Some notes on equivalence relations

Ernie Croot

January 23, 2012

1 Introduction

Certain abstract mathematical constructs get defined because they are useful in unifying and making sense of a large number of seemlingly unrelated concepts. The notion of an 'equivalence relation' is one such construct, as it unifies the notion of the equality of numbers, congruent and similar triangles, similar matrices, congruence of numbers, and many, many other concepts.

Let's start with congruence of triangles, which is a type of equivalence relation: in high school you learned the definition of *congruence*, which said that two triangles ABC and DEF (the letters are labels of the vertices) in the plane are *congruent*, and you write $ABC \cong DEF$, if the lengths of the corresponding sides and the sizes of the corresponding angles of these triangles are equal. Without knowing it, you learned a type of *equivalence relation*. Congruence is an equivalence relation. Here are the three properties that make it an equivalence relation:

1. Reflexive. A triangle ABC is congruent to itself; that is, $ABC \cong ABC$.

2. Symmetric. If $ABC \cong DEF$, then $DEF \cong ABC$.

3. Transitive. If $ABC \cong DEF$ and $DEF \cong GHI$, then $ABC \cong GHI$.

Equality of real numbers is another example of an equivalence relation. Here, rather than working with triangles we work with numbers: we say that the real numbers x and y are *equivalent* if we simply have that $x = y$. Obviously, then, we will have that:

1. **Reflexive.** $x = x$.

2. **Symmetric.** If $x = y$, then $y = x$.

3. **Transitive.** If $x = y$ and $y = z$, then $x = z$.

2 On the need for formal definitions

Mathematicians are nuts about *formal* definitions, which are more abstract and specific than the way you have seen terms defined up till now. One reason for this desire for formality is that it allows them to see what the key issues are underly a phenomenon, idea, or pattern, so that it can be more easily generalized to larger contexts. A second reason is that working purely formally sometimes uncovers hidden analogies that can be used to link different fields together. And a third reason is that sometimes we can become confused by our language, and not realize when we are making logical errors. Here are two examples to illustrate this third point:

Example 1. Let N be the largest number that cannot be defined in 20 words or less.

And now here is a "problem" for you to ponder: it appears as though that sentence just defined, in 20 words or less, a number that can't be defined in 20 words or less! So it seems we have a connundrum on our hands.

Example 2. A teacher announces to her class that there will be a surprise exam next week. On hearing this, one of the students reasons that this is impossible, using the following logic: if there is no exam by Thursday, then it would have to occur on Friday; and by Thursday night the class would know this, making it not a surprise. So, the exam could not be on Friday. Similarly, one can eliminate Thursday as a possibility, since if there were no exam by Wednesday, then on Wednesday night the class would know it has to be on Thursday (Friday having been already eliminated as a possibility). Continuing in this manner, one can elminate every day of the week as a possibilty for when to have the "surprise" exam... and so the exam promptly vanishes in a puff of logic.

3 The formal definition of an equivalence relation

After that digression, we are now ready to state the formal definition of an equivalence relation: given a non-empty set $\mathcal{U}$, we say that $E \subseteq \mathcal{U} \times \mathcal{U}$ is an *equivalence relation* if it has the following properties: [1]

1. **Reflexive.** If $x \in \mathcal{U}$, then $(x, x) \in E$.

2. **Symmetric.** If $(x, y) \in E$, then $(y, x) \in E$.

3. **Transitive.** If $(x, y) \in E$ and if $(y, z) \in E$, then $(x, z) \in E$.

In the case of the congruent triangles example above, the set $\mathcal{U}$ is the set of all triangles in the plane, and the variables x, y and z represent triangles. And in the case of the real numbers example, the set $\mathcal{U}$ is $\mathbb{R}$ and x, y and z are real numbers.

Now, it may seem a little weird using set language like that, and there are uses for working with sets, though mostly for reasons stemming from mathematical logic. Perhaps a more natural way to define an equivalence relation abstractly is to use the "twiddle (tilde) definition": given a non-empty set A, we say that $\sim$ is an equivalence relation if the following all hold:

1. **Reflexive.** If $x \in A$, then $x \sim x$.

2. **Symmetric.** If $x \sim y$, then $y \sim x$.

3. **Transitive.** If $x \sim y$ and $y \sim z$, then $x \sim z$.

In the case of the congruent triangles, $\sim$ takes the place of $\cong$, and A is the set of triangles in the plane. And in the real numbers example, $\sim$ is just the equals symbol $=$ and A is the set of real numbers.

4 Some further examples

Let us see a few more examples of equivalence relations. All the proofs will make use of the $\sim$ definition above:

[1] The notation $\mathcal{U} \times \mathcal{U}$ means the set of all ordered pairs (x, y), where x, y belong to $\mathcal{U}$.

4.1 Example 1

This example comes from number theory: fix a non-zero integer d. We say that $x \sim y$ (more properly we should perhaps use the notation $x \sim_d y$ to indicate that we are going to produce infinitely many equivalence relations, one for each choice of d), and we write $x \equiv y \pmod d$, if the difference $x - y$ is evenly divisible by the number d.

For example, $1 \equiv 16 \pmod 5$, because $1 - 16 = -15$ is evenly divisible by 5; $-17 \equiv 5 \pmod{11}$, because $-17 - 5 = -22$ is evenly divisible by 11; but, $5 \not\equiv 13 \pmod{10}$, since $5 - 13 = -8$ is *not* evenly divisible by 10.

Before we show that this gives an equivalence relation, it is a good idea to express the notion of being ”evenly divisible” in a way that is more accessible to us: we say that an integer x is evenly divisible by a non-zero integer d if there exists an integer k, such that

$$x \;=\; d \cdot k.$$

For example, 12 is evenly divisible by -3, since

$$12 \;=\; (-3) \cdot (-4),$$

so that $k = -4$ in this example.

Ok, so now let us tackle the problem of showing that $\sim$ is an equivalence relation: (remember... we assume that d is some fixed non-zero integer in our verification below) Our set A in this case will be the set of integers $\mathbb{Z}$.

1. Reflexive. If $x \in \mathbb{Z}$, then $x \sim x$, since $x - x = 0 = 0 \cdot 0$ (here $k = 0$).

2. Symmetric. If $x \sim y$, then $x - y = d \cdot k$, for some integer k. But this means that $y - x = d \cdot (-k)$; and since $-k$ is *also* an integer, this means that $y \sim x$.

3. Transitive. If $x \sim y$ and $y \sim z$, then we have that there exist integers k and ℓ such that

$$x - y \;=\; d \cdot k \quad \text{and} \quad y - z \;=\; d \cdot \ell.$$

Adding these together gives

$$x - z \;=\; (x - y) + (y - z) \;=\; d \cdot k + d \cdot \ell \;=\; d(k + \ell).$$

Since k and ℓ were integers, so must be $k + \ell$; and so, we deduce that $x - z$ is evenly divisible by d, resulting in $x \sim z$.

4.2 Example 2

The second example comes from what are called "similar matrices": fix an integer $n \geq 1$. Given two $n \times n$ matrices M and N, we say that M is *similar* to N if there exists an $n \times n$ invertible matrix L satisfying

$$M \;=\; L^{-1}NL.$$

Suppose we just restrict ourselves to 2×2 matrices; that is, we only consider the case $n = 2$. So, we will let A be the set of all 2×2 matrices, and we will use the notation $M \sim N$ to mean that the matrix M is similar to the matrix N. Now we do the verification:

Reflexive. Letting I denote the 2×2 identity matrix, we observe that since $M = I^{-1}MI$, we must have $M \sim M$.

Symmetric. Suppose $M \sim N$. Then, by definition, there exists an invertible matrix L satisfying $M = L^{-1}NL$. But this means that $N = (L^{-1})^{-1}M(L^{-1})$; and therefore $N \sim M$.

Transitive. Suppose that $M \sim N$ and $N \sim P$. Thus, there exist invertible matrices L and K such that $M = L^{-1}NL$ and $N = K^{-1}PK$; and so,

$$M \;=\; L^{-1}NL \;=\; L^{-1}(K^{-1}PK)L \;=\; (L^{-1}K^{-1})P(KL) \;=\; (KL)^{-1}P(KL), \tag{1}$$

which implies that $M \sim P$.

In the chain of equalities in (1) we used several facts about matrices: for the third equality we used associativity of matrix multiplication; and for the last equality we used the fact that $(KL)^{-1} = L^{-1}K^{-1}$ (this is what I called the "shoes and socks principle" in class).

Since we know that if $M \sim N$ then $N \sim M$, it makes sense to say "M and N are similar matrices", without specifying whether you are talking about $M \sim N$ or $N \sim M$.

5 Equivalence classes

Once you have an equivalence relation on a set A, you can use that relation to decompose A into what are called *equivalence classes*: given an element

$x \in A$, let A_x be the set of all elements of A that are equivalent to x; that is, A_x is the set of all $y \in A$ such that $y \sim x$ (or $x \sim y$). We say that this set A_x is the *equivalence class of* x.

For example, in the case of congruence mod d above, if we were to fix $d = 5$, say, then A_7 would be the set of all numbers

$$\{..., -13, -8, -3, 2, 7, 12, 17, 22, 27, 32, ...\};$$

basically, all numbers of the form $5k + 7$, where k ranges over the integers. All these numbers are equivalent to 7; that is, $-13 \sim 7$, $-8 \sim 7$, etc.

A very useful fact about equivalence classes is that they can be used to *partition* the set A. What does that mean? Well, a collection of sets $S_1, S_2, S_3, ...$ is said to form a *partition* of a set S if

1. The sets S_i are non-empty.

2. They are disjoint – meaning that $S_i \cap S_j = \emptyset$ for all $i \neq j$ (i.e. the sets S_i and S_j have no elements in common).

3. And, S is the union of these sets S_i; that is, $S = S_1 \cup S_2 \cup \cdots$.

The sequence of sets $S_1, S_2, ...$ here could be finite or infinite, depending on the situtation at hand.

Let us see that the case of integer congruence gives us a partition of $A = \mathbb{Z}$ for $d = 5$: consider, for instance, the sets

$$\begin{aligned}
A_1 &= \{..., -9, -4, 1, 6, 11, 16, 21, ...\} \\
A_2 &= \{..., -13, -8, -3, 2, 7, 12, 17, ...\} \\
A_3 &= \{..., -12, -7, -2, 3, 8, 13, 18, ...\} \\
A_4 &= \{..., -11, -6, -1, 4, 9, 14, 19, ...\} \\
A_5 &= \{..., -10, -5, 0, 5, 10, 15, ...\}
\end{aligned}$$

First notice (by inspection) that all these sets are non-empty and no pair of sets can have an element in common; this establishes the first two parts of the definition of a partition. Also notice that $A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$ contains all the numbers from -13 to 19; and in fact, it contains all the integers $A = \mathbb{Z}$ (takes a little work to see). So, these sets form a partition of $\mathbb{Z}$.

And, any other partition of $\mathbb{Z}$ using $d = 5$ and congruence *must* use these same sets. The reason for this is that the equivalence class A_x is the same as the class A_{x+5} – e.g.

$$\cdots \ = \ A_{-3} \ = \ A_2 \ = \ A_7 \ = \ A_{12} \ = \ \cdots$$

6 More on partitions

Not only do equivalence classes give a natural partition of a given set A, but you can *form* an equivalence relation by starting with a partition. It is a little artificial, but it does mean that:

> Equivalence relations on a set A are in 1-1 correspondance with partitions of A

And now you'll see why I used the word 'artificial' above: starting with a partition of a set A as

$$A \ = \ S_1 \cup S_2 \cup \cdots$$

we form an equivalence relation $\sim$, such that $x \sim y$ whenever x and y belong to the same set S_i.

For example, suppose my set $A := \{1, 2, 3, 4, 5, 6, 7, 8\}$, and suppose I just arbitrarily partition A into $S_1 \cup S_2 \cup S_3$, where I take

$$S_1 \ := \ \{1, 7, 8\}, \ S_2 \ := \ \{2, 3, 4\}, \ \text{and } S_3 \ := \ \{5, 6\}.$$

Then, I have the following relations:

$$\begin{aligned}
1 &\sim 7 \\
7 &\sim 1 \\
1 &\sim 8 \\
8 &\sim 1 \\
7 &\sim 8 \\
8 &\sim 7 \\
2 &\sim 3 \\
3 &\sim 2
\end{aligned}$$

$$
\begin{aligned}
2 &\sim 4 \\
3 &\sim 4 \\
4 &\sim 3 \\
5 &\sim 6 \\
6 &\sim 5.
\end{aligned}
$$