Level Three Systems of Linear Equations

Alex Elzenaar

January 15, 2018

Contents

Preface

To the instructor

These notes are terse and computational — this is mainly because the underlying theory (elementary linear algebra) is not a part of the standard L3 curriculum. If you want pretty proofs, these are not the notes for you! I have proved a few of the easier theorems on simultaneous linear equations, but of course without the proper linear algebra notation the actual inspiration behind the proofs is impenetrable.

A majority of the exercises I have included are standard and often adapted from other texts (e.g. Whipkey, Whipkey, & Conway's *The Power of Mathematics*); I have not included a large number of computational problems as almost all are very similar in style and are easily found either online or in textbooks. A quick search in the University library catalogue reveals a vast number of books on linear programming, generally located around [519.72].

To the student

These notes are in much more detail than is actually required for even an excellence at Level 3. Essentially, examiners are looking for your ability to solve simple computational problems and to show geometric and/or contextual understanding. While this material is actually very applicable to industrial applications (e.g. in economics or biology), it is very dry as we do not have time to go off on a tangent to actually discuss the ideas behind it. As such, these are perhaps the most boring topics presented at this level; if you are actually interested in the theory and the mathematics behind the techniques, I recommend picking up a book on linear algebra instead.

Section 1: Introduction

Anna, an American, empties her purse and finds that it contains only nickels (worth 5 cents each) and dimes (worth 10 cents each). If she has a total of 7 coins and they have a combined value of 45 cents, how many of each coin does she have?

Situations like this often come up when solving problems — we have two variables (the total amount of each coin), and two linear equations that they satisfy:

$$\begin{cases} n + d = 7 \\ 5n + 10d = 45. \end{cases} \tag{1.1}$$

Our goal in this standard is to come up with a procedure for deciding if there are any solutions of such a linear system of equations, and for finding those solutions if they do exist.

We will begin by formalising the terms which we will be using throughout this standard.

1.2 Definition (Linear equation)**.** A **linear equation** in n variables x_1, x_2, ..., x_n is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

for suitable constants a_1, a_2, ..., a_n, and b.

For example, $2x + 3y = 3$ is a linear equation; $37x^2 + 23y + 4 = 2$ is not (because the variable x is squared); $\sin x = 3$ is not (because $\sin x$ is not in the given form, and cannot be put into the given form).

1.3 Definition (Linear system of equations)**.** A **linear system of equations** is a set of linear equations in n variables x_1, x_2, ..., x_n. A **solution** of the linear system is a set of values for each variable such that every equation is satisfied.

For example, the system

$$\begin{cases} x + y = 2 \\ 2x + 2y + z = 4 \end{cases} \tag{1.4}$$

has the solution $(x, y, z) = (1, 1, 0)$.

Section 2: Graphing Linear Systems

Since each equation is linear, we can easily graph them. We will begin by looking at systems of two variables only, since we can graph these on a plane.

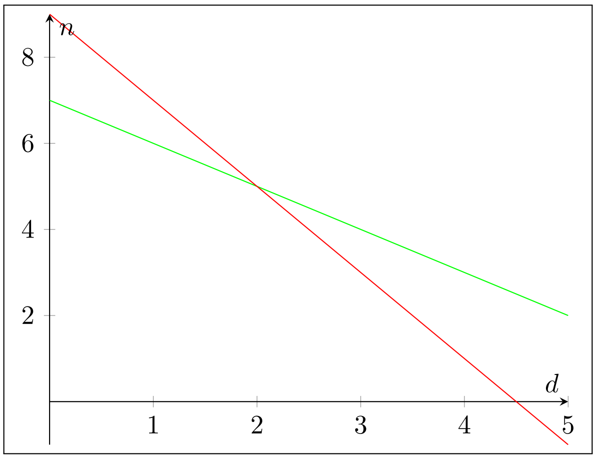
2.1 Example. Consider Anna's system of equations from above,

$$\begin{cases} n + d = 7 \\ 5n + 10d = 45. \end{cases} \tag{2.2}$$

We can rearrange these to be easier to solve:

$$\begin{cases} n = 7 - d \\ n = 9 - 2d, \end{cases} \tag{2.3}$$

and we can graph these:



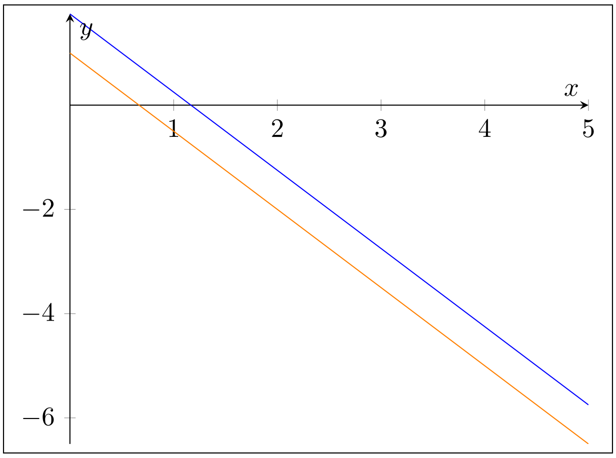
It follows that there is a unique solution to this system: the point of intersection of the two lines, namely $(d, n) = (2, 5)$.

2.4 Example. Now, consider the following system:

$$\begin{cases} 3x + 2y = 2 \\ 6x + 4y = 7. \end{cases} \tag{2.5}$$

Rearranging to make y the subject, we have

$$y = \frac{1}{2}(2 - 3x)$$
$$y = \frac{1}{4}(7 - 6x).$$

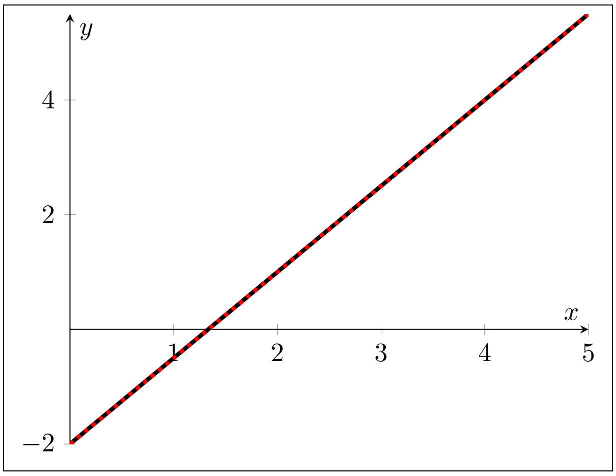


From this graph, we can see that the two lines are parallel and hence the system has no solution. We can also prove this algebraically: note that the second equation in 2.7 implies that $3x + 2y = 7/2$, which contradicts the first.

2.6 Example. For a third example, let us ponder

$$\begin{cases} 3x - 2y = 4 \\ 4y - 6x = -8. \end{cases} \tag{2.7}$$

Rearranging and graphing:



In this example, the two equations conincide everywhere and so there are an infinite number of solutions; you should verify that $(2t, 4 - 3t)$ is a solution for every possible value of t.

These examples suggest that (for two variables, anyway) there are three possibilities:

- A system can have no solutions (it can be **impossible**).
- A system can have exactly one solution.
- A system can have infinitely many solutions.

We will prove this in the next section.

Exercises

1. For the following situations, translate the situations into systems of simultaneous equations. *Do not solve the resulting systems!*

   (a) Goldie has twice as many red blocks as blue blocks. The total number of blocks that she has is three times the number of blue blocks that she has. How many of each type of block does she have?

   (b) Adam breeds chickens and ducks. Last month, he sold 50 chickens and 30 ducks for \$550. This month, he sold 44 chickens and 36 ducks for \$532. How much does he sell each bird for?

   (c) Emma is comparing telecom companies. Phoneworld has an initial *monthly* cost of \$30, and then charges 5 cents per minute of calling. PhoneCity has an initial *yearly* cost of \$200, and charges 20 cents per minute of calling. How many minutes per year must Emma be on the telephone for both plans to cost her exactly the same?

   (d) Jack rides his bicycle at 15 kilometres per hour; Rowan rides his bicycle at 17 kilometres per hour. They are initially 18 kilometres apart when they start to ride directly towards each other. How far from Jack's starting point do they meet?

2. Graph the following systems of equations. How many solutions does each system have?

   (a) $\begin{cases} 2x + 3y = 2 \\ 7x - 14y = 7. \end{cases}$ (b) $\begin{cases} x + y = 2 \\ x + y = 4. \end{cases}$ (c) $\begin{cases} x - 3y = 2 \\ 2x - 6y = 4. \end{cases}$

3. Consider the system

$$\begin{cases} x + y = 2 \\ 4x + 4y = n. \end{cases}$$

Is it possible to pick any value of n such that this system has a unique solution?

4. Use a computerised graphing system (such as GeoGebra) to graph the following systems in three variables. Make a conjecture about the possible number of solutions for a system of equations with three variables.

(a) $\begin{cases} -x + y + z = 1 \\ x - y + z = 1 \\ x + y - z = 1. \end{cases}$ (b) $\begin{cases} -x + y + z = 1 \\ x - y - z = 1 \\ -x + y - z = 1. \end{cases}$ (c) $\begin{cases} 2x + y - z = 4 \\ 4x + 2y - 2z = 5 \\ 8x - 4y - 4z = 16. \end{cases}$

Section 3: The Main Theorem

Never underestimate the power of a theorem that counts something. — Unknown.

The goal in this section is to prove the following theorem that we observed in the last section:

3.1 Theorem (Main Theorem on Linear Equalities)**.**
Given any system of linear equations, we have three possibilities:

- *The system is inconsistent (that is, it has no solutions).*
- *The system is consistent and has a single unique solution.*
- *The system has infinitely many solutions.*

The proof is quite technical, however, so readers who wish to skip it may do so. We begin by proving the theorem for the two smallest cases (systems of a single variable, and systems of two variables) in the following two lemmata — they are not actually necessary for our main proof, but they are much easier to understand.

3.2 Lemma. *Consider any system of a single variable. Then we have exactly two possibilities:*

- *The system is inconsistent (that is, it has no solutions).*
- *The system is consistent and has a single unique solution.*

Proof. A system of a single variable will look like

$$\begin{cases} a_1 x = b_1 \\ a_2 x = b_2 \\ \qquad \vdots \\ a_m x = b_x. \end{cases} \tag{3.3}$$

It is clear that it is possible for such a system to have either no solutions (e.g. take $a_1 = 0$ and $b_1 = 1$) or exactly one solution. Now, suppose that the system has two solutions x and x'. Then, in particular, we have $a_1 x = b_1$ and $a_1 x' = b_1$; so $x = x'$. □

3.4 Lemma. *Consider any system of two variables. Then we have exactly three possibilities:*

- *The system is inconsistent (that is, it has no solutions).*
- *The system is consistent and has a single unique solution.*
- *The system has infinitely many solutions.*

Proof. It is clear that it is possible for such a system to have either no solutions, exactly one solution, or at least two solutions. Now, suppose that the system has two distinct solutions (x, y) and (x', y'). Let t be any number, and suppose that $a_i x + b_i y = c_i$ is any equation in the system. Then $\left(t, \frac{y-y'}{x-x'} t + \frac{c_i}{b_i}\right)$ is a solution:

$$\begin{aligned} a_i t + b_i \left(\frac{y - y'}{x - x'} t + \frac{c_i}{b_i} \right) &= a_i t + t \frac{b_i y - b_i y'}{x - x'} + c_i \\ &= a_i t + t \frac{(c_i - a_i x) - (c_i - a_i x')}{x - x'} + c_i \\ &= a_i t + t \frac{c_i - c_i + a_i x' - a_i x}{x - x'} + c_i \\ &= a_i t + t \frac{a_i (x' - x)}{x - x'} + c_i \\ &= a_i t - a_i t + c_i = c_i. \end{aligned}$$

Hence, if we have at least two distinct solutions, then we have an infinite number of solutions (one for each t) — and we are done. □

Now for the killing blow. The proof we give here will probably seem to be pulled out of thin air (i.e. there seems to be no motivation — how did I come up with it?). The reason for this is that the standard proof of this uses a branch of mathematics known as *linear algebra*. I have translated the proof from the notation of that area of study into the kind of arithmetical algebra that you see below.

Essentially the idea is the same as for the $n = 2$ case in the lemma above; we note that $(x_1 - x'_1, ..., x_n - x'_n)$ is almost a solution, in the sense that it makes all the constant terms zero (this set of values is in the **nullspace** of our system), and then we add on one of our original solutions to get the correct constant terms.

Proof of the main theorem. Let us consider a system of n variables:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn-1}x_{n-1} + a_{mn}x_n = b_m \end{cases} \tag{3.5}$$

Suppose that we have two distinct solutions, $(x_1, ..., x_n)$ and $(x'_1, ..., x'_n)$; let t be a number, and consider $(x_1 + t(x_1 - x'_1), x_2 + t(x_2 - x'_2), ..., x_n + t(x_n - x'_n))$. Then we see that this is a solution to the ith equation in the system:

$$\begin{aligned} &a_{i1}(x_1 + t(x_1 - x'_1)) + \cdots + a_{in}(x_n + t(x_n - x'_n)) \\ &\quad= a_{i1}x_1 + a_{i1}t(x_1 - x'_1) + \cdots + a_{in}x_n + a_{in}t(x_n - x'_n) \\ &\quad= (a_{i1}x_1 + \cdots + a_{in}x_n) + t(a_{i1}(x_1 - x'_1) + \cdots + a_{in}(x_n - x'_n)) \\ &\quad= b_i + t[(a_{i1}x_1 + \cdots + a_{in}x_n) - (a_{i1}x'_1 + \cdots + a_{in}x'_n)] \\ &\quad= b_i + t[b_i - b_i] = b_i. \end{aligned}$$

Since this works for all the solutions in the system, we have an infinite number of solutions: one for each value of t. As above, it is obvious that examples of all three possibilities in the statement of the theorem above exist; by showing that if we have more than one distinct solution we must by necessity have an infinite number of solutions, we have proved that those three possibilities are the only possibilities. [1] □

3.6 Example. In three dimensions, the main theorem has a nice geometric interpretation like our two dimensional discussion in the last section; this time, however, we are talking about planes rather than lines. See figure 1 for an illustration of all three scenarios.

Exercises

1. Consider the following system of two equations.

$$\begin{cases} -x + y = 3 \\ 4x - 4y = 12. \end{cases}$$

   Write down all possible solutions.

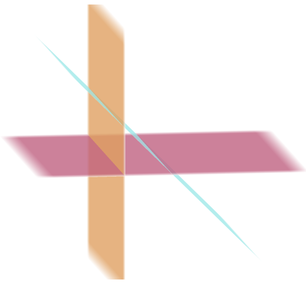
2. Consider the system of equations

$$\begin{cases} -x + y + 2z + 4t = 17 \\ 2x + 0y + z - 7t = 23. \end{cases}$$

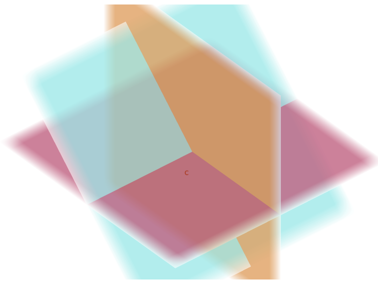
   Given that $(1, -42, 28, 1)$ and $(2, -55, 33, 2)$ are solutions to this system, find infinitely many values (x, y, z, t) satisfying

$$\begin{cases} -x + y + 2z + 4t = 0 \\ 2x + 0y + z - 7t = 0. \end{cases}$$

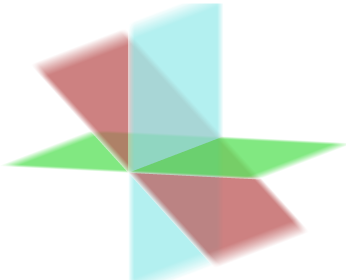
[1] The main idea in this proof is to take an individual solution and to add to it something in the nullspace of the coefficient matrix; if A is the coefficient matrix and B is the constant vector, we have $Ax = B$ and $Ax' = B$, so $t(x - x') \in \operatorname{null} A$ for all t, and hence $A(x + t(x - x')) = Ax = B$ — so $x + t(x - x')$ is a solution vector for every t.



(a) The system is inconsistent.



(b) The system has a unique solution.



(c) The system has an infinite family of solutions.

Figure 1: All three cases of the Main Theorem in 3D.

3. Consider a system of two equations in two variables,

$$\begin{cases} ax + by = c \\ a'x + b'y = c'. \end{cases}$$

Suppose further that this system has no solution. How must a, b, and c be related to a', b', and c'?

Section 4: Algorithms for Solving Systems

Just do it. — Shia LaBeouf.

We have now seen that there are only a small number of possibilities for the number of solutions; we will now look at two ways to actually compute these solutions, if they exist.

4.1 Substitution

4.1 Theorem. *If we have a system of two variables,*

$$\begin{cases} ax + by = c \\ a'x + b'y = c', \end{cases} \tag{4.2}$$

and it has a unique solution then the solution is

$$(x, y) = \left(\frac{b'c - bc'}{ab' - a'b}, \frac{c' - a'x}{b'} \right). \tag{4.3}$$

It is not the result itself (which can be checked by direct substitution) that is important for us here; it is the method that we use to obtain that result. The idea is to rearrange one of the equations to make a variable the subject, and then substitute that equation into the other. This eliminates the main difficulty with linear systems: the fact that there are multiple variables.

Proof of theorem 4.1. The proof is purely algebraic. First, rearrange the second equation to obtain

$$y = \frac{c' - a'x}{b'};$$

substitute it into the first, and then solve for x:

$$\begin{aligned} ax + b\frac{c' - a'x}{b'} &= c \\ \left(a - \frac{b}{b'}a'\right) x &= c - \frac{b}{b'}c' \\ x &= \frac{c - \frac{b}{b'}c'}{a - \frac{b}{b'}a'} = \frac{b'c - bc'}{ab' - a'b}. \end{aligned}$$

□

This method of substitution is quite useful for systems of two variables, but for larger systems it quickly becomes unwieldy: we must substitute the third into the first two in order to reduce them to two variables each, and then substitute the second into the first to obtain a system of one variable which is solvable by inspection. We will work through a simple example, and then one a little more complicated.

4.4 Example. Consider the following system:

$$\begin{cases} -x + y + z = 1 \\ x - y + z = 1 \\ x + y - z = 1. \end{cases} \tag{4.5}$$

We first solve the final equation for z, obtaining $z = x + y - 1$. Substituting this into the first two, we have the following system of two variables:

$$\begin{cases} 2y = 2 \\ 2x = 2, \end{cases} \tag{4.6}$$

which is easily solvable by inspection; so the unique solution is $(x, y, z) = (1, 1, 1)$.

4.7 Example. Suppose we wish to find the equation of the parabola that passes through the points $(-1, 9)$, $(1, 5)$, and $(2, 12)$. We set up our system of linear equations as follows:

$$\begin{cases} a(-1)^2 + b(-1) + c = 9 \\ a(1)^2 + b(1) + c = 5 \\ a(2)^2 + b(2) + c = 12; \end{cases}$$

simplifying, we have

$$\begin{cases} a - b + c = 9 \\ a + b + c = 5 \\ 4a + 2b + c = 12. \end{cases} \tag{4.8}$$

We will begin by substituting out b. The first equation tells us that $b = a + c - 9$, so we can substitute into the second two equations:

$$\begin{cases} a + (a + c - 9) + c = 5 \\ 4a + 2(a + c - 9) + c = 12, \end{cases}$$

or

$$\begin{cases} a + c = 7 \\ 2a + c = 10. \end{cases}$$

Since $a = 7 - c$, we have $10 = 2(7 - c) + c = 14 - c$. Hence $c = 4$, $a = 3$, and $b = -2$; the equation of the parabola is simply

$$3x^2 - 2x + 4 = 0.$$

4.2 Row reduction

For larger systems, a more systematic approach is needed.

4.9 Definition (Elementary row operations)**.** The following operations on equations in a system are called **elementary**:

1. Swapping two equations.
2. Multiplying an equation by a (non-zero) number.
3. Adding one equation to another.

The reason that these operations in particular are useful is the following lemma.

4.10 Lemma. *If we perform an elementary row operation on a system of linear equations, then the resulting system has exactly the same solutions as the original system.*

Proof. Operation (1) obviously does not change the solution set. Let the ith equation in the system be

$$a_{i1}x_1 + \cdots + a_{in}x_n = b_i;$$

if we multiply the whole equation by some non-zero number λ on both sides, then the result has the same equations; hence (2) cannot change the solutions.

Further, let the jth equation be

$$a_{j_1}x_1 + \cdots + a_{j_n}x_n = b_j$$

so that the sum of the two equations is

$$(a_{i1} + a_{j_1})x_1 + \cdots + (a_{in} + a_{j_n})x_n = b_i + b_j.$$

It should be obvious that if we have a set of values for $x_1, ..., x_n$ that satisfies both equations then this resulting equation should be satisfied by that same value (and note also that no new solutions can be introduced). Hence (3) cannot change the solutions. $\square$

Our goal is to use these elementary operations to reduce any system to one which can be easily solved by inspection.

4.11 Example. We will use row reduction to solve the following system:

$$\begin{cases} 2x + 2y + 2z = 2 \\ 2x - y - z = 2 \\ y + z = 0 \end{cases} \tag{4.12}$$

First, we will subtract the second equation from the first (more formally, multiply the second equation by -1 and add it to the first):

$$\begin{cases} 3y + 3z = 0 \\ 2x - y - z = 2 \\ y + z = 0 \end{cases}$$

Subtracting the final equation from the first, three times in a row:

$$\begin{cases} 0 = 0 \\ 2x - y - z = 2 \\ y + z = 0 \end{cases}$$

And adding the final equation to the second:

$$\begin{cases} 0 = 0 \\ 2x = 2 \\ y + z = 0 \end{cases} \tag{4.13}$$

From this, we can read out immediately that $x = 1$; since we have no condition on y, we can set it to be any number t and find that $z = -t$. We therefore have an infinite number of solutions, one for each value of t: $(x, y, z) = (1, t, -t)$.s

An important observation is that, in a linear system, the names of the variables do not affect the solutions: we can get away without writing the variables down, and just write down the coefficients of the equations in an array (called an **augmented matrix**):

4.14 Example. We will solve the same system as in the previous example:

$$\begin{cases} 2x + 2y + 2z = 2 \\ 2x - y - z = 2 \\ y + z = 0 \end{cases} \tag{4.15}$$

Writing this as a matrix, we have

$$\left[\begin{array}{rrr|r} 2 & 2 & 2 & 2 \\ 2 & -1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{array}\right] \tag{4.16}$$

Then, applying the same elementary row operations as above (where $R_1 = R_2 + R_3$ means "replace the first row with the sum of the first and the second rows") we can reduce this matrix:

$$\left[\begin{array}{rrr|r} 2 & 2 & 2 & 2 \\ 2 & -1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_1 = R_1 - R_2} \left[\begin{array}{rrr|r} 0 & 3 & 3 & 0 \\ 2 & -1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_1 = R_1 - 3R_3} \left[\begin{array}{rrr|r} 0 & 0 & 0 & 0 \\ 2 & -1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

$$\left[\begin{array}{rrr|r} 0 & 0 & 0 & 0 \\ 2 & -1 & -1 & 2 \\ 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_2 = R_2 + R_3} \left[\begin{array}{rrr|r} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_2 = (1/2) R_2} \left[\begin{array}{rrr|r} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

So $x = 1$, and $y = -z$ (as above).

We are almost at a stage where we can write down a general algorithm for solving systems of linear equations; we need a technical definition first:

4.17 Definition (Row echelon form)**.** A set of linear equations is in row echelon form if the first variable in each row is to the right of the leading variable in the row above. If a variable is the leftmost variable in some row then we call it a **pivot**; otherwise we call it a **free variable**.

An example of a free variable is the variable z in our matrix example above; we are free to pick any value t for it that we like.

4.18 Example. The following system is in row echelon form:

$$\begin{cases} x + y + 2z + w = 0 \\ +3y - w = 4 \\ +w = 2 \end{cases}$$

Here x, y, and w are pivots; z is a free variable.

It is easy to see that all systems in row echelon form can be solved easily by substituting the lower ones into the higher ones. We will now write down a general algorithm that can be used to solve all consistent systems of linear equations.

4.19 Algorithm (Gaussian Elimination)**.** To solve a system of linear equations:

1. Write the augmented matrix of the system.
2. Use elementary row operations to convert the matrix to row echelon form, by trying to eliminate the leftmost entry of the bottom row of the matrix, and moving up to the next row when it becomes impossible to do so.
3. Use back substitution to solve the system of linear equations corresponding to the row-reduced matrix.

Exercises

1. Using theorem 4.1, give a criteria for a system of two equations in two variables to not have a unique solution (i.e. is there an expression in the coefficients whose value can be used to determine if such a system has exactly one solution).
2. Using substitution, give an analogue of theorem 4.1 for a system of three variables.
3. For each of the following systems, describe *all* the solutions or prove that the system is inconsistent. Can you give a graphical interpretation for each?

(a) $\begin{cases} x + 4y + 3z = 0 \\ 2x + 10y + 18z = 0. \end{cases}$ (b) $\begin{cases} 3x + 2y + z = 1 \\ 2x - 5y + 2z = 0 \\ x + 7y - z = 1 \\ x + y + z = 15. \end{cases}$ (c) $\begin{cases} x + y = 1 \\ 2x + 2y = 4. \end{cases}$ (d) $\begin{cases} -a + b + c + d = 12 \\ a - b + c + d = 12 \\ a + b - c + d = 12 \\ a + b + c - d = 12. \end{cases}$

4. Write down the corresponding system of equations (in the variables x_1, x_2, ..., x_n) for each of the following augmented matrices.

(a) $\left[\begin{array}{rr|r} 2 & 3 & 1 \\ -7 & 2 & 2 \end{array}\right]$ (b) $\left[\begin{array}{rrr|r} 1 & -1 & 1 & -1 \\ 2 & -2 & -2 & -3 \end{array}\right]$ (c) $\left[\begin{array}{rr|l} 1 & 2 & -3 \\ 4 & -5 & 6 \\ -7 & 8 & 9 \end{array}\right]$

5. Goldie has \$8.10 in her pocket, and wants to buy 13 pieces of chocolate. There are three kinds of chocolate available: nutritious, magical, and whimsical. Per piece, the price of each is (respectively) \$0.50, \$0.60, and \$0.90. Goldie wants to buy twice as many nutritious chocolate pieces as whimsical chocolate pieces, and wants to use up all her money so she doesn't have to pay tax to the magical pixie government. To her surprise, she finds that there is only one possible combination of purchases she can make. How many pieces of each type can she buy?

Section 5: Determining with Determinants

In the previous section, we solved all consistent systems of equations. However, we still cannot tell whether any given system has zero, one, or infinite solutions without either solving it or graphing it. In this section, we will develop a method for classifying systems into two categories: those with precisely one solution, and all the others. This does not actually sound like a particularly useful result: “Surely”, (I hear you shout), “it would be more useful to be able to tell if any given system has *no* solutions!” Well, too bad — this is the result that we're going to be discussing.

As with our discussion of the main theorem on linear equalities, we will start with a specific case: the “two variables, two equations” case. You should have guessed this lemma already, based on some of the exercises from the previous sections.

5.1 Lemma. *The system of equations*

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases}$$

has a unique solution if and only if $ad - bc \neq 0$*.*

Proof. Suppose that the system has a unique solution. From theorem 4.1, we know that the unique solution must have x value

$$x = \frac{de - bf}{ad - bc},$$

which is clearly undefined when $ad - bc = 0$. On the other hand, if $ad - bc = 0$ then the system of equations cannot have a unique solution. □

The following little lemma is trivial; it deals with so-called **homogenous** systems.

5.2 Lemma. *If we have a system of m equations in n variables, where the constant term of each equation is zero (as in the following system), and the system has at least one non-zero solution, then the system has an infinite number of solutions.*

Proof. Consider the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n-1}x_{n-1} + a_{2n}x_n = 0 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn-1}x_{n-1} + a_{mn}x_n = 0 \end{cases} \tag{5.3}$$

and suppose that $(x_1, ..., x_n)$ is a solution. Then $(kx_1, ..., kx_n)$ is a solution for all k. □

Obviously $(0, ..., 0)$ will be a solution for any such system.

5.4 Example. Consider the system

$$\begin{cases} 3x + 4y + z = 0x + y + z = 0. \end{cases} \tag{5.5}$$

One non-zero solution to this is $(-3, 2, 1)$; an infinite number of solutions is given by $(-3k, 2k, k)$ for constants k.

Let us now make a set of arbitrary definitions.

5.6 Definition. Consider a system of n equations in n variables (i.e. a system in which the number of equations is the same as the number of variables).

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn-1}x_{n-1} + a_{nn}x_n = b_n \end{cases} \tag{5.7}$$

1. The **determinant** of this $n \times n$ system is a number, denoted by

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & a_{nn} \end{vmatrix}, \tag{5.8}$$

that is defined recursively by the following items.

2. An **element** of the determinant is a single component (e.g. a_{11}) of the determinant.
3. The determinant of a 1×1 system,

$$\begin{vmatrix} a \end{vmatrix} \tag{5.9}$$

is defined to be a.

4. The **minor** of an element of the determinant is the determinant remaining when the row and column containing the element are deleted. For example, given

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}, \tag{5.10}$$

the minor of a is $\begin{vmatrix} e & f \\ h & i \end{vmatrix}$.

5. The **cofactor** of an element of the determinant is the minor of that element multiplied by $+1$ or -1 depending on its position in the matrix:

$$\begin{vmatrix} +1 & -1 & +1 & \cdots \\ -1 & +1 & -1 & \cdots \\ +1 & -1 & +1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}. \tag{5.11}$$

6. The determinant of an $n \times n$ system, where $n \geq 2$, can be calculated by multiplying all the entries in any row or column by their cofactors, and summing the result.

Determinants, as you can see, are horrible things (especially when we meet them here, outside of their natural habitat). The main purpose of introducing them is the following theorem.

5.12 Theorem (Cramer)**.** *Consider the generic system*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn-1}x_{n-1} + a_{nn}x_n = b_n, \end{cases}$$

and let the determinant of this system be $\mathcal{D}$. If this system has a unique solution, then that solution is given by $(x_1, ..., x_n)$ where

$$x_i = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1i-1} & b_1 & a_{1i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i-1} & b_2 & a_{2i+1} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{ni-1} & b_n & a_{ni+1} & \cdots & a_{nn} \end{vmatrix}}{\mathcal{D}}. \tag{5.13}$$

Proof. We omit the proof, because it requires development of properties of determinants that are, while elementary, tedious and not worth the time to spell out algebraically. ☐

It follows directly that

5.14 Corollary. *A linear system that has the same number of equations as variables has a unique solution exactly when its determinant is nonzero.*

After all that, let us consider a couple of special cases.

5.15 Theorem. *The determinant of a* 2×2 *system*

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases}$$

is given by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{5.16}$$

This was previously implied by lemma 5.1.

Proof.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \begin{vmatrix} d \end{vmatrix} - b \begin{vmatrix} c \end{vmatrix} = ad - bc. \tag{5.17}$$

□

5.18 Theorem. *The following analagous equality holds in the* 3×3 *case.*

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh. \tag{5.19}$$

Proof. This is a straighforward calculation; if the proof were written here then you would not read it, and (besides) you will gain far more by doing the computation yourself. □

Exercises

1. Compute: (a) $\begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix}$ (b) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \\ 1 & 1 & 1 \end{vmatrix}$ (c) $\begin{vmatrix} 0 & 0 & 0 \\ 1 & 2 & 17 \\ 3 & 2 & 3 \end{vmatrix}$ (d) $\begin{vmatrix} 1 & 2 & -1 & 3 \\ 2 & -3 & 5 & 1 \\ -2 & 4 & 1 & -4 \\ 3 & 4 & -2 & 8 \end{vmatrix}$

2. Prove theorem 5.18

3. Use the theorem of Cramer to solve the linear system

$$\begin{cases} x + 2y + 3z = 4 \\ 5x + 6y + 7z = 8 \\ 9x + 10y + 11z = 12. \end{cases}$$

4. Explain why, in the proof of lemma 5.1, we ignore the case that $d = 0$ causes the solution given in the earlier theorem to be undefined. [Hint: if $d = 0$, but $ad - bc \neq 0$, can you solve the system of equations explicitly?]

5. Consider lemma 5.2.

   (a) Exhibit a homogenous linear system with two non-zero solutions that are not multiples of each other.

   (b) (Difficult) Show that *all* non-zero solutions to a homogenous system must be linear combinations of d specific solutions that are not multiples of each other, where d is the number of free variables in the row-reduced augmented matrix.

Section 6: Systems of Inequalities

We now know some things about systems of linear equations. Can we also know some things about systems of linear inequalities?

In this section and the next, we will only consider systems of linear equations in two variables.

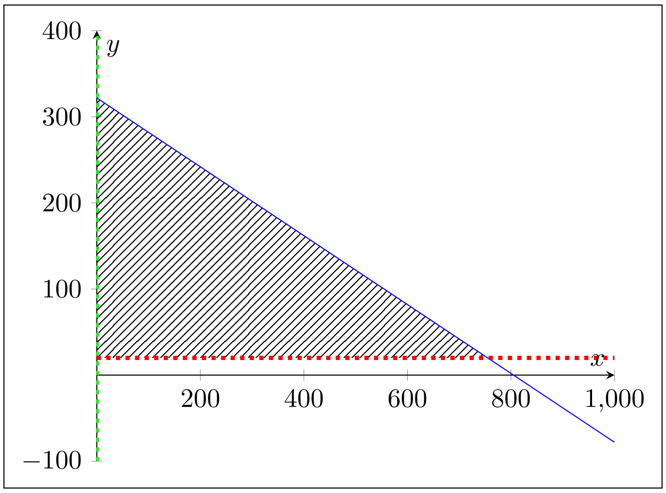
6.1 Example.

> *Arthur is selling bracelets and stuffed animals in order to fund a holiday. Return flights from WLG to PVG cost \$1610 in total. He makes a profit of \$2 per bracelet, and of \$5 per stuffed animal. He needs to sell more than 20 stuffed animals.*

The associated system of linear inequalities is

$$\begin{cases} 2x + 5y \geq 1610 \\ x > 0 \\ y > 20. \end{cases} \tag{6.2}$$

We can graph these constraints, as we did with linear equations.



In diagrams, we use dotted lines for strict inequalities ($<$ and $>$), and solid lines for normal inequalities ($\geq$ and $\leq$). The the set of all points (x, y) that satisfy every inequality in the system is called the **feasible region**; we have shaded it in the above diagram.

We should also the following definition and theorem.

6.3 Definition. We write that $x \geq y$ if there exists some non-negative real number a such that $x = y + a$. If $a \neq 0$, then we write $x > y$.

6.4 Theorem. *Suppose that $x \geq y$.*

1. *If $y \geq z$, then $x \geq z$.*
2. *If r is any real number, then $x + r \geq y + r$.*
3. *If λ is any* positive *real number, then $\lambda x \geq \lambda y$.*
4. *If λ is any* negative *real number, then $\lambda y \geq \lambda x$.*

Proof. One at a time:

1. We have $x = y + a$ and $y = z + b$ (for non-negative a and b), so $x = z + (a + b)$ and $x \geq z$.
2. We have $x = y + a$ so $x + r = (y + r) + a$ and $x + r \geq y + r$.
3. We have $x = y + a$ so $\lambda x = \lambda y + (\lambda a)$. Since both λ and a are nonnegative, their product is nonnegative; hence $\lambda x \geq \lambda y$.

4. We have $x = y + a$ so $\lambda x = \lambda y + (\lambda a)$, where λa is nonpositive; hence $-\lambda a$ is nonnegative, and $\lambda x + (-\lambda a) = \lambda y$ implies that $\lambda y \geq \lambda x$.

□

There is not much else to say about systems of linear inequalities with as little structure as those we have seen in this section; we can find the 'corners' of the system by solving the corresponding system of linear equations, we can check if a point is within the feasible region, and that's about it.

Final note: for each system of linear inequalities, we have a corresponding system of linear equalities obtained by replacing each inequality sign with an equality sign.

Exercises

1. Write down the linear constraints corresponding to the following situations, and graph the feasible regions.

   (a) Hayden is taking a mathematics exam and needs to complete at least ten geometry problems and algebra problems within three hours. It will take him thirty minutes to complete a geometry problem and ten minutes to complete an algebra problem.

   (b) Fuel from petrol pump A costs \$2.19 per litre and from petrol pump B costs \$1.69 per litre. Richard Seddon has at most \$20 to spend on fuel.

   (c) John is packing apples and oranges into boxes. Each box can hold either 15 apples or 8 oranges. He needs to pack at least 40 boxes and at least 360 pieces of fruit.

2. During a family trip, André Weil shared the driving with his father. At most, he was allowed to drive for three hours in total. While driving, your maximum speed is eighty kilometres per hour. Write a system of inequalities describing the possible number of hours t and distance d that he may have driven; is it possible that he drove for 160 kilometres?

Section 7: Linear Programming

We can introduce some additional structure to a linear inequalities problem by adding an additional constraint: suppose we want to find the point within the feasible region that maximises or minimises some other quantity.

7.1 Example. A company produces two models of sewing machine. Each unit of model A requires four hours of work on the first assembly line, and five on the second; each unit of model B requires seven hours of work on the first line, but only three on the second. During a week, the first line is allowed to spend up to 160 man-hours on sewing machines, and the second line is allowed to spend up to 80 man-hours on sewing machines. The total profit resulting from the sale of a single unit of model A is \$60; the total profit resulting from the sale of a unit of model B is \$50.

In this case the quantity to be maximised is the profit, given by the function

$$\theta(a, b) = 60a + 50b, \tag{7.2}$$

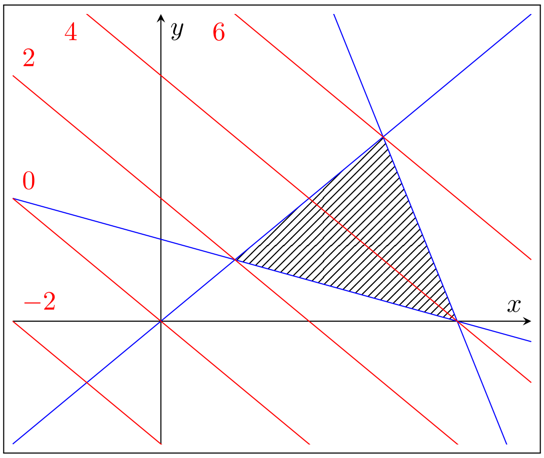
where a is the total number of model A units produced, and b is the total number of model B units produced. We call a constraint function such as this an **objective function**.

Our main result is the following theorem.

7.3 Theorem (Main Theorem on Linear Inequalities, two variable case)**.** *Consider any system of linear inequalities in two variables x and y and any linear objective function $\theta(x, y)$. Then the maximum and minimum values of θ occur at the 'corners' of the linear system; that is, the intersection points of the individual linear functions of the system.*

We will not give a proof of this here, as it is a little technical and our goal is to apply the theorem rather than to justify it.[2] However, we can look at a simple example to convince ourselves that it is reasonable.

7.4 Example. Consider the following system; it does not matter precisely what the actual linear inequations are. Suppose that we wish to maximise the value C given by some linear equation $ax+by = C$. Let us graph this situation for different values of C:



Since an increase in C means either in increase or a decrease in the y-intercept of $C = ax + by$ (depending on the sign of a and b) but no change of slope of the function, it follows that the highest value of C such that $ax + by = C$ still passes through the feasible region causes the function to pass through one of the corners of the feasible region.

7.5 Example. Consider the linear programming scenario pictured in figure 2. Suppose we have further been given an objective function $f(x, y) = 3x - 4y$. By the main theorem on linear programming, the maximum and minimum values of f must occur at corners. The maximum value, for example, occurs at $(12, 0)$.

Since this topic is perhaps the most boring of all the L3 mathematics topics, we will not dwell further on it.

[2] A proof can be found in most applied linear algebra textbooks, or one can be constructed with multivariate calculus.

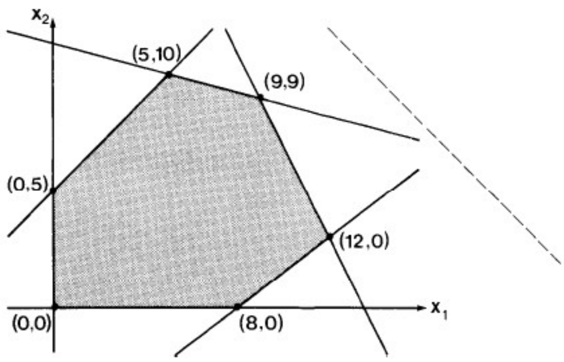


Figure 2: Yet another feasible region.

Exercises

1. A manufacturer produces two different models, X and Y, of the same product. There are two different raw materials, m and n, required for each model. At least 18 units of raw material m and 12 units of raw material n must be used per day. There are at most 34 hours of labour allowed each day. For model X, 2 units of m and 1 of n is required; for model Y, 1 unit of each is required. It requires 3 hours to manufacture one model X unit, and 2 hours to manufacture one model Y unit. The profit is \$3 for each model X, and \$5 for each model Y. How many units of each model should be produced each day to maximise the profit?

2. Before opening a new factory, the following information (based on a single hour of operation) is considered.

	No. of Type I Produced	**No. of Type II Produced**	**Sulfur Oxides Released (kg)**
Prod. Line A	10	20	3
Prod. Line B	10	30	5

   The factory must produce at least 50 units of type I and at least 120 of type II per hour. How many of each type of production line should be installed to minimise the amount of sulfur oxides released?

3. A company produces boxes of pens and of pencils. There is an expected demand of at least 100 boxes of pens and 80 boxes of pencils each day. No more than 200 boxes of pens and 170 boxes of pencils can be made daily, but a total of at least 200 boxes must be shipped each day. If each box of pens sold results in a \$1 loss, but each box of pencils produces a \$2 profit, how many of each type should be made daily?

Section 8: Linear Inequalities in n Variables

Our results also apply naturally to higher-dimensional cases.

8.1 Example (Scholarship 2016)**.** A student is to sit an examination. The questions are divided into three groups. The student may answer any question from any group so long as the total number of questions answered does not exceed 100. The groups are characterised as follows:

- Group 1 — easy, worth four marks each, and will take an average time of two minutes per question to answer.
- Group 2 — moderate difficulty, worth five marks each, and will take an average time of three minutes per question to answer.
- Group 3 — the most difficult, worth six marks each, and will take an average time of four minutes per question to answer.

The total time available to the student is 3.5 hours. The questions in groups 1 and 2 are the most mechanical and the student can tolerate only 2.5 hours of this kind of work before losing motivation.

What combination of questions should the student answer for a maximum grade, assuming all answered questions are correct?

Instead of having a feasible region that is geometrically a plane, we now have a feasible region that is a volume. Our constraints are (in three variables each):

$$2x_1 + 3x_2 + 4x_3 \leq 210, \tag{L1}$$
$$2x_1 + 3x_2 \leq 150, \tag{L2}$$
$$x_1 + x_2 + x_3 \leq 100, \tag{L3}$$
$$x_1 \geq 0, \tag{L4}$$
$$x_2 \geq 0, \text{ and} \tag{L5}$$
$$x_3 \geq 0. \tag{L6}$$

Each represents a plane in 3-space.

Our objective function is

$$G = 4x_1 + 5x_2 + 6x_3. \tag{8.2}$$

Since we are in three dimensions, a 'corner' is now the intersection point of three (or more) of our constraint surfaces. Since we have 6 constraint surfaces, there are at most $\binom{6}{3} = 20$ corners.

If we do this, we find 18 corners (combinations) — there are two missing, since the combinations of L1/L2/L6, and L2/L4/L5 are inconsistent (the three planes do not meet in either case). By drawing a suitable diagram of the situation (figure 3), it is possible to minimise the amount of further work required by throwing out all the other corners that miss out on being within the feasible region.

Table 1 shows the final results. Hence, for the highest grade (of 390) the student should answer 75 questions from group 1, none from group 2, and 15 from group 3.

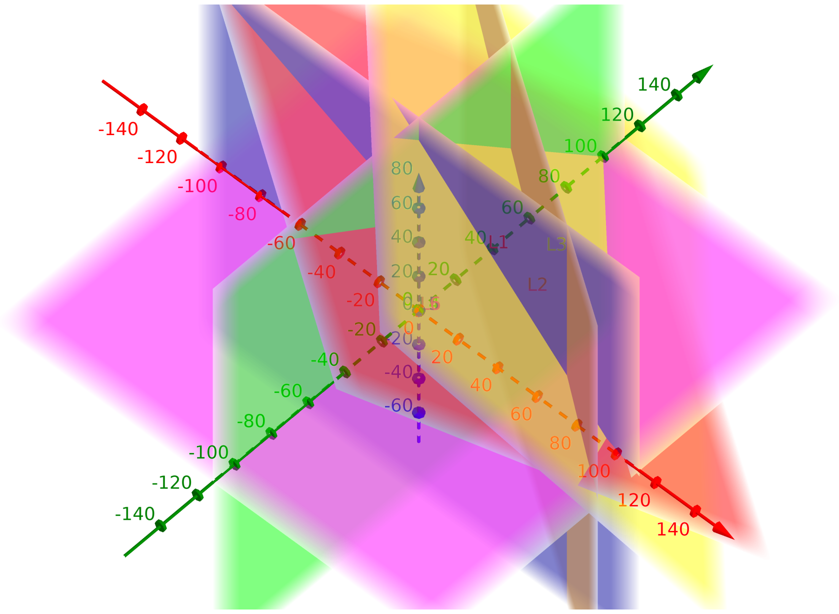


Figure 3: The inequations L1-6 correspond to red, orange, yellow, green, blue, and pink respectively.

Planes	**Corner**	**Feasible?**	**Objective function**
L1/L2/L3	$(105, -20, 15)$	No	N/A
L1/L2/L4	$(0, 50, 15)$	Yes	340
L1/L2/L5	$(75, 0, 15)$	Yes	390
L1/L2/L6	inconsistent	No	N/A
L1/L3/L4	$(0, 190, -90)$	No	N/A
L1/L3/L5	$(95, 0, 5)$	No	N/A
L1/L3/L6	$(90, 10, 0)$	No	N/A
L1/L4/L5	$(0, 0, 52.5)$	Yes	315
L1/L4/L6	$(0, 70, 0)$	No	N/A
L1/L5/L6	$(105, 0, 0)$	No	N/A
L2/L3/L4	$(0, 50, 50)$	No	N/A
L2/L3/L5	$(75, 0, 25)$	No	N/A
L2/L3/L6	$(150, -50, 0)$	No	N/A
L2/L4/L5	inconsistent	No	N/A
L2/L4/L6	$(0, 50, 0)$	Yes	250
L2/L5/L6	$(75, 0, 0)$	Yes	300
L3/L4/L5	$(0, 0, 100)$	No	N/A
L3/L4/L6	$(0, 100, 0)$	No	N/A
L3/L5/L6	$(100, 0, 0)$	No	N/A
L4/L5/L6	$(0, 0, 0)$	Yes	0

Table 1

Section 9: More Exercises

1. For each of the following, describe the system graphically and solve if possible.

$$\text{(a)} \begin{cases} 2x+3y=6 \\ 9x+2y=19 \\ 4x+6y=14. \end{cases} \qquad \text{(b)} \begin{cases} x+2y+z=6 \\ 2x-y+z=8 \\ -2x+y+z=4. \end{cases} \qquad \text{(c)} \begin{cases} x+y+z=3 \\ x+y+z=4 \\ x-y+z=3 \\ -x+y-z=4 \end{cases}$$

2. Find the row-reduced echelon form of each of the following matrices, and write down the corresponding systems of equations.

$$\text{(a)} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 16 \end{array}\right] \qquad \text{(b)} \left[\begin{array}{cc|c} 3 & 1 & 4 \\ 1 & 5 & 9 \\ 2 & 6 & 5 \end{array}\right] \qquad \text{(c)} \left[\begin{array}{cc|c} \sqrt{2} & \sqrt{3} & \sqrt{5} \\ 2 & 3 & 5 \\ 2^2 & 3^2 & 5^2 \end{array}\right]$$

3. Compute $\begin{vmatrix} 2 & 3 \\ 3 & 2 \end{vmatrix}$ and $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$.

4. Agatha and Arthur are buying cooking supplies. In total, they wish to buy 3 kilograms of flour and sugar. They have \$24 available; flour costs \$6 per kilogram, and sugar costs \$9 per kilogram. How much flour and sugar do they buy?

5. Solve the following system.

$$\begin{cases} 9^{2x+y} - 9^x \times 3^y = 6 \\ \log_{x+1}(y+3) + \log_{x+1}(y+x+4) = 3 \end{cases}$$

6. Explain why we cannot subtract true linear inequalities from each other to obtain another true linear inequality:

$$\begin{aligned} 1 &\leq 2 \text{ and} \\ 4 &\leq 8 \text{ but} \\ -3 &\not\leq -6. \end{aligned}$$

(This explains why we cannot simply solve linear programming problems by row reduction.)

7. For each of the following, draw the feasible region.

$$\text{(a)} \begin{cases} 2x+3y \leq 6 \\ 9x+2y \geq 19 \\ 4x+6y < 14. \end{cases} \qquad \text{(b)} \begin{cases} x < y \\ 2x > 3 \\ y < 4 \end{cases} \qquad \text{(c)} \begin{cases} x+y+1 < 0 \\ 2-x > 6 \\ (x+1)(y+2) > 0 \end{cases}$$

8. Come up with an example of a 2-variable linear programming problem involving a *quadratic* objective function such that the objective function is maximised at a point inside the feasible region.

9. A publisher has orders for 600 copies of a book from Wellington and 400 copies from Christchurch. The company has 700 copies in a warehouse in Dunedin and 800 copies in a warehouse in Auckland. It costs \$5 to ship a text from Dunedin to Christchurch, but it costs \$10 to ship it to Wellington. It costs \$15 to ship a text from Auckland to Christchurch, but it costs \$4 to ship it from Auckland to Wellington. How many copies should the company ship from each warehouse to Wellington and Christchurch to fill the order at the least cost?

10. A company produces three items, A, B, and C. The company has three factories, each of which produces the three items in the quantities per hour indicated in the following table:

		Plant		
		I	II	III
Item	A	1	2	3
	B	2	1	4
	C	3	1	1

It costs \$1000 per hour to operate plant I, \$400 per hour to operate plant II, and \$2400 per hour to operate plant III.

An order is placed with the company for three units of item A, five of item B, and six of item C. Determine the number of hours each plant should be operated to produce at least the required number of items for the order at minimum cost.