in row 1 . There it must stay! The work on x_1 is finished. This corresponds to the fact that at the top of the first column of the augmented matrix there is 1 and directly below it only 0's. This means that work on the first column is finished !

The next transformations aim at ensuring that the unknown x_2 will be represented only in row 2 and nowhere else. First we make sure that the coefficient of x_2 in row 2 switches coefficient from -1 to 1 by use of the operation

$$(-1)\cdot R_2\ :$$

$$\begin{aligned} x_1+2x_2-x_3&=1\\ x_2-x_3&=-2\\ -2x_2+4x_3&=6 \end{aligned}\qquad \left[\begin{array}{rrr|r} 1&2&-1&1\\ 0&1&-1&-2\\ 0&-2&4&6 \end{array}\right] \tag{6-27}$$

We now remove the occurrences of x_2 from row 1 and row 3 with the operations

$$R_1-2\cdot R_2\quad\text{and}\quad R_3+2\cdot R_2:$$

$$\begin{aligned} x_1+x_3&=5\\ x_2-x_3&=-2\\ 2x_3&=2 \end{aligned}\qquad \left[\begin{array}{rrr|r} 1&0&1&5\\ 0&1&-1&-2\\ 0&0&2&2 \end{array}\right] \tag{6-28}$$

Now the work with x_2 is finished, which corresponds to the fact that in row 2 in the augmented matrix the number in the second column is 1, all the other numbers in the second column being 0. This column must not be altered by subsequent operations.

Finally we wish that the unknown x_3 is represented in row 3 by the coefficient 1 and that x_3 is removed from row 1 and row 2. This can be accomplished in two steps. First

$$\frac{1}{2}\cdot R_3:$$

$$\begin{aligned} x_1+x_3&=5\\ x_2-x_3&=-2\\ x_3&=1 \end{aligned}\qquad \left[\begin{array}{rrr|r} 1&0&1&5\\ 0&1&-1&-2\\ 0&0&1&1 \end{array}\right] \tag{6-29}$$

Then

$$R_1-R_3\quad\text{and}\quad R_2+R_3:$$

$$\begin{aligned} x_1&=4\\ x_2&=-1\\ x_3&=1 \end{aligned}\qquad \left[\begin{array}{rrr|r} 1&0&0&4\\ 0&1&0&-1\\ 0&0&1&1 \end{array}\right] \tag{6-30}$$

Now x_3 only appears in row 3. This corresponds to the fact that in column 3 in the third row of the augmented matrix we have 1, each of the other elements in the column being 0. We have now completed a *total reduction* of the system, and from this we can conclude that there exists exactly one solution to the system viz :

$$\mathbf{x}=(x_1,x_2,x_3)=(4,-1,1). \tag{6-31}$$

Let us remember what a solution is: an n-tuple that satisfies all the equations in the system! Let us prove that formula (6-31) actually is a solution to equation (6-24):

$$\begin{aligned} -(-1)+1&=2\\ 2\cdot 4+4\cdot(-1)-2\cdot 1&=2\\ 3\cdot 4+4\cdot(-1)+1&=9 \end{aligned}$$

As expected all three equations are satisfied!

In (6-30) after the row operations the augmented matrix of the system of linear equations has achieved a form of special beauty with three so-called leading 1's in the ***diagonal*** and zeros everywhere else. We say that the transformed matrix is in *reduced row echelon form*. It is not always possible to get the simple representation shown in (6-30). Sometimes the leading 1 in the next row is found more than one column to the right, as one moves down. The somewhat complex definition follows below.

|||| Definition 6.13 Reduced Row Echelon Form

A system of linear equations is denoted to be in ***reduced row echelon form***, if the corresponding augmented matrix fulfills the following four conditions:

1. The first number in a row that is not 0, is a 1. This is called the ***leading*** 1 or the ***pivot*** of the row.

2. In two consecutive rows which both contain a pivot, the upper row's leading 1 is further to the left than the leading 1 in the following row.

3. In a column with a leading 1, all other elements are 0.

4. Any rows with only 0's are placed at the bottom of the matrix.

|||| Example 6.14 Reduced Row Echelon Form

Consider the three matrices

$$\mathbf{A} = \begin{bmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{1} & 2 & 0 \\ 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{og} \quad \mathbf{C} = \begin{bmatrix} \mathbf{1} & 3 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{6-32}$$

The three matrices shown are all in row reduced echelon form. In **A** all the leading 1's are nicely placed in the *diagonal*. **B** has only leading two leading 1's and you have to go two steps to the right to go from the first to the second step. In **C** there is only one leading 1.

|||| Example 6.15

None of the following four matrices is in reduced row echelon form because each violates exactly one of the rules in the definition 6.13 – which, is left to reader to figure out!

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{6-33}$$

Note the following important theorem about the relationship between a matrix on the one hand, and the reduced row echelon form of the same matrix produced through the use of row operations, on the other.

|||| Theorem 6.16 Reduced Row Echelon Form

If a given matrix **M** is transformed by two different sequences of row operations into a reduced row echelon form, then the two resulting reduced row echelon forms are identical.

The unique reduced row echelon form a given matrix **M** can be transformed into this way is termed the ***reduced row echelon form***, and given the symbol $\text{rref}(\mathbf{M})$.

|||| Proof

We use the following model for the six matrices that are introduced in the course of the proof:

$$\begin{array}{ccccc} \mathbf{A} & \overset{f_1}{\longleftarrow} & \mathbf{M} & \overset{f_2}{\longrightarrow} & \mathbf{B} \\ & & \downarrow & & \\ \mathbf{A}_1 & \overset{f_1}{\longleftarrow} & \mathbf{M}_1 & \overset{f_2}{\longrightarrow} & \mathbf{B}_1 \end{array} \tag{6-34}$$

Suppose a matrix $\mathbf{M}$ has been transformed, by two different series of row operations f_1 and f_2, into two different reduced row echelon forms $\mathbf{A}$ and $\mathbf{B}$. Let column number k be the first column of $\mathbf{A}$ and $\mathbf{B}$ where the two matrices differ from one another. We form a new matrix $\mathbf{M}_1$ from $\mathbf{M}$ in the following way. First we remove all the columns in $\mathbf{M}$ whose column numbers are larger than k. Then we remove just the columns in $\mathbf{M}$ whose column numbers are less than k , and have the same column numbers as a column in $\mathbf{A}$ (and thus $\mathbf{B}$) which does not contain a leading 1.

Now we transform $\mathbf{M}_1$ by the series of row operations f_1 and f_2 , and the resulting matrices formed hereby are called $\mathbf{A}_1$ and $\mathbf{B}_1$, respectively. Then $\mathbf{A}_1$ necessarily will be the same matrix that would result if we remove all the columns from $\mathbf{A}$, similar to those we took away from $\mathbf{M}$ to produce $\mathbf{M}_1$. And the same relationship exists between $\mathbf{B}_1$ and $\mathbf{B}$. $\mathbf{A}_1$ and $\mathbf{B}_1$ will therefore have a leading 1 in the diagonal of all columns apart from the last, which is the first column where the two matrices are different from one another. In this last column there are two possibilities: Either one of the matrices has a leading 1 in this column or neither of them has. An example of how the situation in the first case could be is:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{B}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \tag{6-35}$$

We now interpret $\mathbf{M}_1$ as the augmented matrix for a system of linear equations $\mathcal{L}$. Both $\mathbf{A}_1$ and $\mathbf{B}_1$ will then represent a totally reduced system of equations with the same solution set as $\mathcal{L}$. However, this leads to a contradiction since one of the totally reduced systems is seen to be inconsistent due to one of the equations now being invalid and the other will have just one solution. We can therefore rule out that one of $\mathbf{A}_1$ and $\mathbf{B}_1$ contains a leading 1 in the last column.

We now investigate the other possibility, that neither of $\mathbf{A}_1$ and $\mathbf{B}_1$ contains a leading 1 in the last column. The situation could then be like this:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{B}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \tag{6-36}$$

Both the totally reduced system of equations as represented by $\mathbf{A}_1$, and that which is represented by $\mathbf{B}_1$, will in this case have exactly one solution. But when the last column is different in the two matrices the solution for $\mathbf{A}_1$'s system of equations will be different from the solution for $\mathbf{B}_1$'s system of equations, whereby we again have ended up in a contradiction.

We conclude that the assumption that **M** might be transformed into two different reduced row echelon forms cannot be true. Hence, to **M** corresponds a unique reduced row echelon form: rref(**M**).

■

From Theorem 6.16 it is relatively easy to obtain the next result about matrices that can transformed into each other through row operations:

|||| Corollary 6.17

If a matrix **M** has been transformed by an arbitrary sequence of row operations into the matrix **N**, then

$$\operatorname{rref}(\mathbf{N}) = \operatorname{rref}(\mathbf{M}). \tag{6-37}$$

|||| Proof

Let s be a sequence of row operations that transforms the matrix **M** to the matrix **N**, and let t be a sequence of row operations that transforms the the matrix **N** to $\operatorname{rref}(\mathbf{N})$. Then the sequence of row operations consisting of s followed by t, transform **M** to $\operatorname{rref}(\mathbf{N})$. But since **M** in accordance with 6.16 has a unique reduced row echelon form, $\operatorname{rref}(\mathbf{M})$ must be equal to $\operatorname{rref}(\mathbf{N})$.

■

If, in the preceding corollary, we interpret **M** and **N** as the augmented matrices for two systems of linear equations, then it follows directly from definition (6.13) that:

|||| Corollary 6.18

If two systems of linear equations can be transformed into one another by the use of row operations, then they are identical in the reduced row echelon form (apart from possible trivial equations).

6.5 Gauss-Jordan Elimination

We are now able to precisely introduce the method of elimination that is applied in these eNotes.

|||| Definition 6.19 Gauss-Jordan Elimination

A system of linear equations is totally reduced by ***Gauss-Jordan elimination*** when the corresponding augmented matrix after the use of the three row operations (see theorem 6.9) is brought into the reduced row echelon form by the following procedure:

> We proceed from left to right : First we treat the first column of the augmented matrix so that it does not conflict with the reduced row echelon form, then the second column is treated so as not to conflict with the reduced row echelon form and so on, as far as and including the last column in the augmented matrix .

This is always possible!

> When you are in the process of reducing systems of linear equations, you are free to deviate from the Gauss-Jordan method if it is convenient in the situation at hand. If you have achieved a reduced row echelon form by using other sequences of row operations, it is the same form that would have been obtained by using the Gauss-Jordan method strictly. This follows from corollary 6.18.

In Example 6.12 it was possible to read the solution from the totally reduced system of linear equations. In the following main example the situation is a bit more complicated owing to the fact that the system has infinitely many solutions.

|||| Example 6.20 Gauss-Jordan Elimination

We want to reduce the following system of four linear equations in five unknowns:

$$\begin{aligned} x_1 + 3x_2 + 2x_3 + 4x_4 + 5x_5 &= 9 \\ 2x_1 + 6x_2 + 4x_3 + 3x_4 + 5x_5 &= 3 \\ 3x_1 + 8x_2 + 6x_3 + 7x_4 + 6x_5 &= 5 \\ 4x_1 + 14x_2 + 8x_3 + 10x_4 + 22x_5 &= 32 \end{aligned} \tag{6-38}$$

We write the augmented matrix for the system:

$$\mathbf{T} = \left[\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 5 & 9 \\ 2 & 6 & 4 & 3 & 5 & 3 \\ 3 & 8 & 6 & 7 & 6 & 5 \\ 4 & 14 & 8 & 10 & 22 & 32 \end{array}\right] \tag{6-39}$$

Below we reduce the system using three row operations. This we will do by only looking at the transformations of the augmented matrix!

$$R_2 - 2 \cdot R_1\,, \quad R_3 - 3 \cdot R_1 \quad \text{and} \quad R_4 - 4 \cdot R_1 :$$

$$\left[\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 5 & 9 \\ 0 & 0 & 0 & -5 & -5 & -15 \\ 0 & -1 & 0 & -5 & -9 & -22 \\ 0 & 2 & 0 & -6 & 2 & -4 \end{array}\right] \tag{6-40}$$

Now we have completed the treatment of the first column, because we have a leading 1 in the first row and only 0's on the other entries in the column.

$$R_2 \leftrightarrow R_3 \quad \text{and} \quad (-1) \cdot R_2 :$$

$$\left[\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 5 & 9 \\ 0 & 1 & 0 & 5 & 9 & 22 \\ 0 & 0 & 0 & -5 & -5 & -15 \\ 0 & 2 & 0 & -6 & 2 & -4 \end{array}\right] \tag{6-41}$$

$$R_1 - 3 \cdot R_2 \quad \text{and} \quad R_4 - 2 \cdot R_2 :$$

$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & -11 & -22 & -57 \\ 0 & 1 & 0 & 5 & 9 & 22 \\ 0 & 0 & 0 & -5 & -5 & -15 \\ 0 & 0 & 0 & -16 & -16 & -48 \end{array}\right] \tag{6-42}$$

The work on the second column is now completed. Now a deviation from the standard situation follows, where leading 1's are established in the diagonal, because it is not possible to produce a leading 1 as the third element in the third row. We are *not* allowed to swap row 1 and row 3, because by doing so the first column would be changed in conflict with the principle that the treatment of the first column is complete. This means that we have also completed the treatment of the third column (the number 2 in the top row cannot be removed). To continue the reduction we move on to the fourth element in row three, where it *is* possible to provide a leading 1.

$$-\tfrac{1}{5}\cdot R_3 :$$
$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & -11 & -22 & -57 \\ 0 & 1 & 0 & 5 & 9 & 22 \\ 0 & 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & -16 & -16 & -48 \end{array}\right] \tag{6-43}$$

$$R_1 + 11\cdot R_3\,, \quad R_2 - 5\cdot R_3 \quad \text{and} \quad R_4 + 16\cdot R_3 :$$
$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -11 & -24 \\ 0 & 1 & 0 & 0 & 4 & 7 \\ 0 & 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{6-44}$$

Now the Gauss-Jordan elimination has ended and we can write the totally reduced system of equations:

$$\begin{aligned} 1x_1 + 0x_2 + 2x_3 + 0x_4 - 11x_5 &= -24 \\ 0x_1 + 1x_2 + 0x_3 + 0x_4 + 4x_5 &= 7 \\ 0x_1 + 0x_2 + 0x_3 + 1x_4 + 1x_5 &= 3 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= 0 \end{aligned} \tag{6-45}$$

First, we note that the original system of equations has actually been reduced (made easier) by the fact that many of the coefficients of the equation system are replaced by 0's. But moreover the system with four equations can now be replaced by a system consisting of only three equations. The last row is indeed a ***trivial*** equation that has the whole $\mathbb{L}^5$ as its solution set. Therefore, the solution set of the system system will not change if the last equation is omitted in the reduced system (since the intersection of the solutions sets of all four equations equals that of the solution sets from the first three equations alone). Quite simply , we can therefore write the totally reduced system of equations as:

$$\begin{aligned} x_1 + 2x_3 - 11x_5 &= -24 \\ x_2 + 4x_5 &= 7 \\ x_4 + x_5 &= 3 \end{aligned} \tag{6-46}$$

But how do we proceed from the reduced system of equations to writing down the solution set in a comprehensible form? We shall return to this example later, see Example 6.30. Before that we need to introduce the concept of *rank*.

6.6 The Concept of Rank

In the example 6.20 a system of linear equations consisting of 4 equations with 5 unknowns has been totally reduced, see equation (6-46). Only three equations are left, because the trivial equation $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 0$ has been left out since it only expresses the fact that $0 = 0$. That the reduced system of equations contains a trivial equation means that the reduced row echelon form of the the augmented matrix contains a 0-row, as in equation (6-44). This leads to the following definition.

|||| Definition 6.21 Rank

By the ***rank*** ρ of a *matrix* we understand the number of rows that are not 0-***rows***, in the reduced row echelon form of the matrix. The rank thereby corresponds to the number of leading 1's in the reduced row echelon form of the matrix.

From the definition 6.21 and corollary 6.18 together with corollary 6.17 we obtain:

|||| Theorem 6.22 Rank and Row Operations

Two matrices that can be transformed into each other by row operations have the same rank.

The rank gives the least possible number of non-trivial equations that a system of equations can be transformed into using row operations. You can never transform a system of linear equations through row operations in such a way that it will contain fewer non-trivial equations than it does when it is totally reduced. This is a consequence of theorem 6.22.

|||| Example 6.23 The Rank of Matrices

A matrix $\mathbf{M}$ with 3 rows and 4 columns is brought into the reduced row echelon form as follows:

$$\mathbf{M} = \begin{bmatrix} 3 & 1 & 7 & -2 \\ -1 & -3 & 3 & 1 \\ 2 & 3 & 0 & -3 \end{bmatrix} \rightarrow \operatorname{rref}(\mathbf{M}) = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6-47}$$

Since $\text{rref}(\mathbf{M})$ does not contain 0-rows, $\rho(\mathbf{M}) = 3$.

A matrix $\mathbf{N}$ with 5 rows and 3 columns is brought into reduced row echelon form like this:

$$\mathbf{N} = \begin{bmatrix} 2 & 2 & 1 \\ -2 & -5 & -4 \\ 3 & 1 & -7 \\ 2 & -1 & -8 \\ 3 & 1 & -7 \end{bmatrix} \quad \rightarrow \quad \text{rref}(\mathbf{N}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{6-48}$$

Since $\text{rref}(\mathbf{N})$ contains three rows that are not 0-rows, $\rho(\mathbf{N}) = 3$.

If we interpret $\mathbf{M}$ and $\mathbf{N}$ as augmented matrices for linear systems of equations we see that for both coefficient matrices the rank is 2, this is less than the ranks of the augmented matrices.

We now investigate the relationship between rank and the number of rows and columns. First we notice that from the definition of 6.21 it follows that the rank of a matrix can never be larger than the number of matrix rows.

In Example 6.23 the rank of $\mathbf{M}$ equals the number of rows in $\mathbf{M}$, while the rank of $\mathbf{N}$ is less than the number of rows in $\mathbf{N}$.

Analogously the rank of a matrix cannot be larger than the number of columns. The rank is in fact equal to the number of leading 1's in the reduced row echelon form . And if the echelon form of the matrix contains more leading 1's than there are columns, then there must be at least one column containing more than one leading 1. But this contradicts condition number 3 in the definition 6.13.

In the example 6.23 the rank of $\mathbf{M}$ is less than the number of columns in $\mathbf{M}$, while the rank of $\mathbf{N}$ equals the number of columns in $\mathbf{N}$.

We summarize the above observations in the following theorem:

|||| Theorem 6.24 Rank, Rows and Columns

For a matrix $\mathbf{M}$ with m rows and n columns we have that

$$\rho(\mathbf{M}) \leq \min\{m, n\}\,. \tag{6-49}$$

6.7 From Reduced Row Echelon Form to the Solution Set

Sometimes it is possible to write down the solution set for a system of linear equations immediately when the corresponding augmented matrix is brought into its reduced echelon form. This applies when the system has no solution or when the system has exactly one solution. If the system has infinitely many solutions, work is needed in order to be able to characterize the solution set. This can be achieved by writing the solution in a standard parametric form. The concept of rank proves well suited to give an instructive overview of the classes of solution sets.

6.7.1 When $\rho(\mathbf{A}) < \rho(\mathbf{T})$

The augmented matrix $\mathbf{T}$ for a system of linear equations has the same number of rows as the coefficient matrix $\mathbf{A}$ but one column more, which contains the right hand sides of the equations. There are two possibilities. Either $\rho(\mathbf{T}) = \rho(\mathbf{A})$, or $\rho(\mathbf{T}) = \rho(\mathbf{A}) + 1$, corresponding to the fact that the last column in $\mathrm{rref}(\mathbf{T})$ contains a leading 1. The consequence of the last possibility is investigated in Example 6.25.

|||| Example 6.25 Inconsistent Equation (No Solution)

The augmented matrix for a system of linear equations consisting of three equations in two unknowns is brought into reduced row echelon form

$$\mathrm{rref}(\mathbf{T}) = \left[\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right] \tag{6-50}$$

The system is thereby reduced to

$$\begin{aligned} x_1 - 2x_2 &= 0 \\ 0x_1 + 0x_2 &= 1 \\ 0x_1 + 0x_2 &= 0 \end{aligned} \tag{6-51}$$

Notice that the equation in the second row is *inconsistent* and thus has no solutions. Because the solution set for the system is the intersection of the solution sets for all the equations, the system has no solutions at all.

Let us look at the reduced row echelon form of the coefficient matrix

$$\mathrm{rref}(\mathbf{A}) = \begin{bmatrix} 1 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{6-52}$$

We note that $\rho(\mathbf{A}) = 1$. This is less than $\rho(\mathbf{T}) = 2$, and this is exactly due to the inconsistency of the equation in the reduced system of equations.

The considerations in example 6.25 can be generalized to the following theorem.

|||| Theorem 6.26 When $\rho(\mathbf{A}) < \rho(\mathbf{T})$

If a system of linear equations with coefficient matrix $\mathbf{A}$ and augmented matrix $\mathbf{T}$ has

$$\rho(\mathbf{A}) < \rho(\mathbf{T})\,, \tag{6-53}$$

then the totally reduced system has an inconsistent equation. Therefore the system has no solutions.

If $\mathrm{rref}(\mathbf{T})$ has a row of the form $\left[\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right]$, then the system has no solutions.

|||| Exercise 6.27

Determine the reduced rwo echelon form of the augmented matrix for the following system of linear equations, and determine the solution set for the system.

$$\begin{aligned} x_1 + x_2 + 2x_3 + x_4 &= 1 \\ -2x_1 - 2x_2 - 4x_3 - 2x_4 &= 3 \end{aligned} \tag{6-54}$$

6.7.2 When $\rho(\mathbf{A}) = \rho(\mathbf{T}) =$ Number of Unknowns

Let n denote the number of unknowns in a given system of linear equations. Then by the way the coefficient matrices are formed there must be n columns in $\mathbf{A}$.

Further we assume that $\rho(\mathbf{A}) = n$. Then $\mathrm{rref}(\mathbf{A})$ contains exactly n leading 1's. Therefore the leading 1's must be placed in the ***diagonal*** in $\mathrm{rref}(\mathbf{A})$, and all other elements of $\mathrm{rref}(\mathbf{A})$ are zero.

Finally we assume that in the given example $\rho(\mathbf{A}) = \rho(\mathbf{T})$. Then the solution set can be read directly from $\text{rref}(\mathbf{T})$. The first row in $\text{rref}(\mathbf{T})$ will correspond to an equation where the first unknown has the coefficient 1 while all the other unknowns have the coefficient 0. Therefore the value of the first unknown is equal to to the last element in the first row (the right hand side). Similarly with the other rows, row number i corresponds to an equation where unknown number i is the only unknown, and therefore its value is equal to the last element in row number i. Since each unknown there corresponds to exactly one value, and since $\rho(\mathbf{A}) = \rho(\mathbf{T})$ we are certain that there is no inconsistent equation in the given system of equations. Thus the given system of equations has exactly one solution.

|||| Example 6.28 Exactly One Solution

The augmented matrix for a system of linear equations consisting of three equations in two unknowns has been brought onto the reduced row echelon form

$$\text{rref}(\mathbf{T}) = \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{array}\right] \tag{6-55}$$

Consider the reduced row echelon form of the coefficient matrix for the system

$$\text{rref}(\mathbf{A}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{6-56}$$

This has a leading 1 in each column and 0 in all other entries. We further note that $\rho(\mathbf{A}) = \rho(\mathbf{T}) = 2$.

From $\text{rref}(\mathbf{T})$ we can write the totally reduced system of equations as

$$\begin{aligned} 1x_1 + 0x_2 &= -3 \\ 0x_1 + 1x_2 &= 5 \\ 0x_1 + 0x_2 &= 0 \end{aligned} \tag{6-57}$$

which shows that this system of equations has exactly one solution $\mathbf{x} = (x_1, x_2) = (-3, 5)$.

The argument given just before the example proves the following theorem:

Theorem 6.29 When $\rho(\mathbf{A}) = \rho(\mathbf{T}) =$ Number of Unknowns

If a linear system with coefficient matrix $\mathbf{A}$ and augmented matrix $\mathbf{T}$ has:

$$\rho(\mathbf{A}) = \rho(\mathbf{T}) = \text{number of unknowns}, \tag{6-58}$$

then the system has exactly one solution, and this can be immediately read from $\mathrm{rref}(\mathbf{T})$.

6.7.3 When $\rho(\mathbf{A}) = \rho(\mathbf{T}) <$ the Number of Unknowns

We are now ready to resume the discussion of our main example 6.20, a system of linear equations with 5 unknowns, for which we found the totally reduced system of equations consisting of 3 non-trivial equations. Let us now find the solution set and investigate its properties!

Example 6.30 Infinitely Many Solutions

In the example 6.20 the augmented matrix $\mathbf{T}$ for a system of linear equations with 4 equations in 5 unknowns was reduced to

$$\mathrm{rref}(\mathbf{T}) = \left[\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -11 & -24 \\ 0 & 1 & 0 & 0 & 4 & 7 \\ 0 & 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{6-59}$$

We see that $\rho(\mathbf{A}) = \rho(\mathbf{T}) = 3$, i.e. less than 5, the number of unknowns.

From $\mathrm{rref}(\mathbf{T})$ we can write the totally reduced system of equations

$$\begin{aligned} x_1 + 2x_3 - 11x_5 &= -24 \\ x_2 + 4x_5 &= 7 \\ x_4 + x_5 &= 3 \end{aligned} \tag{6-60}$$

The system has infinitely many solutions. For every choice of values for x_3 and x_5 we can find exactly one new value for the other unknowns x_1, x_2 and x_4. This can be made more clear by isolating x_1, x_2 and x_4 in the following way

$$\begin{aligned} x_1 &= -24 - 2x_3 + 11x_5 \\ x_2 &= 7 - 4x_5 \\ x_4 &= 3 - x_5 \end{aligned} \tag{6-61}$$

If we, for example, choose $x_3 = 1$ and $x_5 = 2$, we find the solution $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5) = (-4, -1, 1, 1, 2)$. More generally, *any* choice of values for x_3 and x_5 will, in the same way, produce a solution, whilst the other three variables are uniquely determined by the cohice. Therefore we can consider x_3 and x_5 as *free parameters* that determine the value of the three other unknowns, and therefore on the right hand side we rename x_3 and x_5 the parameter names t_1 and t_2, respectively. Then we can write the solution set as:

$$\begin{aligned} x_1 &= -24 - 2t_1 + 11t_2 \\ x_2 &= 7 - 4t_2 \\ x_3 &= t_1 \\ x_4 &= 3 - t_2 \\ x_5 &= t_2 \end{aligned} \tag{6-62}$$

or more clearly in the *standard parameter form*:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -24 \\ 7 \\ 0 \\ 3 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 11 \\ -4 \\ 0 \\ -1 \\ 1 \end{bmatrix} \quad \text{where} \quad t_1, t_2 \in \mathbb{L}. \tag{6-63}$$

With geometry-inspired wording we term the vector $(-24, 7, 0, 3, 0)$ the *initial point* of the solution set and the two vectors $(-2, 0, 1, 0, 0)$ and $(11, -4, 0, -1, 1)$ its *directional vectors*. Letting $\mathbf{x}_0$, $\mathbf{v}_1$ and $\mathbf{v}_2$ denote the initial point, and the directional vectors, respectively, we can write the parametric representation in this way:

$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 \quad \text{hvor} \quad t_1, t_2 \in \mathbb{L}. \tag{6-64}$$

Since the solution set has two free parameters corresponding to two directional vectors, we say that it has a *double -infinity* of solutions.

Line 3 and 5 in (6-63) only express that $x_3 = t_1$ and $x_5 = t_2$.

Let us, inspired by example 6.30, formulate a general method for changing the solution set to standard parametic form from the totally reduced system of equations:

|||| Method 6.31 From the Augmented Matrix to the Solution in Standard Parameter Form

We consider a system of linear equations with n unknowns with the coefficient matrix $\mathbf{A}$ and the augmented matrix $\mathbf{T}$. In addition we assume

$$\rho(\mathbf{A}) = \rho(\mathbf{T}) = k < n. \tag{6-65}$$

The solution set of the system is brought into standard parametric form in this way:

1. We find $\operatorname{rref}(\mathbf{T})$ and from this we write the totally reduced system of equations (as is done in (6-60)).

2. In each of the k non-trivial equations in the totally reduced system of equations we isolate the *first* unknowns on the left hand side (as is done in (6-61)).

3. In this way we have isolated k different unknowns on the left hand side of the total system. The other $(n-k)$ unknowns, that are placed on the right hand side are *renamed* the parameter names $t_1, t_2, \ldots, t_{n-k}$.

4. We can now write the solution set in *standard parametric form*:

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) = \mathbf{x}_0 + t_1\,\mathbf{v}_1 + t_2\,\mathbf{v}_2 + \cdots + t_{n-k}\,\mathbf{v}_{n-k}\,, \tag{6-66}$$

   where the vector $\mathbf{x}_0$ denotes the *initial point* of the parameter representation, while $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-k}$ are its *directional vectors* (as is done in (6-63)).

Notice that the numbers $t_1, t_2, \ldots, t_{n-k}$ can be chosen freely. Regardless of the choice equation (6-66) will be a valid solution. Therefore they are called ***free parameters***.

If the algorithm of the Gauss-Jordan elimination has been followed perfectly, one arrives at a certain initial point and a certain set of directional vectors for the solution set, see equation (6-66). But the solution set can be written with another choice for the initial point (if the system is inhomogeneous), and with a different choice of directional vectors. However, the *number* of directional vectors will always be $(n-k)$.

Solution sets in which some of the unknowns have definite values are possible. In the following example the free parameter only influences one of the unknowns. The other two are locked:

|||| Example 6.32 Infinitely Many Solutions with a Free Parameter

For a given system of linear equations it is found that

$$\operatorname{rref}(\mathbf{T}) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 5 \end{array}\right] \tag{6-67}$$

We see that $\rho(\mathbf{A}) = \rho(\mathbf{T}) = 2 < n = 3$. Accordingly we have one free parameter. We write the solution set as:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{6-68}$$

where t is a scalar that can be chosen freely.

In general you can prove the following theorem:

|||| Theorem 6.33 When $\rho(\mathrm{A}) = \rho(\mathrm{T}) <$ Number of Unknowns

If a system of linear equations with n unknowns and with the coefficient matrix $\mathbf{A}$ and augment matrix $\mathbf{T}$ has

$$\rho(\mathbf{A}) = \rho(\mathbf{T}) = k < n \tag{6-69}$$

Then the system has infinitely many solutions that can be written in standard parameter form with an initial point and $(n-k)$ directional vectors.

6.8 On the Number of Solutions

Let us consider a system of three linear equations in two unknowns:

$$\begin{aligned} a_1 \cdot x + b_1 \cdot y &= c_1 \\ a_2 \cdot x + b_2 \cdot y &= c_2 \\ a_3 \cdot x + b_3 \cdot y &= c_3 \end{aligned} \tag{6-70}$$

We have previously emphasized that the solution set for a system of equations is the ***intersection*** of the solution sets for each of the equations in the system. Let us now interpret the given system of equations as equations for three straight lines in a coordinate system in the plane. Then the solution set corresponds to a set of points that are *common* to all the three lines. In order to answer the question about "number" of solutions we draw the different situations in Figure 6.1. In situation 2 two of the lines are parallel,

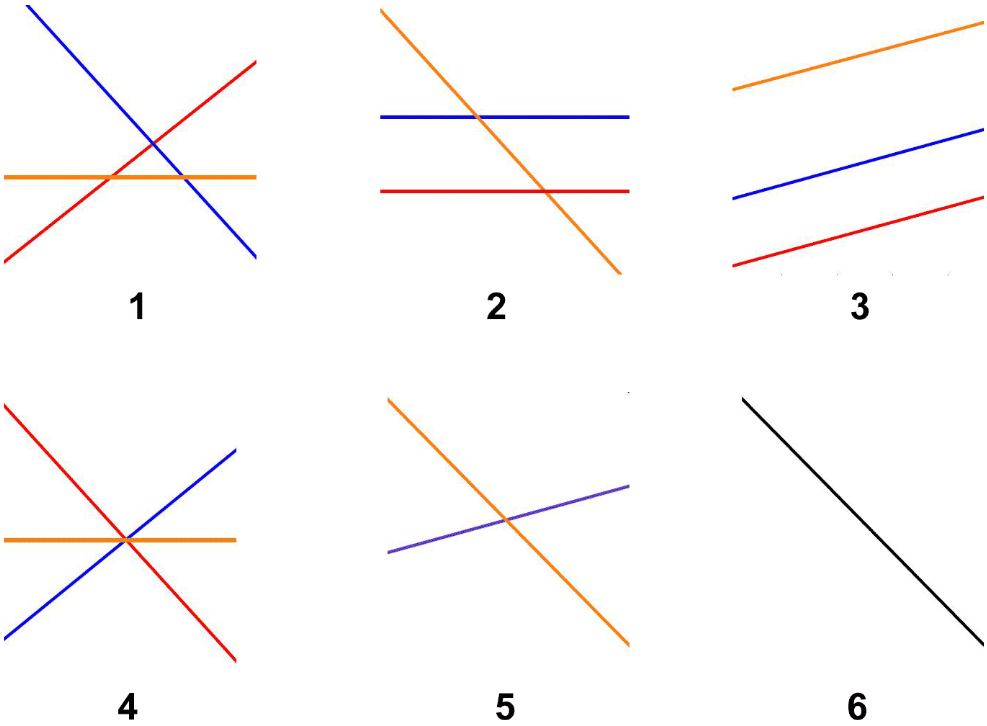

Figure 6.1: The six possible structures of the solutions for three linear equations in two unknowns.

and in situation 3 all three lines are parallel. Therefore there are no points that are part of all three lines in the situations 1, 2 and 3. In situation 5 two of the lines are identical (the blue and the red line coincide in the purple line). Hence there is exactly one common point in the situations 4 and 5. In the situation 6 all the three lines coincide (giving the black line). Therefore in this situation there are infinitely many common points.

The example with three equations in two unknowns illustrates the following theorem which follows from our study of the solution sets in the previous section, see the theorems 6.26, 6.29 and 6.33:

|||| Theorem 6.34 Remark about the Number of Solutions

A system of linear equations either has no, exactly one, or infinitely many solutions. There are no other possibilities.

6.9 The Linear Structure of the Solution set

In this section we dig a little deeper into the question about the *structure* of the solution set for systems of linear equations. It is particularly important to observe the correspondence between the solution set for an inhomogeneous system of equations and the *corresponding homogeneous system of equations*. We start by investigating the homogenous system.

6.9.1 The Properties of Homogeneous Systems of Equations

A homogenous system of linear equations of m linear equations in n unknowns is written in the form

$$\begin{array}{c} a_{11}\cdot x_1 + a_{12}\cdot x_2 + \ldots + a_{1n}\cdot x_n = 0 \\ a_{21}\cdot x_1 + a_{22}\cdot x_2 + \ldots + a_{2n}\cdot x_n = 0 \\ \vdots \\ a_{m1}\cdot x_1 + a_{m2}\cdot x_2 + \ldots + a_{mn}\cdot x_n = 0 \end{array} \tag{6-71}$$

In the following theorem we describe an important property of the structure of the solution set for homogeneous systems of linear equations.

|||| Theorem 6.35 Solutions to a Homogeneous System of Linear Equations

Let L_{hom} denote the solution set of a homogeneous system of linear equations. Then there exists at least one solution to the system, namely the zero or *trivial* solution. If

$$\mathbf{x} = (x_1, x_2, \ldots x_n) \quad \text{and} \quad \mathbf{y} = (y_1, y_2, \ldots y_n) \tag{6-72}$$

are two arbitrary solutions, and k is an arbitrary scalar then both the sum

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots x_n + y_n) \tag{6-73}$$

and the product

$$k\cdot\mathbf{x} = (k\cdot x_1, k\cdot x_2, \ldots k\cdot x_n) \tag{6-74}$$

belong to L_{hom}.

|||| Proof

An obvious property of the system (6-71) is that $\rho(\mathbf{A}) = \rho(\mathbf{T})$ (because the right hand side consists of only zeros). Therefor the system has at least one solution - it follows from theorem 6.29. We can also immediately find a solution, viz. the zero vector, $\mathbf{0} \in \mathbb{L}^n$. That this is a solution is evident when we replace all the unknowns in the system with the number 0, then the system consists of m equations of the form $0 = 0$.

Apart from this the theorem comprises two parts that are proved separately:

1. If

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0 \quad \text{for every} \quad i = 1, 2, \ldots, m \tag{6-75}$$

and

$$a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = 0 \quad \text{for every} \quad i = 1, 2, \ldots, m \tag{6-76}$$

then by addition of the two equations and a following factorization with respect to the coeficients we get

$$a_{i1}(x_1 + y_1) + a_{i2}(x_2 + y_2) + \cdots + a_{in}(x_n + y_n) = 0 \quad \text{for every} \quad i = 1, 2, \ldots, m \tag{6-77}$$

which shows that $\mathbf{x} + \mathbf{y}$ is a solution.

2. If

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0 \quad \text{for every} \quad i = 1, 2, \ldots, m \tag{6-78}$$

and k is an arbitrary scalar, then by multiplying both sides of the equation by k and a following factorization with respect to the coefficients we get

$$a_{i1}(k \cdot x_1) + a_{i2}(k \cdot x_2) + \cdots + a_{in}(k \cdot x_n) = 0 \quad \text{for every} \quad i = 1, 2, \ldots, m \tag{6-79}$$

which shows that $k \cdot \mathbf{x}$ is a solution.

■

|||| Remark 6.36

If you take an arbitrary number of solutions from L_{hom}, multiply these by arbitrary constants and add the products then the so-called ***linear combination*** of solutions also is a solution. This is a consequence of theorem 6.35.

6.9.2 Structural Theorem

We will now consider a decisive relation between an inhomogeneous system of linear equations of the form

$$\begin{aligned} a_{11}\cdot x_1 + \ a_{12}\cdot x_2 + \ldots + \ a_{1n}\cdot x_n &= b_1 \\ a_{21}\cdot x_1 + \ a_{22}\cdot x_2 + \ldots + \ a_{2n}\cdot x_n &= b_2 \\ &\vdots \\ a_{m1}\cdot x_1 + a_{m2}\cdot x_2 + \ldots + a_{mn}\cdot x_n &= b_m \end{aligned} \tag{6-80}$$

and *the corresponding homogeneous system of linear equations*, by which we mean the equations (6-80) after all the right hand sides b_i have been replaced by 0. The solution set for the inhomogeneous system of equations is called L_{inhom} and the solution set for the corresponding homogeneous system of equations is called L_{hom}.

|||| Theorem 6.37 Structural Theorem

If you have found just one solution (a so-called ***particular*** solution) $\mathbf{x}_0$ to an inhomogeneous sytem of linear equations, then L_{inhom} can be found as the sum of $\mathbf{x}_0$ and L_{hom}.

In other words

$$L_{inhom} = \left\{ \mathbf{x} = \mathbf{x}_0 + \mathbf{y} \;\middle|\; \mathbf{y} \in L_{hom} \right\}. \tag{6-81}$$

or in short

$$L_{inhom} = \mathbf{x}_0 + L_{hom}. \tag{6-82}$$

|||| Proof

Note that the theorem contains two propositions. One is that the sum of $\mathbf{x}_0$ and an arbitrary vector from L_{hom} belongs to L_{inhom}. The other is that an arbitrary vector from L_{inhom} can be written as the sum of $\mathbf{x}_0$ and a vector from L_{hom}. We prove the two propositions separately:

1. Assume $\mathbf{y} \in L_{hom}$. We want to show that

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{y} = (x_{0_1} + y_1, x_{0_2} + y_2, \ldots, x_{0_n} + y_n) \in L_{inhom}. \tag{6-83}$$

Since

$$a_{i1}x_{0_1} + a_{i2}x_{0_2} + \cdots + a_{in}x_{0_n} = b_i \quad \text{for any} \quad i = 1, 2, \ldots, m \tag{6-84}$$

and

$$a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = 0 \quad \text{for any} \quad i = 1, 2, \ldots, m \tag{6-85}$$

then by addition of the two equations and a following factorization with respect to the coeficients we get

$$a_{i1}(x_{0_1} + y_1) + \cdots + a_{in}(x_{0_n} + y_n) = b_i \quad \text{for any} \quad i = 1, 2, \ldots, m \tag{6-86}$$

which proves the proposition.

2. Assume $\mathbf{x} \in L_{inhom}$. We want to show that a vector $\mathbf{y} \in L_{hom}$ exists that fulfills

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{y}. \tag{6-87}$$

Since both $\mathbf{x}$ and $\mathbf{x}_0$ belong to L_{inhom} we have that

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \quad \text{for any} \quad i = 1, 2, \ldots, m \tag{6-88}$$

and

$$a_{i1}x_{0_1} + a_{i2}x_{0_2} + \cdots + a_{in}x_{0_n} = b_i \quad \text{for any} \quad i = 1, 2, \ldots, m \tag{6-89}$$

When we subtract the lower equation from the upper, we get after factorization

$$a_{i1}(x_1 - x_{0_1}) + \cdots + a_{in}(x_n - x_{0_n}) = 0 \quad \text{for any} \quad i = 1, 2, \ldots, m \tag{6-90}$$

which shows that the vector $\mathbf{y}$ defined by $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$, belongs to L_{hom} and satisfies: $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$.

■