Systems of Linear Equations Attachment

Systems of linear equations are first introduced at the grade 10 level in the current Ontario curriculum. Generally, just two equations are considered, and solved simultaneously, by one of three methods: by graphing the lines to determine their point of intersection, and algebraically using the substitution method or the elimination method. In the past, we also taught a third algebraic method – the comparison method (which can really be considered to be either the substitution or the elimination method in a slightly modified form).

For example, the system of equations given by, $5x - y = 7$ ① and $3x + y = 9$ ② is solved by the comparison method by solving for y from ① to obtain $y = 5x - 7$ ③ and solving for y in ② to obtain $y = 9 - 3x$ ④ and then equating ③ and ④ to obtain $5x - 7 = 9 - 3x$ and solving for x.

It should be understood that a linear equation with real coefficients can be multiplied by any non-zero scalar and still represent the same relation. As well, a system of linear equations retains the same solution whenever we add or subtract scalar multiples of one or more or the equations from that system. An example of a system that can be used to illustrate these concepts is: $x - y = 2$ ① and $x + 3y = 6$ ② (which has the solution $x = 3$ and $y = 1$). Adding ① and ② produces the new equation $2x + 2y = 8$, while subtracting ① from ② gives us $4y = 4$ while adding 3 × ① and ② yields $4x = 12$. All of these equations have simple integral intercepts and can be graphed using the slope and y-intercept method quite easily as well. This example provides an illustration to support the basis for using the elimination method for solving a system of linear equations simultaneously.

Hopefully, students will question the necessity for having two different methods for finding a solution to the same problem. In future, the equations will be made increasingly complex by either,
a) using more than two variables (in which case the elimination method is required), or
b) using variables with exponents greater than 1 (which requires the substitution method).

In general, a system with m linear equations and n unknowns can be written as:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n &= b_m
\end{aligned}$$

where $x_1, x_2, x_3, \ldots, x_n$ are the unknowns and the numbers $a_{11}, a_{12}, a_{13}, \ldots, a_{mn}$ are the coefficients of the system. We can collect the coefficients in a matrix as follows:

A standard problem is to decide if any assignment of values for the unknowns $x_1, x_2, x_3, \ldots$ can satisfy all three equations simultaneously, and to find such an assignment if it exists. The existence of, and the nature of a solution depends on the equations, and also on the available values (whether integers, real numbers, complex numbers, and so on).

Systems of linear equations have many applications, such as in digital signal processing, estimation, forecasting and generally in linear programming and in the approximation of non-linear problems in numerical analysis. There are many different ways to solve systems of linear equations; However, one of the most efficient ways is given by Gaussian elimination (also called Gauss reduction, row-reduction, row elimination, reduced echelon form, etc.) or by the extended version of this same process, Gauss-Jordan reduction (or row-reduced echelon form, etc.). Incidentally, the term "**echelon form**" is derived from the name given to a military formation in which members (planes, ships, tanks, etc.) are arranged diagonally. The pictures on the next page illustrate such formations.

The use of the term echelon form as it applies to the matrix $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ should be obvious.

We can write the coefficients and variables as a matrix product as follows:

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

If we represent each matrix by a single letter, this becomes $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ where A is an $m \times n$ matrix, $\mathbf{x}$ is a column vector with n entries, and $\mathbf{b}$ is a column vector with m entries. If the field is infinite (as in the case of the real or complex numbers), then only the following three cases are possible (exactly one will be true) for any given system of linear equations:

- the system has no solution (in this case, we say that the system is overdetermined or inconsistent)
- the system has a single solution (the system is exactly determined or consistent and independent)
- the system has infinitely many solutions (the system is underdetermined or consistent and dependent).

If the coefficient matrix and column vector containing the constant terms are combined to form a single matrix such as: $\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$, it is called an augmented matrix.

A system of the form $\boldsymbol{A}\mathbf{x} = \mathbf{0}$ is called a *homogeneous* system of linear equations. The solution of such a homogeneous system is always consistent (i.e. it exists), but may be dependent or independent. The independent solution is also called the **trivial solution** (because it always exists).

Each entry in an $m \times n$ matrix (or element) is defined in terms of its position with respect to the row and column numbers of the matrix. For example, for matrix $\boldsymbol{A}$, the element in the third row and fifth column would be designated as a_{35} (using the same lower case letter as the upper case letter specifying the matrix). A matrix with the same number of rows and columns is called a square matrix. A matrix in which every entry is 0, is called the zero matrix (for example the zero matrix having two rows and four columns would be designated as $\mathbf{0}_{24}$).

The following elementary row operations can be applied to a matrix in order to reduce it (using either Gauss or gauss-Jordan reduction):

- Multiply any row by a non-zero scalar (real number)
- Interchange rows or columns
- Add a scalar multiple of a row to another row (this includes subtraction in that the multiple can be a negative number)

Some texts (particularly high school texts) indicate that you can divide any row by a non-zero number while others restrict this operation to multiplication only. It does not really make a difference in that one can multiply by "$\frac{1}{a}$" instead of dividing by "a".

A matrix is in **reduced row echelon form** (Gauss-Jordan reduced form) if it satisfies the following requirements:

- All nonzero rows are above any rows of all zeroes.
- The leading coefficient of a row is always to the right of the leading coefficient of the row above it.
- All leading coefficients are 1.
- All entries above a leading coefficient in the same column are zero.

A matrix in **reduced echelon form** (Gauss reduction or row echelon form) satisfies only the first three conditions listed above. It is readily seen that these conditions are stronger than those for row echelon form. Therefore, every matrix in reduced row echelon form is also in row echelon form. Unlike row echelon form, every matrix **reduces to a unique matrix** in reduced row echelon form by elementary row operations.

In determining the nature of the solution of a linear system (from its corresponding matrix) we require the definition of the **rank of a matrix**. The **column rank** of a matrix *A* is the maximal number of linearly independent **columns** of *A*. Likewise, the **row rank** is the maximal number of linearly independent **rows** of *A*. In linear algebra, a family of vectors is **linearly independent** if none of them can be written as a linear combination of finitely many other vectors in the collection. A family of vectors which is not linearly independent is called **linearly dependent**. The column rank and the row rank are always equal thus they are simply called the **rank** of *A*. It is commonly denoted by either rk(*A*) or rank *A*.

The rank of a matrix is most easily discerned from a reduced matrix. When a matrix has been reduced (either to Gauss or Gauss-Jordan reduction) then its **rank is equal to the number of non-zero rows**.

From this definition, we can also find the number of parameters (or "free variables") necessary to describe the solution to a dependent system of linear equations. We have: **The number of parameters needed equals the number of variables subtract the rank of the system.**

The matrices shown below are all in **reduced row echelon form**.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 3 & 0 & -8 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 3 & -8 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 5 & -2 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Applying the definition stated above, the first system does not require any parameters (i.e. 3 variables subtract the rank of 3 equals 0), the second requires 1 parameter, the third 2 parameters while the fourth has no solution (and it is meaningless to talk of the number of parameters here).

Consider the system of equations: $\begin{array}{l} 3x - y + 2z = 7 \\ x + 4y + z = 12 \\ 4x + 2z = 10 \end{array}$. The augmented matrix which corresponds to this system is $\left[\begin{array}{ccc|c} 3 & -1 & 2 & 7 \\ 1 & 4 & 1 & 12 \\ 4 & 0 & 2 & 10 \end{array}\right]$, in which the coefficient matrix is given by $\begin{bmatrix} 3 & -1 & 2 \\ 1 & 4 & 1 \\ 4 & 0 & 2 \end{bmatrix}$ and the column vector (representing the constant terms) is given by $\begin{bmatrix} 7 \\ 12 \\ 10 \end{bmatrix}$.

The matrix will be reduced to **reduced row echelon form** using the following steps:

$$\left[\begin{array}{ccc|c} 3 & -1 & 2 & 7 \\ 1 & 4 & 1 & 12 \\ 4 & 0 & 2 & 10 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 3 & -1 & 2 & 7 \\ 4 & 0 & 2 & 10 \end{array}\right] \quad \text{(interchange row 1 and row 2)}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 0 & 13 & 1 & 29 \\ 0 & 16 & 2 & 38 \end{array}\right] \quad \begin{array}{l}\text{(multiply row 1 by 3 and subtract row 2;} \\ \text{multiply row 1 by 4 and subtract row 3)}\end{array}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 0 & 13 & 1 & 29 \\ 0 & 8 & 1 & 19 \end{array}\right] \quad \text{(multiply row 3 by } \tfrac{1}{2}\text{)}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 0 & 104 & 8 & 232 \\ 0 & 104 & 13 & 247 \end{array}\right] \quad \text{(multiply row 2 by 8 and multiply row 3 by 13)}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 0 & 104 & 8 & 232 \\ 0 & 0 & -5 & -15 \end{array}\right] \quad \text{(subtract row 3 from row 2)}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 1 & 12 \\ 0 & 13 & 1 & 29 \\ 0 & 0 & 1 & 3 \end{array}\right] \quad \text{(multiply row 2 by } \tfrac{1}{8} \text{ and row 3 by } -\tfrac{1}{5}\text{)}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 0 & 9 \\ 0 & 13 & 0 & 26 \\ 0 & 0 & 1 & 3 \end{array}\right] \quad \text{(subtract row 3 from row 2 and subtract row 3 from row 1)}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 0 & 9 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \text{(multiply row 2 by } \tfrac{1}{13}\text{)}$$

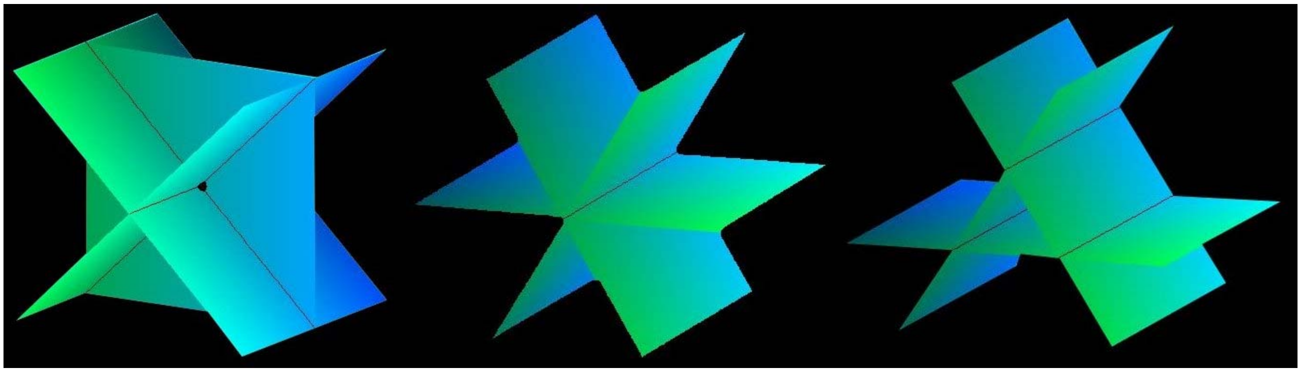
$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \qquad \text{(row 1 subtract 4 times row 2)}$$

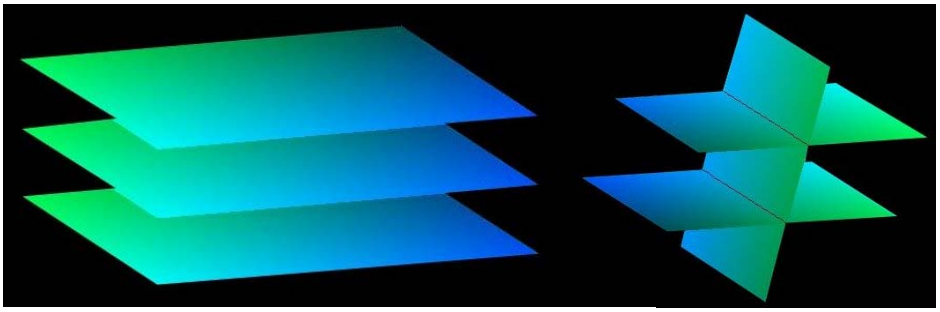
The matrix is now reduced to reduced row echelon form. The solution is $x = 1$, $y = 2$ and $z = 3$ (and is consistent and independent). The symbol "~" is used to denote that the subsequent matrices are equivalent (represent the same equivalent equations with the same solution). The symbol "⇒" is often used instead of the symbol "~".

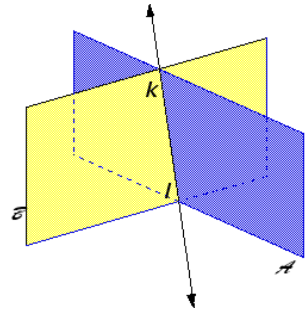
In general (particularly for a system of linear equations that is consistent and independent) the strategy involved is to obtain zero elements in the positions in the matrix illustrated here, in the order **1** first, **2** second, **3** third and **4** fourth. $\left[\begin{array}{ccc|c} a & 4 & 3 & d \\ 1 & b & 3 & e \\ 1 & 2 & c & f \end{array}\right]$ It is most helpful one begins the process by obtaining a "1" in the cell occupied by element "*a*" in the example shown. This can be accomplished by interchanging rows or columns, but one must exercise care in interchanging columns because the order of variables also change. Students tend to have the greatest difficulty reducing matrices that involve dependent systems (and will require parameters in the solution). This is always the case for systems in which there are more variables than there are equations, but can also arise when a zero row is obtained during the process of reduction. Occasionally students incorrectly obtain a zero row by performing the same row operation twice (e.g. row 2 subtract row 3 and row 3 subtract row 2). Often, students attempt to obtain zero elements in a position where they will revert back to non-zero elements later on in the reduction process (e.g. by attempting to get a "0" in the position occupied by element **2.**

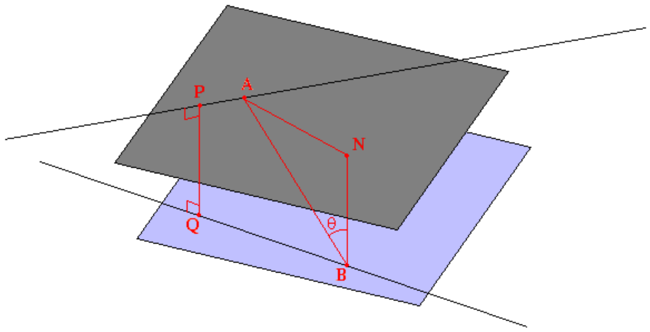
The solution of systems involving dependent systems will be examined in greater detail during the presentation given at the session on Saturday. Many other methods for solving linear systems are considered in discrete mathematics courses at university – particularly those involving determinants.

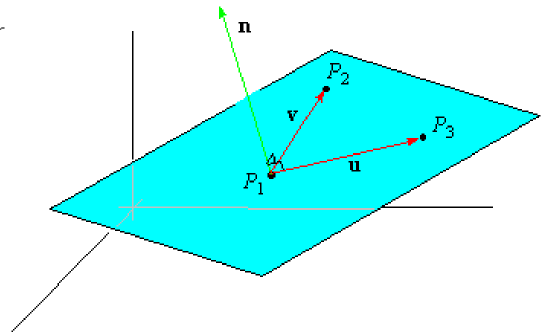
Often, secondary school students are required to illustrate the geometric interpretation of the systems of linear equations (in particular, those involving three equations in three variables). There are actually eight possible cases that can arise, as illustrated by the diagrams below.











The last three diagrams are not considered if we are dealing with equations which represent distinct planes. These cases will be explained more fully in the presentation Saturday.

The examples given in question 1.) below illustrate extensions of the elimination method to special non-linear systems which often appear in contest questions.

1.) Solve the systems of equations given by:

a) $\begin{aligned} x^2 + 3y^2 &= 31 \\ 5x^2 - 2y^2 &= 2 \end{aligned}$

b) $\begin{aligned} \frac{8}{x} + \frac{12}{y} &= 6 \\ \frac{4}{x} - \frac{9}{y} &= -2 \end{aligned}$

c) $\begin{aligned} 4\left(3^x\right) - 2^y &= 28 \\ 3^x + 3\left(2^y\right) &= 33 \end{aligned}$

d) $\begin{aligned} 5\sin x - 3\cos y &= 1.1 \\ 2\sin x + 7\cos y &= 7 \end{aligned}$

2.) Solve the system of equations given by:

a) $\begin{aligned} x + 3y + 2z &= 10 \\ 2x - y + 3z &= 8 \\ 3x + 2y + 5z &= 18 \end{aligned}$

b) $\begin{aligned} x + 3y + 2z &= 10 \\ 2x - y + 3z &= 8 \\ 3x + 2y + 5z &= 19 \end{aligned}$

c) $\begin{aligned} x_1 + x_2 + 3x_3 + x_4 - x_5 &= 10 \\ x_1 + 2x_2 + x_4 &= 12 \\ x_3 + 2x_4 + x_5 &= 16 \end{aligned}$

d) $\begin{aligned} 3x + y - 2z &= 0 \\ 2x + 2y - 5z &= 0 \\ 5x + 3y + 2z &= 0 \end{aligned}$

3.) A parabola passes through the points (1, 0), (2, 2) and (3, 10). Determine the equation of the parabola. (This is a good question for combining a system of linear equations with the quadratic function.)