

Solving simple equations, a review/primer

© 2010 Yonatan Katznelson

Differential and integral calculus do not replace algebra, they build on it. When working problems in calculus, you will find yourself faced with the usual algebraic tasks of simplifying some expressions, expanding others, and of course, solving equations.

There is no unified set of formulas for solving general algebraic equations. On the other hand there are simple methods and formulas for the equations that you will encounter most frequently: linear and quadratic equations, or systems of these equations.

1. *Equations in one variable*

A solution of an equation in one variable is a *number* that satisfies the equation. In general, equations can have no solutions, one solution, two solutions or more. In this course we will only be considering *real numbers*, and so when we speak of *the* solution(s) of an equation we will only be considering solutions that are real numbers. For example, the equation

$$(1.1) \quad x^2 - 2x + 2 = 0$$

has two solutions given by $x_1 = 1 - \sqrt{-1}$ and $x_2 = 1 + \sqrt{-1}$. These solutions are *complex* numbers because they involve $\sqrt{-1}$, and while they are perfectly good numbers, they are not real numbers.[†]

If we consider only real values of the variable x then

$$x^2 - 2x + 2 = (x - 1)^2 + 1 \geq 1$$

for any real x , since the square of the *real* number is always nonnegative. This implies that Equation (1.1) has no real solutions.

1.1 *Linear equations*

The simplest equation that you will encounter is a linear equation in one variable, e.g.,

$$ax + b = c,$$

where x is the variable (unknown) and a, b and c are (known) constants. You have all seen these, and solved these before.[‡] As long as $a \neq 0$ the solution is given by

$$x = \frac{c - b}{a}.$$

If $a = 0$ then there is no variable, and the equation either reduces to a tautology or to a contradiction, depending on whether $b = c$ or $b \neq c$, respectively.

1.2 *Quadratic equations*

[†]The terminology ‘real’ number is unfortunate, because it often leads people to think that numbers that aren’t real don’t actually exist. This is *not* the case. Complex numbers exist in exactly the same way that real numbers exist — as mathematical abstractions that we use to describe the world around us.

[‡]If you haven’t, then you’re in the wrong course.

The quadratic equation in one variable is usually written like this

$$(1.2) \quad ax^2 + bx + c = 0,$$

where, as usual, we assume that a , b and c are real numbers.

The solutions, if they exist, can be found by *completing the square*

$$ax^2 + bx + c = a \left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} \right) - \frac{b^2}{4a} + c = a \left(x + \frac{b}{2a} \right)^2 + \frac{4ac - b^2}{4a}.$$

Plug the right hand side above into Equation (1.2)

$$a \left(x + \frac{b}{2a} \right)^2 + \frac{4ac - b^2}{4a} = 0,$$

and move things around to get

$$\left(x + \frac{b}{2a} \right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Now, take square roots on both sides and move things around again to obtain the **quadratic formula** for the solutions

$$(1.3) \quad x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

The expression $b^2 - 4ac$ is called the discriminant, and we have the following three possibilities:

- If $b^2 - 4ac > 0$ then Equation (1.2) has two distinct real solutions, given by the quadratic formula, (1.3).
- If $b^2 - 4ac = 0$ then Equation (1.2) has one real solution. That is the two solutions in (1.3) coincide.
- If $b^2 - 4ac < 0$ then there are **no** real solutions, and the two solutions given by the quadratic formula are *complex* numbers.

1.3 Factoring

Many polynomial equations,

$$a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0,$$

can be simplified or even completely solved by *factoring* the polynomial on the left into its simplest factors. Then we use a very basic and very important property of real (and complex) numbers.

Fact 1.

If A_1, A_2, \dots, A_n are real numbers and

$$A_1 \cdot A_2 \cdots A_n = 0$$

then

$$A_1 = 0 \quad \text{or} \quad A_2 = 0 \quad \text{or} \quad \dots \quad \text{or} \quad A_n = 0.$$

In words, *if the product of n real numbers is 0 then at least one of the numbers must be 0.*

Example 1. Solve the equation

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

This is a *cubic* equation, so the *quadratic* formula doesn't apply. To solve this we factor the polynomial into *linear* factors:

$$x^3 - x^2 - 4x + 4 = (x - 1)(x - 2)(x + 2).$$

Thus the original equation is the same as

$$(1.4) \quad (x - 1)(x - 2)(x + 2) = 0.$$

Now, in view of Fact 1, equation (1.4) implies that

$$x - 1 = 0 \quad \text{or} \quad x - 2 = 0 \quad \text{or} \quad x + 2 = 0,$$

so the solutions of Equation (1.4), and so the original equation are $x_1 = 1$, $x_2 = 2$ and $x_3 = -2$.

Example 2. Solve the equation

$$x^4 - 1 = 0.$$

The quartic polynomial on the left factors as follows

$$x^4 - 1 = (x^2 - 1)(x^2 + 1) = (x - 1)(x + 1)(x^2 + 1),$$

so the quartic equation above is equivalent to

$$(x - 1)(x + 1)(x^2 + 1) = 0.$$

There are only two real solutions in this case, $x_1 = 1$ and $x_2 = -1$, since the third factor, $x^2 + 1$, cannot be 0 if x is real.

Comment: Polynomials can sometimes be factored easily by trial and error. However, factoring

$$a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + a_1 x + a_0$$

is actually equivalent to solving the equation

$$a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0,$$

and is generally a hard problem. In practice, people who need to solve complicated polynomial equations often resort to clever approximation techniques or fancy computer programs to solve the equations.

2. Systems of equations

A system of equations is a collection of two or more equations, typically involving two or more variables. Generically, we can represent a system of equations like this

$$\begin{aligned} F_1(x_1, x_2, \dots, x_k) &= d_1 \\ F_2(x_1, x_2, \dots, x_k) &= d_2 \\ &\vdots \\ F_m(x_1, x_2, \dots, x_k) &= d_m, \end{aligned}$$

where F_1 through F_m are m algebraic expressions, d_1 through d_m are m constants, and x_1 through x_k are the variables. A **solution** of such a system is a k -tuple of numbers, $(x_1^*, x_2^*, \dots, x_k^*)$, that satisfy all of the m equations simultaneously.

For example

$$(2.1) \quad \begin{aligned} x^2 + 2xy - 2y^2 &= 1 \\ x + 3y &= 2 \end{aligned}$$

is a system of two equations in two variables, and

$$(2.2) \quad \begin{aligned} 2x + 3y + z &= 1 \\ x - 2y - 3z &= 2 \\ x + y + 2z &= 3 \end{aligned}$$

is a system of three equations in three variables.

There is no rule saying that the number of equations must equal the number of variables, but this is the case that we will encounter most frequently (perhaps exclusively) in this course.

2.1 Successive substitution and reduction

The most elementary approach to solving a system of equations is to successively solve each equation for one of the variables in terms of the remaining variables, and then substitute the expression you find in the remaining equations. In this way you successively reduce the number of equations and the number of variables, until you reach a system that is (we hope) easy to solve.

This is an *ad hoc* method, with no fixed set of rules. This is both a strength and a weakness. A strength, because the method works in principal for *any* system of equations. A weakness, because there is no fixed set of criteria that you can apply to determine whether or not you are on the right track before you reach the end. I'll illustrate the method with two simple examples.

Example 3. Consider the system (2.1) above. We can solve the second equation for x :

$$x = 2 - 3y$$

and substitute this into the first equation to obtain

$$x^2 + 2xy - 2y^2 = (2 - 3y)^2 + 2(2 - 3y)y - 2y^2 = 1,$$

which simplifies to

$$y^2 - 12y + 7 = 0.$$

This quadratic equation in y has two real solutions

$$y_1 = \frac{12 - \sqrt{144 - 28}}{2} = 6 - \sqrt{29} \approx 0.615 \text{ and } y_2 = \frac{12 + \sqrt{144 - 28}}{2} = 6 + \sqrt{29} \approx 11.385.$$

Corresponding to each y value there is an x value

$$x_1 = 2 - 3y_1 \approx 0.155 \text{ and } x_2 = 2 - 3y_2 \approx -32.155,$$

so the two solutions of the system are $(0.155, 0.615)$ and $(-32.155, 11.385)$.

Example 4. We can solve the system (2.2) in the same way. Solving the first equation for z gives

$$z = 1 - 2x - 3y.$$

Substitute this into both of the remaining equations reduces the original system of equations to a system of *two* equations in *two* variables,

$$\begin{aligned}x - 2y - 3z &= x - 2y - 3(1 - 2x - 3y) = 2 \\x + y + 2z &= x + y + 2(1 - 2x - 3y) = 3,\end{aligned}$$

which simplifies to

$$(2.3) \quad 7x + 7y = 5$$

$$(2.4) \quad -3x - 5y = 1.$$

Next, solve (2.3) for y

$$y = \frac{5}{7} - x,$$

and substitute this into (2.4)

$$-3x - 5\left(\frac{5}{7} - x\right) = 2x - \frac{25}{7} = 1,$$

giving

$$x = \frac{16}{7}, \quad y = \frac{5}{7} - \frac{16}{7} = -\frac{11}{7} \text{ and } z = 1 - \frac{32}{7} + \frac{33}{7} = \frac{8}{7}.$$

In this case there is exactly one solution $\left(\frac{16}{7}, -\frac{11}{7}, \frac{8}{7}\right)$.

Comment: When you use this method you can start with any of the equations and solve for any of the variables. Some choices may make your work easier and others may make your work more difficult. The more you practice, the more obvious the ‘good’ choices will be.

3. Systems of linear equations

An equation of the form

$$c_1x_1 + c_2x_2 + \cdots + c_kx_k = b$$

is called a **linear** equation in the variables x_1, \dots, x_k . The constants c_1, \dots, c_k and b are called the *coefficients* of the equation.

When all of the equations in a system of equations are linear, there is a rich theory that describes the solutions of the system as well as methods for finding them. The field that studies linear systems (and related problems) is called **linear algebra**. One of the features of linear systems is the following fact about the set of solutions to such a system.

Fact 2.

Any system of linear equations with real coefficients has exactly one solution, infinitely many solutions or no solutions.

The method of successive substitution and reduction works for linear systems of equations, but it is not the most efficient method.[§] The most efficient method for the solution of linear systems is called *Gaussian elimination*.

[§]One method of solution is more efficient than another if it requires fewer steps and fewer arithmetic operations.

Two systems are called *equivalent* if they have the same solutions. The idea of Gaussian elimination is to combine (and rearrange) the equations to replace the original system with an equivalent system of equations that is easier to solve. What makes the new system simpler than the original system is that most of the variables have been *eliminated* from most of the equations. Rather than describe the method in detail, I will illustrate it with an example.

Example 5. Consider the system of equations (2.2) again. We already solved it using substitution, now we'll try elimination. The idea is to add multiples of one equation to the others to produce new equations with fewer variables. There is more than one way to proceed, as with substitution, and the more you practice, the more you will develop an eye for which way is quickest.

We begin with the original system:

$$\begin{aligned} 2x + 3y + z &= 1 \\ x - 2y - 3z &= 2 \\ x + y + 2z &= 3 \end{aligned}$$

First, we subtract the third equation from the second equation, and subtract 2 times the third equation from the first equation. After swapping the third equation with the (new) first equation, we have the system

$$\begin{aligned} x + y + 2z &= 3 \\ -3y - 5z &= -1 \\ y - 3z &= -5 \end{aligned}$$

Note that the variable x has been eliminated from two of the equations. Next, add 3 times the (new) third equation to the second equation, then swap the second and third equations to obtain the system

$$\begin{aligned} x + y + 2z &= 3 \\ y - 3z &= -5 \\ -14z &= -16 \end{aligned}$$

Now we can solve for z directly in the third equation (since both y and x have been eliminated from that equation) giving

$$z_0 = \frac{-16}{-14} = \frac{8}{7}.$$

Plugging z_0 in the second equation, and solving for y we have

$$y_0 = -5 + 3z_0 = -5 + \frac{24}{7} = \frac{-35 + 24}{7} = -\frac{11}{7}.$$

finally, substituting z_0 and y_0 into the first equation and solving for x , we have

$$x_0 = 3 - y_0 - 2z_0 = 3 - \left(-\frac{11}{7}\right) - \frac{16}{7} = \frac{21 + 11 - 16}{7} = \frac{16}{7}.$$

The solution is no surprise — we already solved this system using the substitution method. The point of this example is to illustrate a better method for dealing with linear systems of equations.